Technische Universität München
Fakultät für Medizin

Evidence-based pharmacotherapy and drug development in autism spectrum disorder: a systematic review, network meta-analysis and meta-regression of placebo-effects

Spyridon Siafis

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung eines

Doktors der Medizin (Dr. med.)

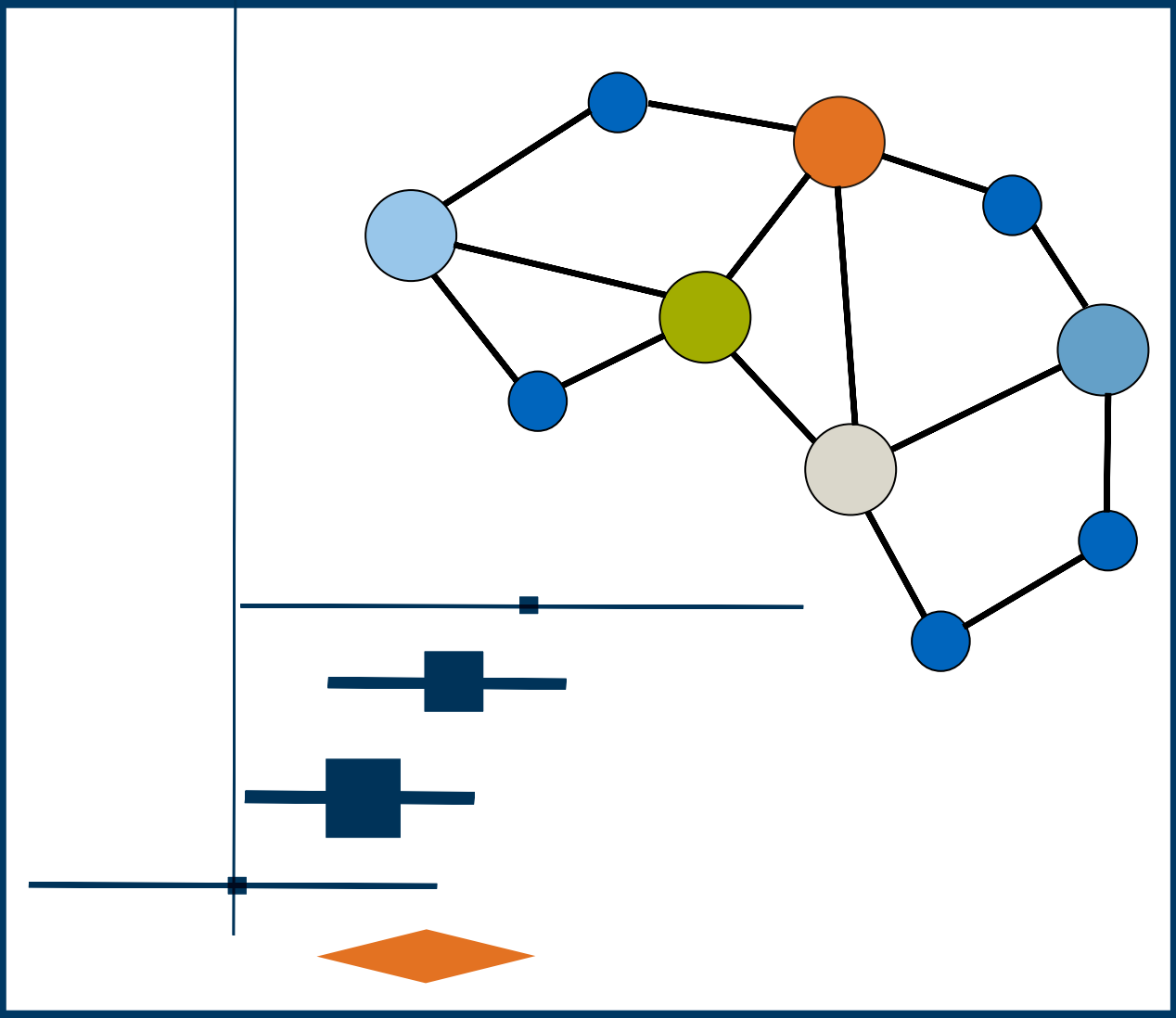
genehmigten Dissertation.

Vorsitz: Prof. Dr. Lars Mägdefessel

Prüfer*innen der Dissertation:

1. apl. Prof. Dr. Stefan M. Leucht
2. Priv.-Doz. Dr. Andrea Welling
3. apl. Prof. Dr. Andreas Dinkel

Die Dissertation wurde am 05.08.2022 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 03.01.2023 angenommen.



Autism spectrum disorder from evidence-based pharmacotherapy to drug development

Spyridon Siafis

School of Medicine, Department of Psychiatry and Psychotherapy

Section for Evidence-Based Medicine in Psychiatry and Psychotherapy

Evidence-based pharmacotherapy and drug development in autism spectrum disorder: a systematic review, network meta-analysis and meta-regression of placebo-effects

Spyridon Siafis

Doctoral degree in human medicine (Dr. med.)

“Σὰ βγεῖς στὸν πηγαιμὸ γιὰ τὴν Ἰθάκη,

νὰ εὔχεσαι νἆναι μακρὺς ὁ δρόμος,

γεμάτος περιπέτειες, γεμάτος γνώσεις.

……..

Πάντα στὸ νοῦ σου νἄχῃς τὴν Ἰθάκη.

Τὸ φθάσιμον ἐκεῖ εἶν’ ὁ προορισμός σου.

……..

Ἀλλὰ μὴ βιάζῃς τὸ ταξείδι διόλου.

Καλλίτερα χρόνια πολλὰ νὰ διαρκέσει.”

Κ.Π. Καβάφης, Ἰθάκη, 1911

Abstract

Previous late-stage randomized controlled trials (RCTs) failed to identify effective medications for the core symptoms of autism spectrum disorder (ASD), i.e., social-communication difficulties and repetitive behaviors, and as a consequence, there is still no approved medication. Thus, I conducted a systematic review of 203 RCTs with 12111 participants in order to inform evidence-based pharmacotherapy and drug development in ASD.

The thesis consists of three parts:

First, I conducted a network meta-analysis to investigate the efficacy and tolerability of pharmacological and dietary-supplement treatments. Some medications, e.g., the antipsychotics aripiprazole and risperidone, might be effective for the core symptoms and/or co-occurring difficulties (e.g., irritability), albeit associated with adverse events. Nevertheless, the evidence was generally preliminary and with low certainty. Therefore, routine prescription of medications for the core symptoms cannot be recommended and further investigation is warranted.

Second, I conducted a meta-analysis of placebo-effects. The magnitude of placebo-effects was considerable and predictors of higher placebo-effects were identified, e.g., caregiver-ratings and larger trials. However, there were limited and scattered data for participant-level factors, e.g., age, sex, and baseline severity of symptoms.

Third, I validated an imputation method to estimate the number of responders from continuous data of the Clinical Global Impression Improvement (CGI-I) scale. This method could facilitate the comparability and combination of findings across RCTs in ASD. However, sensitivity analyses are necessary given the relatively wide limits of agreement between imputed and original values.

The findings and future implications of my thesis would hopefully facilitate better support and care for individuals with ASD and their families.

Zusammenfassung

Es gibt noch keine offiziell zugelassenen Medikamente für die Kernsymptome von Autismus-Spektrum-Störungen (ASD), d. h. Schwierigkeiten in der sozialen Interaktion und Kommunikation sowie repetitive Verhaltensmuster. Daher habe ich eine systematische Überprüfung von 203 randomisiert-kontrollierten Studien (RCTs) mit 12111 Teilnehmern durchgeführt, mit dem Ziel eine evidenzbasierte Arzneimittelentwicklung und Pharmakotherapie bei ASD zu unterstützen.

Die Arbeit umfasst drei Teile:

Erstens führte ich eine Netzwerk-Metaanalyse über die Wirksamkeit und Verträglichkeit von pharmakologischen Behandlungen und Nahrungsergänzungsmitteln bei ASD durch. Einige Medikamente, z.B. die Antipsychotika Aripiprazol und Risperidon, waren gegen die Kernsymptome und/oder die damit einhergehende Begleitsymptome (z.B. Aggressivität) Placebo überlegen, aber sie verursachten auch Nebenwirkungen. Ferner waren die Ergebnisse vorläufig und noch wenig verläßlich. Daher kann die routinemäßige Verschreibung von Medikamenten gegen die Kernsymptome aktuell noch nicht empfohlen werden und weitere Untersuchungen sind erforderlich.

Zweitens habe ich eine Metaanalyse der Placeboeffekte (Ansprechen auf Placebo) durchgeführt. Das Ausmaß der Placeboeffekte war beträchtlich, und ich konnte verschiedene Prädiktoren für höhere Placeboeffekte identifizieren. Insbesondere waren die Placeboeffekte größer, wenn Eltern- anstatt Behandlerfragebögen eingesetzt wurden und wenn die Studien größer waren. Für patientenbezogene Faktoren gab es nur begrenzte Daten, wie Alter, Geschlecht und Schweregrad der Symptome bei Studienbeginn.

Drittens validierte ich eine Imputationsmethode zur Schätzung der Anzahl der Responder aus kontinuierlichen Daten der Clinical Global Impression of Improvement (CGI-I) Skala. Diese Methode kann die Vergleichbarkeit und Kombination von Ergebnissen verschiedener RCTs bei ASD erleichtern. Allerdings sind bei ihrer Anwendung Sensitivitätsanalysen erforderlich, da die Grenzen der Übereinstimmung zwischen den geschätzten und den ursprünglichen Werten relativ groß sind.

Die Ergebnisse und die sich aus ihnen ergebenden Implikationen meiner Dissertation werden hoffentlich einen Beitrag zur Behandlung von Menschen mit ASD und ihren Familien leisten.

Publications included in the thesis

The publications included in the thesis are listed according to the order presented in the thesis, and not the chronological order.

Publication I: Network meta-analysis on the effects of pharmacological and dietary-supplement treatments for autism spectrum disorder

Siafis, S., Çıray, O., Wu, H., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., San José Cáceres, A., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2022. 'Pharmacological and dietary-supplement treatments for autism spectrum disorder: a systematic review and network meta-analysis', *Mol Autism*, 13: 10. doi: 10.1186/s13229-022-00488-4.

Publication II: Meta-analysis of placebo-effects in the core symptoms of autism spectrum disorder

Siafis, S., Çıray, O., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2020. 'Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis', *Mol Autism*, 11: 66. doi: 10.1186/s13229-020-00372-z.

Publication III: Validation of an imputation method to the estimate the number of responders using the mean and standard deviation of the Clinical Global Impression Improvement scale in autism spectrum disorder

Siafis, S., Rodolico, A., Çıray, O., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2021. 'Imputing the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder', *Brain Sci*, 11. doi: 10.3390/brainsci11070908.

Table of contents

List of abbreviations

5-$HT_{1A/2A}$: serotonin 1A, and 2A receptors

ABC: Aberrant Behavior Checklist

ABC-H: Aberrant Behavior Checklist - Hyperactivity/noncompliance

ABC-I: Aberrant Behavior Checklist - Irritability

ABC-L/SW: Aberrant Behavior Checklist - Lethargy/Social Withdrawal

ADHD: Attention deficit and hyperactivity disorder

ALDH5A1: Aldehyde dehydrogenase 5 family member A1

alpha-2: Norepinephrine alpha-2 receptor

ALS: Amyotrophic lateral sclerosis

APA: American Psychiatric Association

ASD: Autism Spectrum Disorder

BPS: British Pharmacological Society

CARS: Childhood Autism Rating Scale

CASI: Child and Adolescent Symptom Inventory

CBD: Cannabidiol

CBDV: Cannabidivarin

CBT: Cognitive behavioral therapy

CCC: Concordance correlation coefficient

CENTRAL: Cochrane Central Register of Controlled Trials

CGI-I: Clinical Global Impression - Improvement

CGI-S: Clinical Global Impression - Severity

CINeMA: Confidence in Network Meta-analysis

CY-BOCS-PDD: Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders

D_2: Dopamine 2 receptor

DALY: Disability-Adjusted Life Years

DAT: Dopamine transporter

DD-CGAS: Developmental Disabilities Modification of Children's Global Assessment Scale

DSM-5: Diagnostic and Statical Manual of Mental Disorders version 5

EBM: Evidence-Based Medicine

EMA: European Medicines Agency

FDA: U.S. Food and Drug Administration

FFAR: Free fatty acid receptor

GBD: Global Burden of Disease

GRADE: Grading of Recommendations Assessment, Development and Evaluation

GSK-3b: Glycogen synthase kinase 3 beta

HDAC: Histone deacetylase

ICD-11: International Classification of Diseases version 11

ICTRP: International Clinical Trials Registry Platform

ITT: Intention-to-treat

IUPHAR: International Union of Basic and Clinical Pharmacology

LOCF: Last-observation carried forward

m: Mean score

MCID: Minimum clinically important difference

MDMA: 3,4-methylenedioxymethamphetamine

M_{drug}: Mean improvement of symptoms in the medication group

MMRM: Mixed-models of repeated measurement

$M_{placebo}$: Mean improvement of symptoms in the placebo group

MRS: Magnetic resonance spectroscopy

MSE: Mean squared error

MT: Melatonin receptor

n: Number of participants

Nav: Voltage gated sodium channel

NbN: Neuroscience-based Nomenclature

NET: Norepinephrine transporter

NIMH-RUPP:National Institute of Mental Health (NIMH) Research Units on Pediatric Psychopharmacology

NMDA: N-methyl-D-aspartate receptor

OACIS: Ohio Autism Clinical Impressions Scale

OCD: Obsessive-compulsive disorder

OR: Odds ratio

OT: Oxytocin receptor

PedsQL: Pediatric Quality of Life Inventory

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PSI: Parental Stress Index

r: Pre and post correlation or correlation among scales

RB: Repetitive Behaviors

RCT: Randomized controlled trial

RXRa: Retinoid X receptor a

SCD: Social-communication difficulties

SD: Standard Deviation

SERT: Serotonin transporter

SMC: Standardized mean change

SMD: Standardized mean difference

SRS: Social Responsiveness Scale

SSRI: Selective serotonin reuptake inhibitors

SUCRA: Surface under the cumulative ranking curve

THC: Δ9-tetrahydrocannabinol

TRP: Transient receptor potential cation channel

VP : Vasopressin receptor

vs. : Versus

WHO: World Health Organization

ρ: Spearman's rho correlation

List of tables and figures

1. Introduction

1.1. Autism spectrum disorder

Autism spectrum disorder (ASD) is a group of common (about 1-2% of the population) and early-emerging neurodevelopmental conditions that are characterized by difficulties in social communication and interaction as well as by restricted interests, repetitive behaviors and/or sensory abnormalities (APA 1980; Lord et al. 2022). The aetiology is multifactorial that lies on the genetic structure, with a hereditability of about 50-60% explained mainly by common variations (Gaugler et al. 2014), and environmental interactions such as prenatal exposures (Kim et al. 2019). Its clinical manifestation is a cluster of heterogeneities in terms of neurobiology, severity of symptoms, presence of co-occurring conditions, and outcome trajectories. Accordingly, the care needed for people with ASD across the lifespan varies, which is also reflected by differences across families, regions and cultures (Lord et al. 2022).

ASD has also a dual nature consisting of a medical condition that may be associated with disabilities due to atypical neurodevelopment (medical viewpoint), and a form of human variation that may be associated with difficulties because of misfit with a particular environment (neurodiversity viewpoint) (Lai et al. 2020). Therefore, a holistic framework that combines the medical and neurodiversity viewpoints has been proposed for the multidisciplinary and collaborative support and care for individuals with ASD (Lai et al. 2020). The thesis examined the evidence on pharmacotherapy, and therefore, it was eventually focused on the medical viewpoint.

In addition, no term describing autism is universally accepted by the autism community (Kenny et al. 2016). The thesis conventionally used the term “autism spectrum disorder”, which is the formal diagnostic term of the Diagnostic and Statical Manual of Mental Disorders version 5 (DSM-5)(APA 1980) and the International Classification of

Diseases version 11 (ICD-11)(WHO 2018). However, it should be noted that this term was endorsed to a lesser degree by the autism community in comparison to “autism” or “on the autism spectrum” (Kenny et al. 2016).

1.2. Burden, support and care

ASD is a lifelong condition that accounts globally for a substantial burden of disease with 53-58 disability-adjusted life years (DALY) per 100,000 (Baxter et al. 2015). In Germany, the WHO global burden disease report of 2019 found that ASD was ranked 12th and 16th in terms of DALYs for individuals <5 years old (104 DALYs per 100,000) and 5-14 years old (101 DALYS per 100,000)(GBD 2019 Dieases and Injuries Collaborators 2020). ASD is also found to be associated with a higher risk of mortality both due to natural and unnatural causes (Catalá-López et al. 2022). There are also considerable direct and indirect costs for individuals with ASD and their families, e.g., due to healthcare services, loss of productivity and education (Rogge and Janssen 2019).

Therefore, there is a tremendous need to improve the support for individuals with ASD in order to relieve distress, improve adaption, and quality of life (Lai et al. 2020). The required support can vary substantially at the individual level, given the heterogeneity in ASD, and thus, a shared-decision making process among individuals with ASD, their families and service users is required to take into account the preferences and complex needs of the individuals and their families (Lai et al. 2020; Lord et al. 2022). Accordingly, current frameworks propose a personalized, multilateral and collaborative support and care network to maximize the potential of individuals with ASD by facilitating the development of skills, identifying and minimizing barriers and optimizing the person-environment fit (Lai et al. 2020; Lord et al. 2022). To that direction, there is a large number of different modalities of interventions, such as

behavioral, developmental, educational, psychosocial, vocational, lifestyle, dietary, pharmacological interventions, as well as advocacy and policy changes (Hyman et al. 2020; Lord et al. 2022). The thesis was focused on pharmacotherapy, and thus, a comprehensive description of other modalities of interventions is out of the scope, e.g., see other reviews (Ameis et al. 2018; Fuentes, Hervás, and Howlin 2021; Howes et al. 2018; Hyman et al. 2020; Lai et al. 2020; Lord et al. 2020; Lord et al. 2022; Lord et al. 2018).

1.3. Pharmacological and dietary-supplement treatments for ASD

1.3.1. Current status in clinical practice

1.3.1.1. Pharmacological treatments

There is yet no approved medication for the core symptoms of ASD, i.e., social-communication difficulties (SCD) and restricted, repetitive behaviors (RB) (Howes et al. 2018).

Some medications, however, were found efficacious for the treatment of co-occurring difficulties and conditions, i.e., aripiprazole, haloperidol, and risperidone for severe irritability and aggression (with a specific indication by the U.S. Food and Drug Administration, FDA, and/or European Medicines Agency, EMA), methylphenidate, clonidine and guanfacine for attention deficit and hyperactivity disorder (ADHD), and melatonin for sleep disorders (with a specific indication by EMA) (Fuentes, Hervás, and Howlin 2021; Howes et al. 2018; Rodrigues et al. 2021).

In addition, pharmacological treatments could also be useful for other co-occurring conditions following the guidelines for the general population, e.g., antidepressants for anxiety or depression, antiepileptics for epilepsy, albeit specific evidence when these conditions co-occur with ASD is missing (Fuentes, Hervás, and Howlin 2021; Howes et al. 2018).

However, due to the risk of side-effects, medications should be generally considered on a case-by-case basis and as second-line treatments when behavioral interventions have failed, as well as after evaluating and monitoring their benefit-risk trade-offs (Ameis et al. 2018; Howes et al. 2018; Rodrigues et al. 2021).

Nevertheless, psychotropic medications are often used in clinical practice. A systematic review found that about half of the individuals with ASD use at least one medication (median prevalence of 45.7%) and about a fifth use more than one medication (23.0%) (Jobski et al. 2017). The most commonly used medication classes were antipsychotics (median prevalence of 18.1%), antidepressants (17.2%), ADHD medication (16.6%), mood stabilizers or antiepileptics and sleep medications, reflecting the use of psychotropics for co-occurring difficulties (Jobski et al. 2017). In addition, the prevalence and patterns of use seem to vary importantly because of differences in age, sex, presence of co-occurring conditions and regions (Jobski et al. 2017).

1.3.1.2. Dietary-supplements

There is no conclusive evidence about the use of dietary-supplements for core or associated symptoms of ASD (Ameis et al. 2018; Fuentes, Hervás, and Howlin 2021; Lord et al. 2022). Their use is generally not recommended, except for the correction of nutritional deficiencies, since feeding problems could also be more often in children with ASD (Fuentes, Hervás, and Howlin 2021; Hyman et al. 2020; Sharp et al. 2013). Nevertheless, and given their general safety, easy accessibility and availability, about half of the individuals with ASD use dietary-supplements (Höfer, Hoffmann, and Bachmann 2017), most frequently multivitamins, vitamin-D, omega-3, probiotics and magnesium (Trudeau et al. 2019).

1.3.2. Current status in clinical research

Clinical psychopharmacology in ASD is an emerging topic as reflected by a recent increase of randomized controlled trials (RCTs). For example, a survey of clinical trials in ASD found 257 RCTs investigating pharmacological treatments, most of them published after 2008 (Tromans and Adams 2018).

There are currently two main approaches that guide psychopharmacology in ASD, 1) re-purposing medications for other conditions with overlapping symptoms, e.g., antipsychotics for aggression, antidepressants for obsessive and compulsive disorder (OCD), and 2) developing medications that may target the proposed underlying neurobiological processes, e.g., excitatory-inhibitory imbalance, neuroinflammation, serotoninergic, oxytocinergic, endocannabinoid neurotransmission (Anagnostou 2018; Díaz-Caneja et al. 2021; Howes et al. 2018).

Therefore, a plethora of medications and dietary-supplement treatments with diverse mechanisms of action have already been or are being investigated, e.g., see other reviews for proposed neurobiological processes and mechanisms of actions of investigated medications (Baribeau, Vorstman, and Anagnostou 2022; Díaz-Caneja et al. 2021; McCracken et al. 2021; Persico et al. 2021).

However, and despite the recent advances in preclinical and translational research, late-stage RCTs have failed so far to identify efficacious medications for the core symptoms (Berry-Kravis et al. 2018; Díaz-Caneja et al. 2021; Jacob et al. 2022; McCracken et al. 2021).

There could be many reasons for the low success of RCTs in ASD such as 1) a true lack of efficacy of the investigated medication, e.g., poor translational validity of preclinical models, inappropriate dosing and administration, 2) heterogeneity and interpersonal variability of treatment response, and 3) suboptimal trial design, e.g.,

inadequate randomization or blinding, lack of appropriate outcome measures, small sample sizes, short treatment duration, substantial placebo-effects etc. (Berry-Kravis et al. 2018; Díaz-Caneja et al. 2021; Gribkoff and Kaczmarek 2017; Jacob et al. 2022; Lord et al. 2022; McCracken et al. 2021).

1.4. Randomized controlled trials, meta-analysis and placebo-effects

1.4.1. Randomized controlled trials

RCTs are considered the “gold standard” for the evaluation of treatment effects according to the traditional hierarchy of evidence (Djulbegovic and Guyatt 2017). RCTs are prospective clinical trials, in which participants are randomized to an experimental intervention (e.g., a medication) vs. a control intervention (e.g., placebo or inert substance, sham treatment), and usually without the participants, the study personnel and outcome assessors being aware of the treatment assignment (i.e., double-blind RCT).

1.4.2. Systematic reviews and meta-analysis

There are often more than one RCTs investigating a comparison between an experimental and control intervention, e.g., in order to replicate findings, test in different settings, or conducted during different phases of drug development, etc. Their results could be different, and sometimes even conflicting. For this reason, systematic reviews and meta-analysis aim to synthesize the evidence from all available studies (e.g., RCTs) for a research question (Higgins et al. 2019).

The basic methodology of systematic reviews consists of designing and registering the protocol, searching and selecting all available studies, assessing their risk of bias and extracting data, statistically combining data with meta-analytic methods and evaluating the certainty in the evidence (Higgins et al. 2019).

In particular, meta-analyses aim to synthesize effect-sizes across studies. Effect-sizes are statistical measures that quantify the magnitude of a phenomenon under investigation (Kelley and Preacher 2012), e.g., the magnitude of change from baseline after treatment with an intervention or the magnitude of difference between two interventions. Therefore, there are different meta-analytic methods for different research questions, e.g., pairwise meta-analysis (combining the results across studies for a specific comparison), network meta-analysis (combining direct and indirect evidence for three or more interventions across a network of studies), and single-group meta-analysis (e.g., proportions or pre-post changes).

1.4.3. Placebo-effects

An important component of the rigorous design of RCTs is that they could control for placebo-effects, i.e., improvements observed in participants treated with placebo. Placebo-effects consist of unspecific improvements, e.g., spontaneous variation of symptoms or regression to the mean effects, and specific responses, e.g., due to patient or parent attitudes towards the treatment, participation in a trial and a supportive relationship with the doctor (Rodrigues and Ferreira 2020; Schedlowski et al. 2015).

It is suggested that placebo-effects are considerable in psychiatry (Weimer, Colloca, and Enck 2015). Therefore, the results of non-controlled prospective trials should be interpreted with great caution, since any observed improvement could be possibly explained by placebo effects and efficacy may not be found in RCTs. A notable example in ASD is secretin, which was promoted as a “miracle” medication in 1990s based on observational evidence, yet no efficacy was demonstrated later in RCTs (Williams, Wray, and Wheeler 2012).

In addition, substantial placebo-effects in RCTs could potentially dilute effect-sizes, i.e., medication-placebo differences, and thus, hinder assay sensitivity, i.e., the ability to discriminate efficacious medications (Enck et al. 2013; Rodrigues and Ferreira 2020). Accordingly, it is also suggested that a major concern of the low success rate of late-stage clinical trials in ASD is high placebo-effects, and thus, further research on strategies that could optimize placebo effects and subsequently medication-placebo differences are warranted (Berry-Kravis et al. 2018; Díaz-Caneja et al. 2021; Jacob et al. 2022; McCracken et al. 2021).

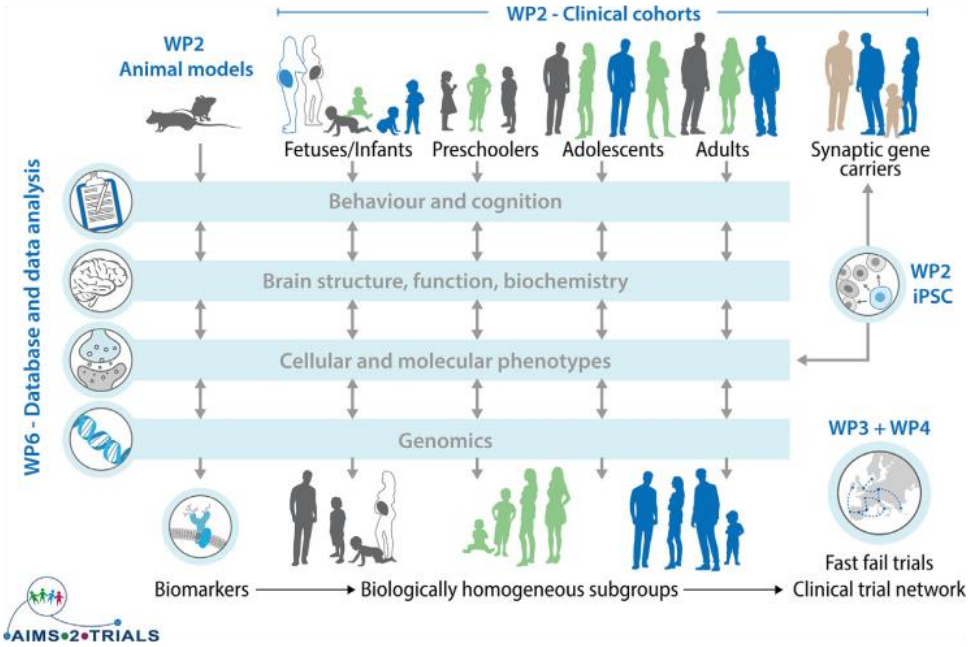
1.5. AIMS-2-TRIALS

According to the above, clinical research in ASD has generally failed to guide evidence-based clinical practice, is flooded by a large number of recently conducted RCTs with a low success rate, potentially due to a suboptimal design and substantial placebo effects, as well as due to the lack of biomarkers and understanding of the neurobiology of ASD.

In order to fill this gap, the Autism Innovative Medicine Studies-2-Trials (AIMS-2-TRIALS) consortium (https://www.aims-2-trials.eu/) is a European project that provides a framework and infrastructure of multidisciplinary and collaborative research in ASD with the aim to elucidate the neurobiology of ASD, facilitate the development of biomarkers and treatments, build a network among researchers, clinicians and the autism community, and explore policies for better support of autistic people (**Figure-1** from (Díaz-Caneja et al. 2021)). Thus, AIMS-2-TRIALS is the largest research grant awarded for ASD, and consists of 48 partners across 14 countries. It was based on the previous consortium of EU-AIMS (Murphy and Spooren 2012).

The work of the thesis was conducted within the Work Package 4 (WP4) of the AIMS-2-TRIALS, which is focusing on improving the design of trials and evaluating new medications.

Figure 1 AIMS-2-TRIALS framework from (Díaz-Caneja et al. 2021). The article was published Open Access and it was licensed under a Creative Commons Attribution 4.0 International License https://creativecommons.org/licenses/by-nc-nd/4.0/ .



1.6. Objectives

The purpose of the thesis was to facilitate current efforts in research by conducting a systematic review and meta-analysis of RCTs on pharmacological and dietary-supplement treatments for ASD. The thesis consisted of three sections, one for each publication, aiming to inform clinical practice (network meta-analysis), the design of RCTs (meta-analysis of placebo effects) and the methodology of systematic reviews (validation of an imputation method).

1.6.1. Network meta-analysis to inform clinical practice

The first part of the thesis was a network meta-analysis on the comparative efficacy and tolerability of medications and dietary-supplements in ASD (Siafis et al. 2022b).

This analysis was the first comprehensive network meta-analysis on this topic and provided an up-to-date synthesis of the evidence suitable for guiding clinical practice (Leucht et al. 2016).

Previous reviews had a limited focus on certain symptoms and specific medications and/or did not use a network meta-analysis, e.g., (Ameis et al. 2018; Fallah et al. 2019; Fraguas et al. 2019; Rodrigues et al. 2021; Salazar de Pablo et al. 2022; Zhou et al. 2021).

1.6.2. Meta-analysis of placebo-effects to inform the design of future RCTs

The second part of the thesis was a single-group meta-analysis and meta-regression analysis of the placebo arms (Siafis et al. 2020).

This analysis was the first to systematically quantify the magnitude of placebo-effects in the core symptoms, and identify predictors that could be considered in the design of future RCTs in order to optimize assay sensitivity.

Previous analyses were not focused on core symptoms (Masi et al. 2015) or were limited to post-hoc analyses of single trials (Arnold et al. 2010; King et al. 2013).

1.6.3. Validation of an imputation method to inform the methodology of future systematic reviews

The third part of the thesis was the validation of an imputation method to estimate the number of responders from the mean and standard deviation of the Clinical Global Impression Improvement scale (CGI-I) by assuming an underlying normal distribution (Siafis et al. 2021).

CGI-I is a clinician-administered 7-point Likert scale measuring global response to treatment ranging from 1 “very much improved” to 7”very much worse” (Guy 1976). Due to the lack of agreement on the outcome measures in ASD trials, it is generally recommended that all trials should use CGI-I for comparability of the results and

combined analyses (Aman et al. 2004; McCracken et al. 2021; Provenzani et al. 2020). However, the results of CGI-I could be reported as dichotomous, i.e., number of participants with a positive response (e.g., at least “much improved” with CGI-I score of 1 and 2) and/or as continuous, i.e., mean and standard deviation.

In order to facilitate the comparability and combination of the results across RCTs, an imputation method to estimate the number of responders from the mean and standard deviation of scale scores have been proposed and validated with depression (Furukawa et al. 2005) and schizophrenia scales (Samara et al. 2013). This method assumes an underlying normal distribution, and thus, its appropriateness could be questioned with potentially skewed data, e.g., due to the limited number of points of CGI-I and the small sample sizes in ASD trials.

Therefore, this analysis was the first to validate the imputation method with CGI-I in ASD trials, and it could inform the methodology of future systematic reviews in ASD, since the imputation method allowed the aforementioned meta-analyses to incorporate data from more trials and provide a more comprehensive synthesis of the evidence.

2. Material and methods

2.1. Protocol of the systematic review

The systematic review and meta-analyses were conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses statement (PRISMA) (Liberati et al. 2009), and the extension for network meta-analysis (Hutton et al. 2015). The protocol of the systematic review was predefined and registered to PROSPERO (ID: CRD42019125317; registration date on March 15, 2019, and updated on October 29, 2020).

The methods across the publications included in thesis were similar, as described below, and differences in the study inclusion criteria and data analysis are noted. More details about the methodology could be found in the respective publications and their appendices (see Publications)(Siafis et al. 2020; Siafis et al. 2022b; Siafis et al. 2021).

2.2. Search strategy and study selection

Comprehensive searches were conducted in multiple electronic databases Cochrane Central Register of Controlled Trials (CENTRAL), EMBASE, MEDLINE, PsychINFO, PubMed, and the clinical trial registries ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform (WHO-ICTRP).

The search strategies combined keywords for ASD, RCTs, pharmacological and dietary-supplement treatments, without any restriction in terms of publication date, language and document type. Searches were conducted in all databases from inception up to July 8, 2018, and updated in PubMed and CENTRAL up to July 4, 2019 (Siafis et al. 2020), up to August 31, 2020 (Siafis et al. 2021), and up to November 3, 2021(Siafis et al. 2022b).

In addition, reference lists of included studies and previous reviews, such as (Howes et al. 2018; Masi et al. 2017; Masi et al. 2015; Posey and McDougle 2000; Rodrigues et al. 2021; Zhou et al. 2021), were inspected for relevant trials.

2.3. Eligibility criteria

2.3.1. Study design

RCTs comparing any medication or dietary-supplement with each other or placebo for ASD were eligible. Head-to-head and placebo-controlled trials were included in the network meta-analysis (Siafis et al. 2022b) and the validation of the imputation method (Siafis et al. 2021), and placebo-controlled trials in the meta-analysis of placebo effects (Siafis et al. 2020).

There was no restriction in terms of blinding (e.g., open, single-, or double-blind) and country of origin. Randomization was assumed if not explicitly reported in case the trial was reported as double-blind. Quasi-randomized trials and trials with a high risk in random sequence generation or allocation concealment (e.g., allocation by date of administration) were excluded (Higgins et al. 2011).

In case of crossover trials, the first phase of crossover was considered in order to avoid carry-over effects (Elbourne et al. 2002).

Cluster-randomized (e.g., randomization at the ward level) and discontinuation studies were excluded. In addition, studies published before 1980 were excluded, since ASD and childhood schizophrenia were clearly separated after the introduction of standardized diagnostic criteria of DSM-III (APA 1980). Studies with a sample size smaller than 10 participants were also excluded in order to avoid potentially inflated effect sizes from clearly underpowered studies (Button et al. 2013), and the cutoff of 10 participants was proposed as a minimum quality indicator of research in ASD (Reichow, Volkmar, and Cicchetti 2008).

2.3.2. Participants

Participants should be diagnosed with ASD according to standardized diagnostic criteria (e.g., DSM-III, ICD-10 and their more recent versions), and/or validated diagnostic tools (Howes et al. 2018) (e.g., Autism Diagnostic Interview Revised, ADI-R)) (Rutter, Le Couteur, and Lord 2003). There were no restrictions in terms of age, sex, ethnicity, setting, baseline severity, presence of co-occurring difficulties or conditions (e.g., irritability, ADHD, sleep disorders, intellectual disability, genetic syndrome). Trials were excluded when more than 20% of the participants had a diagnosis different from ASD (e.g., another psychiatric or neurodevelopmental disorder), participants were characterized as “autistic” or with a genetic disorder and none of the aforementioned diagnostic criteria or tools were applied in the inclusion criteria of the studies.

2.3.3. Interventions and control

Any medication, dietary-supplement and placebo was eligible.

There was no restriction in terms of route of administration (e.g., oral, intranasal, intramuscular, etc.) and dosing schedule (e.g., fixed or flexible dosing), and multiple dose arms were combined according to the Cochrane Handbook (Higgins et al. 2019). The duration of treatment should be at least seven days, and treatments with a shorter-term duration or single-dose interventions were excluded.

Other intervention modalities were excluded, e.g., psychosocial, behavioral interventions, traditional or homeopathic medicine, elimination diets or milk formulations. Augmentation or multimodal interventions were also excluded (e.g., medications or dietary-supplements combined with risperidone or behavioral intervention).

2.4. Outcomes

Outcomes were measured at trial endpoint.

2.4.1. Primary outcomes

Core symptoms of ASD were the co-primary outcomes in the network meta-analysis (Siafis et al. 2022b) and the meta-analysis of placebo-effects (Siafis et al. 2020):

a. Social-communication difficulties (SCD), e.g., as measured with the Aberrant Behavior Checklist – Lethargy/Social Withdrawal (ABC-L/SW)(Aman et al. 1985) or the Vineland Adaptive Behavior Scales – Socialization domain (VABS-S)(Sparrow 2011).
b. Repetitive behaviors (RB), e.g., ABC-Stereotypic Behavior (ABC-S)(Aman et al. 1985) or Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders (CY-BOCS-PDD) (Scahill et al. 2006).
c. Overall core symptoms (OCS) in single scores, e.g., total score of the Social Responsiveness Scale (SRS)(Constantino and Gruber 2012) or the Childhood Autism Rating Scale (CARS) (Schopler et al. 1980).

2.4.2. Secondary outcomes

Co-occurring difficulties, caregiver stress, quality of life, global functioning, dropouts and specific adverse events were secondary outcomes in the network meta-analysis (Siafis et al. 2022b), and response to treatment was investigated in all publications included in the thesis (Siafis et al. 2020; Siafis et al. 2022b; Siafis et al. 2021):

a. Co-occurring difficulties: i) irritability or aggression, e.g., ABC-Irritability (ABC-I) (Aman et al. 1985), ii) ADHD symptoms, e.g., ABC-Hyperactivity/noncompliance (ABC-H)(Aman et al. 1985), iii) anxiety or depressive symptoms, e.g., Child and Adolescent Symptom Inventory (CASI) – anxiety (Lavigne et al. 2009). These three co-occurring difficulties were

prioritized as targets for interventions in a survey of parents with children with ASD (Anixt et al. 2020).

b. Caregiver stress, e.g., Parental Stress Index (PSI) (Abidin 1983).
c. Quality of life, e.g., Pediatric Quality of Life Inventory (PedsQL) (Varni et al. 2003)
d. Global functioning, e.g., Developmental Disabilities Modification of Children's Global Assessment Scale (DD-CGAS) (Wagner et al. 2007).
e. Number of participants with a positive response to treatment, defined by "at least much improved" or a score of ≤2 in CGI-I in the meta-analysis of placebo effects and the validation of the imputation method (Siafis et al. 2020; Siafis et al. 2021), and when not available, other definitions were also accepted in the network meta-analysis (Siafis et al. 2022b). In terms of multiple CGI-I ratings, we preferred ratings anchored to global over core and associated symptoms.
f. Number of participants discontinued prematurely from the study (dropouts) i) due to any reason, and ii) due to adverse events.
g. Number of participants with specific adverse events: i) at least one adverse event, ii) weight gain (preferable defined as ⩾ 7% increase from baseline), iii) sedation, and iv) at least one extrapyramidal symptom.

2.5. Study selection, data extraction and risk of bias

2.5.1. Procedure

Study selection, data extraction and assessment of risk of bias for each record or study was conducted independently in double by myself and at least one additional reviewer or contributor (see Publications). Any disagreements were resolved in discussion or consultation with a third experienced reviewer. Study selection was conducted in two steps, first by inspecting and selecting potentially relevant titles/abstracts in Rayyan

(Ouzzani et al. 2016), second by acquiring and inspecting their full-texts in Citavi (Swiss Academic Software, Zurich, Switzerland). Data extraction was conducted in a Microsoft Access database specifically designed for this project. In case of missing or unclear information, the first and/or corresponding author of the trial was conducted by e-mail, including a reminder.

2.5.2. Data extraction

Data extraction considered information about trial methodology (e.g., publication year, country of origin, funding, number of centers, sample size, duration, number of arms, risk of bias), participant characteristics (e.g., age, sex, presence of intellectual disability and other co-occurring difficulties), intervention (e.g., dose and route of administration), and outcome measures (see below sections 2.5.2.1, and 2.5.2.2).

2.5.2.1 Data extraction of continuous outcomes and rating scales

The continuous outcomes were rating scale scores, for which mean, standard deviations (SD) and number of participants analyzed were extracted.

There is still no universally accepted outcome measure for the core symptoms of ASD in RCTs (Aman et al. 2004; McCracken et al. 2021). Therefore, data from any validated rating scale were used, e.g., as in (Anagnostou et al. 2015; Baxter et al. 2015; EMA 2018; Lecavalier et al. 2014; McCracken et al. 2021; Scahill et al. 2015). Preference was given first to scales filled by clinicians (e.g., clinician observation or semi-structured interviews) than caregivers, teachers and participants, and second to commonly used scales (e.g., as those described in section 2.4), when data from more than one scale by the same informant were available.

Data from intention-to-treat (ITT) analyses and analyses that handle missing data were preferred, e.g., mixed-models of repeated measurement (MMRM) and multiple

imputation over last-observation carried forward (LOCF), yet data from observed case or per-protocol analyses were also eligible.

The results from rating scales could be reported as change or follow-up scores, which were both eligible in the analysis. Nevertheless, preference was given to change scores, and when not reported, they were estimated when possible from baseline and follow-up scores assuming a pre-post correlation (r) of 0.5 (Balk et al. 2013).

Missing SDs were obtained according the methods of the following order i) from standard errors, ii) confidence intervals, p-values, t-values, F-values (Higgins et al. 2019), iii) contacting study authors, iv) median and ranges (Hozo, Djulbegovic, and Hozo 2005; Wan et al. 2014), v) pooling subscale scores assuming r=0.5, and vi) the mean SD of the other studies in the review (Furukawa et al. 2006).

2.5.2.2 Data extraction of dichotomous outcomes

For dichotomous outcomes, the number of participants with an event and the number of participants randomized were extracted. An ITT approach was followed by assuming participants lost to follow-up did not have the event.

When the number of participants with a positive response to treatment or weight gain was not reported, it was estimated from means and SDs of CGI-I and weight, respectively using a validated imputation method (Furukawa et al. 2005; Samara et al. 2013; Siafis et al. 2021) (see 2.6.4. for the imputation method and the validation methodology with CGI-I).

2.5.3. Risk of bias assessment

The risk of bias of each trial was assessed using the Cochrane Risk of Bias tool version 1, which considers the domains of sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, missing outcome data, selecting reporting, and other biases (e.g., termination of the study,

baseline imbalance) (Higgins et al. 2011). Each domain could be rated with a “low”, “unclear” or “high” risk of bias (Higgins et al. 2011). Studies were also classified with an overall “low” (no domain with a high risk of bias and <4 domains with unclear), “moderate” (1 domain with a high risk of bias, or no domain with high risk of bias and >3 domains with unclear) or “high” risk of bias (in all other occasions) (Furukawa et al. 2016).

2.6. Data analysis

2.6.1. Effect-sizes

For continuous outcomes, effect-sizes were standardized mean differences (SMDs, Hedge’s g, treatment contrasts) and standardized mean changes (SMC, single-group), since different rating scales were used (Higgins et al. 2019). SMDs and SMCs standardize mean differences between groups or mean changes within a group, respectively, into SD units, e.g., $SMD = \frac{Mean\ difference}{SD}$. SMDs could be interpreted as “small” (<0.2), “medium” (=0.5) and “large” (>0.8) according to the Cohen’s rule of thumb (Cohen 1988), yet caution is required (Luo et al. 2022), and SMC=0.5 could correspond to a minimum clinically important difference (MCID) (Norman, Sloan, and Wyrwich 2003). About SMCs, the pre-post correlation (r) is also required, yet it is often not reported, and therefore, r=0.5 was assumed (Balk et al. 2013). In addition, SMCs were calculated with raw score standardization using preferably the baseline SD, and therefore, point estimates were not affected by the pre-post correlation, but only the variances (Becker 1988; Viechtbauer 2010).

For dichotomous outcomes, effect-sizes were odds ratios (ORs, treatment contrasts) due their preferred mathematical properties (Bakbergenuly, Hoaglin, and Kulinskaya 2019; Suhail A et al. 2020), and proportion of participants with an event (%, single-group). Natural logarithms of ORs and logit transformed proportions were used in the

meta-analysis, which were back-transformed for presentation (Schwarzer 2007; Schwarzer et al. 2019).

Effect-sizes were presented along with their 95% confidence intervals (95%CI).

2.6.2. Network meta-analysis

For the first objective of the thesis, meta-analyses were conducted separately in children/adolescents and adults, in order to avoid spurious extrapolations between age groups (EMA 2018).

2.6.2.1. Two-step procedure

A two-step procedure was followed by conducting first pairwise meta-analyses for all direct treatment comparisons, and second, if the necessary requirements were met (see 2.6.2.2.), a network meta-analysis to synthesize data from all available RCTs simultaneously and incorporate indirect evidence, i.e., interventions could be compared indirectly, even if there was no RCT comparing them.

A random-effects model was used, since heterogeneity is likely, and assuming a normal distribution for the underlying treatment effects (DerSimonian and Laird 1986; Higgins et al. 2019). Network meta-analyses were conducted using graph-theoretical methods within a frequentist framework (Rücker 2012). Interventions were also ranked using P-scores, an equivalent of the surface under the cumulative ranking curve (SUCRA) (Rücker and Schwarzer 2015).

2.6.2.2. Assessment of the transitivity assumption and incoherence

Transitivity (i.e., jointly randomizable interventions) and coherence (the statistical manifestation of transitivity, i.e., agreement between direct and indirect evidence) are the core assumptions of valid indirect comparisons, and thus of network meta-analysis (Salanti et al. 2014; Salanti et al. 2008).

Transitivity was assumed by expecting that participants included in the eligible RCTs to be equally likely to be randomized to any of the interventions of the review. The assumption was further investigated by examining the distribution of potential effect-modifiers (Salanti et al. 2014) and their effects in sensitivity analysis (see 2.6.2.4.), i.e., study duration, informant of the rating scale, mean age, presence of co-occurring difficulties at baseline, baseline mean scores of CGI-Severity (CGI-S, ranging from 1 ”normal” to 7 “among the most extremely ill”(Guy 1976)) as a measure of global severity of illness, and of ABC-I (ranging from 0-45)(Aman et al. 1985) as a proxy of serious behavioral problems (RUPP 2002). In addition, moderate levels of indirectness were assigned for studies focusing on participants with a co-occurring condition, when evaluating the certainty of the evidence (see 2.6.2.6).

Coherence was examined in closed loops, i.e., for pairings with both direct and indirect estimates, locally for a specific comparison and globally for the entire network (Efthimiou et al. 2016). The employed local method was the Separating Indirect from Direct Evidence (SIDE) approach, which evaluates the agreement of direct and indirect effect estimates for a specific comparison (Dias et al. 2010), and the global method used a design-by-treatment interaction model (Higgins et al. 2012). A descriptive analysis of published network meta-analyses found incoherence in about 10% of closed loops (Veroniki et al. 2013) and 20% of the entire networks (Veroniki et al. 2021). Therefore, and since incoherence tests have low statistical power, alpha was set at 0.1 (Veroniki et al. 2021).

2.6.2.3. Assessment of heterogeneity

Heterogeneity refers to the variation of treatment effects across studies. The amount of heterogeneity as reflected by τ^2 (between-study variance) was assumed to be common across all comparisons within a network (Higgins et al. 2019), and it was

further classified as low, moderate or high according to the empirical distribution of τ^2 (Rhodes, Turner, and Higgins 2015; Turner et al. 2012).

The test statistics χ^2 (alpha at 0.1) and I^2 (substantial heterogeneity when >50%) were also employed to assess heterogeneity in pairwise meta-analyses (Higgins et al. 2019).

2.6.2.4. Sensitivity analysis

Predefined sensitivity analyses were conducted to investigate the robustness of the results, also in terms of the transitivity assumption, incoherence and heterogeneity: i) using a fixed-effects model, ii) excluding studies with concerns in risk of bias (e.g., non-double-blind, presenting only observed case data), iii) with certain methodological factors (e.g., non-clinician ratings, shorter than four weeks), iv) with certain participant characteristics (e.g., presence of co-occurring difficulties by inclusion criteria), and v) with imputed data (e.g., estimated SDs).

2.6.2.5. Assessment of small-study effects and publication bias

Publication and other reporting bias refers to the selection or omission of the reported results, e.g., on the basis of significance or direction of the effect (Higgins et al. 2019). There were efforts to include unpublished studies by searching clinical trial registries and contacting authors. Failure to include data from identified unpublished studies was considered to evaluate the certainty in the evidence (see 2.6.2.6).

In addition, small-study effects were used as a proxy of publication bias (small studies could be more prone to publication bias) and they were indicated by an asymmetry in funnel plots (i.e., scatter plots of point estimates against their standard errors): i) contour-enhanced funnel plots for pairwise meta-analyses that visualize lines and areas concerning the statistical significance of the effect-sizes (Peters et al. 2008) and ii) comparison-adjusted funnel plots for network meta-analyses that account for the different comparisons across studies, and assuming that the direction of bias was

towards more recent interventions (Chaimani et al. 2013). Asymmetry was examined visually and with a linear regression test or Egger's test (alpha at 0.1) (Egger et al. 1997) when there were at least 10 studies (Higgins et al. 2019).

2.6.2.6. Certainty in the evidence

Confidence in the evidence of meta-analytic estimates of placebo-controlled comparisons was evaluated using the framework of the Confidence in Network Meta-analysis (CINeMA) (Nikolakopoulou et al. 2020) and its browser-based tool (Papakonstantinou et al. 2020). Confidence in the evidence was classified as "very low", "low", "moderate" and "high", by considering judgments on the domains of i) risk of bias of the studies, ii) publication and other reporting biases, iii) indirectness of the studies, iv) imprecision, v) heterogeneity, and vi) incoherence (Nikolakopoulou et al. 2020).

2.6.3. Meta-analysis of placebo-effects

For the second objective, single-group meta-analyses were conducted using data from the placebo arms of RCTs.

2.6.3.1. Meta-analysis

In a similar vein to the 2.6.2, a random-effects model was used, heterogeneity was assessed with the χ^2 and I^2 test statistics, small-study effects were assessed with contour-enhanced funnel plots, and sensitivity analyses were conducted.

2.6.3.2. Meta-regression analysis and correlations

Meta-regression analyses were conducted to investigate potential predictors of placebo-effects in the core symptoms. A list of potential predictors of placebo-effects were identified in the literature, e.g., (King et al. 2013; Leucht et al. 2019; Masi et al. 2017; Masi et al. 2015; Sandler and Bodfish 2000; Schedlowski et al. 2015; Weimer,

Colloca, and Enck 2015; Weimer et al. 2013). They were predefined in the protocol and classified into study-, intervention-, and participant-related factors (**Table-1**).

Table 1 List of investigated potential predictors of placebo-effects in core symptoms of ASD.

Study-related factors	**Intervention-related factors**	**Participant-related factors**
• Duration of treatment (in weeks) • Publication year • Washout from previous psychotropic medications (yes vs. no) • Placebo lead-in with exclusion of responders (yes vs. no) • Informant of the rating scale (clinician vs. caregiver) • Sample size (in number of participants) • Number of sites • Academic sites (%) • Number of arms • Number of different experimental interventions • Participants randomized into placebo (%) • Financial interest (industry-sponsored or patent application vs. academic funding) • Country of origin (US vs. not only US) • Risk of bias domains (low vs. unclear/high)	• Route of administration (oral vs. other) • Experimental intervention (pharmacological vs. dietary-supplement) • Dose-schedule (fixed vs. flexible)	• Presence of associated conditions at baseline by inclusion criteria (except for intellectual disability or a genetic syndrome) • Mean age (in years) • Age group (children/adolescents vs. adult or mixed populations) • Participants with intellectual disability (%, study-defined or intellectual quotient <70) • Female participants (%) • Ethnicity (% Caucasian or Hispanic) • Mean baseline BMI (in mg/kg^2) • Threshold of core symptom severity as inclusion criteria (yes vs. no; not just a threshold for the confirmation of the diagnosis) • Mean CGI-Severity • Mean ABC-I Due to inconsistent reporting and large number of different scales used, baseline severity in core symptoms could not be assessed.

In the meta-regression analyses, SMC was the dependent variable, and the potential predictors were the independent variables. Univariable meta-regressions were conducted separately for each factor with enough data (i.e., at least five data points). Multivariable meta-regressions were also conducted similarly to previous analysis (Leucht et al. 2019), by first selecting factors that were found to be significant in the univariable models, and then applying a backward step-wise algorithm by removing factors from the model with p-values>0.15. Alpha was not adjusted for multiple testing, due to the low statistical power and exploratory nature of meta-regression analysis (Higgins 2011).

In addition, Spearman's ρ were conducted to assess correlations between the potential predictors, as well as the relationship between the SMC of placebo and of the experimental intervention.

2.6.4. Validation of the imputation method

For the third objective, the number of participants with a positive response to treatment (as defined by at least "much improved" or a score 1 or 2 in CGI-I) was imputed from the mean and SD of CGI-I. Therefore, studies that provided both continuous and dichotomous data on CGI-I were included in this analysis. This was a *post-hoc* analysis.

2.6.4.1. Imputation method and selection of cutoff

The imputation method was previously validated with depression (Furukawa et al. 2005) and schizophrenia scales (Samara et al. 2013). The aim of the method is to impute the number of responders from the mean (m), SD and number of participants analyzed (n) of a rating scale, given a certain cutoff (θ) and assuming a normal distribution of the continuous data. Thus, the number of responders, i.e., participants with a score of ≤θ, could be calculated as $n * p$, where p is the probability of the lower tail for $Zscore = \frac{(\theta - m)}{SD}$.

CGI-I is a 7-point Likert-scale, therefore, an assumed normal distribution of the underlying latent continuous variable may have different cutoffs to classify responses than the original Likert scale (Liddell and Kruschke 2018), e.g., cutoffs from 2 to 2.5 in the underlying latent variable could be considered to impute the number of participants "at least much improved" (CGI-I of 1 or 2). Therefore, θ=2.5 was used in the primary analysis, and θ=2 in a secondary analysis.

2.6.4.2. Validation of the method

Responder rates and logORs were calculated from the original and imputed number of responders. The performance of the imputation method was evaluated in a similar vein with a previous analysis (Samara et al. 2013): i) The agreement between imputed and original values was examined with the concordance correlation coefficient (CCC), ranging from -1 to 1 and |CCC|=1 indicated perfect agreement. ii) The predictive accuracy was examined with linear regression models, and a slope=1, R^2=100% and mean squared error (MSE)=0 indicated perfect accuracy. iii) The limits of agreement of bias (i.e., difference between a pair imputed and original values) were examined with Bland-Altman plots (Bland and Altman 1986; Bland and Altman 1999). Bland-Altman plots present the bias (y-axis) against the average of imputed and original values (x-axis). The mean bias and 95% limits of agreement were calculated, assuming normally distributed biases (Bland and Altman 1986). For this reason, the normality assumption was assessed by inspection of the distribution of bias and a Shapiro-Wilk test. In addition, the relationship between bias and the average was investigated with linear regression models (Bland and Altman 1999). v) Subgroup analyses compared the pooled estimates between the original and imputed values as obtained from random-effects meta-analysis

2.6.5. Statistical software

Data cleaning and analysis was conducted in R statistical software (R Core Team 2013) using the packages epiR (Stevenson et al. 2020), meta (Schwarzer 2007), metafor (Viechtbauer 2010), netmeta (Rücker et al. 2016) and tidyverse (Wickham et al. 2019). Alpha was set at two-sided 0.05, except otherwise reported.

3. Results and publication summaries

3.1. Network meta-analysis on the effects of pharmacological and dietary-supplement treatments for ASD

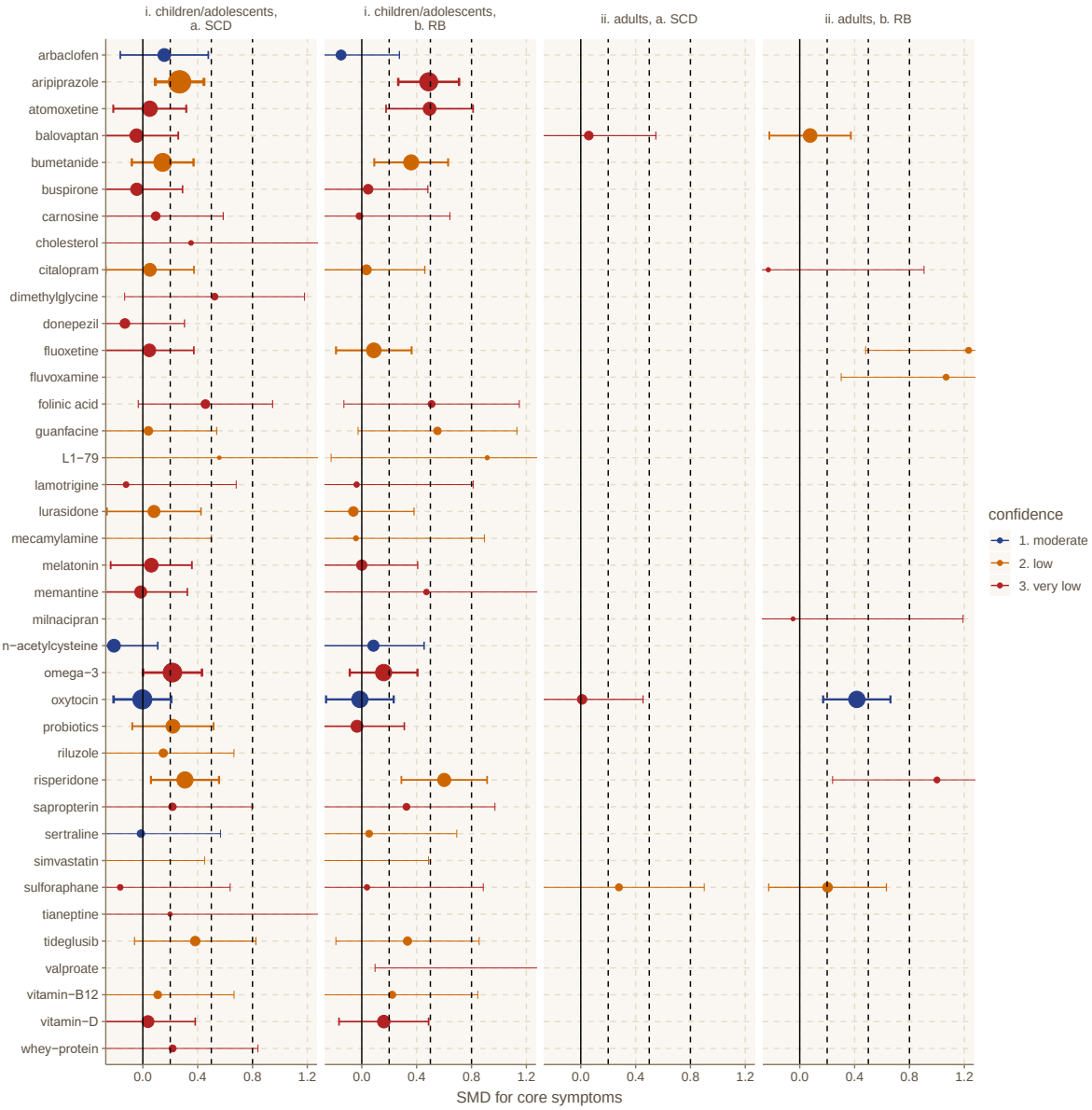
Detailed results of this analysis can be found in the respective publication (Siafis et al. 2022b) (see Publications).

This analysis included data from k=143 RCTs with n=8554 participants (children/adolescents: k=125, n=7450; adults: k=18, n=1104) (search up to November 3, 2021), which investigated 59 different treatments (41 pharmacological, 17 dietary-supplements and placebo). RCTs were mainly double-blind, placebo-controlled, recently published (median 2015), short (8-13 weeks), and small (20-80 participants), as well as a third of them was focused on co-occurring difficulties (e.g., irritability or ADHD symptoms per inclusion criteria) and a fifth had a high risk of bias. Pairwise and network meta-analyses on the effects of pharmacological and dietary-supplement treatment were conducted.

There were some medications that could improve at least one core symptom domain (mainly with small-to-medium effects-sizes) in children/adolescents and/or adults, i.e., aripiprazole, atomoxetine, bumetanide, fluoxetine, fluvoxamine, oxytocin, risperidone. There was also some supporting evidence for haloperidol, folinic acid, guanfacine, omega-3-fatty acids, probiotics, sulforaphane, tideglusib and valproate, yet imprecise based on small samples and a few studies, not robust and the 95%CIs did not exclude the null effect (**Figure-2**).

The confidence in the evidence was generally very low or low according to the CINeMA framework because of risk of bias of the studies, indirectness (e.g., participants with co-occurring difficulties), heterogeneity, imprecision (e.g., due to small sample sizes) and reporting biases (e.g., unpublished data due to negative findings).

Figure 2 Network meta-analytic estimates on the effects of pharmacological and dietary supplement treatments for the core symptoms of ASD (social-communication difficulties, SCD and repetitive behaviors, RB) in children/adolescents and adults. Standardized mean differences (SMD) for comparisons with placebo and their 95% confidence intervals are presented (SMD>0 indicate a favorable outcome for the medication or dietary supplement; |SMD|=0.2 represents small effects, |SMD|=0.5 medium and |SMD|=0.8 large as presented by the dashed lines) The size of the point is proportional to the inverse standard error of the estimate. Confidence in the evidence is presented with a color key. Scales measuring overall core symptoms in single scores were also evaluated, yet not presented here. Not all of the 58 medications or dietary supplements had data for the core symptoms. The results in the plots may be slightly different from the ones in the publications, because they were produced using an updated version of the package netmeta 2.1-0.



The findings on the secondary outcomes, e.g., co-occurring difficulties and side-effects, were generally in agreement with previous studies (e.g., aripiprazole and risperidone

were efficacious for irritability and ADHD symptoms, and they were associated with sedation, weight gain and extrapyramidal symptoms).

According to the findings and their limitations, routine prescription of pharmacological or dietary-supplements for the core symptoms of ASD could not be recommended. In addition, current medications that are indicated for co-occurring difficulties (e.g., aripiprazole, risperidone, ADHD medications) may also improve core symptoms to some degree and probably in an unspecific manner. However, the evidence is generally preliminary and further research is warranted, given also that there were limited data for some medications (e.g., methylphenidate), outcomes (e.g., anxiety) and adults.

3.2. Meta-analysis of placebo-effects in the core symptoms of ASD

Detailed results of this analysis can be found in the respective publication (Siafis et al. 2020) (see Publications).

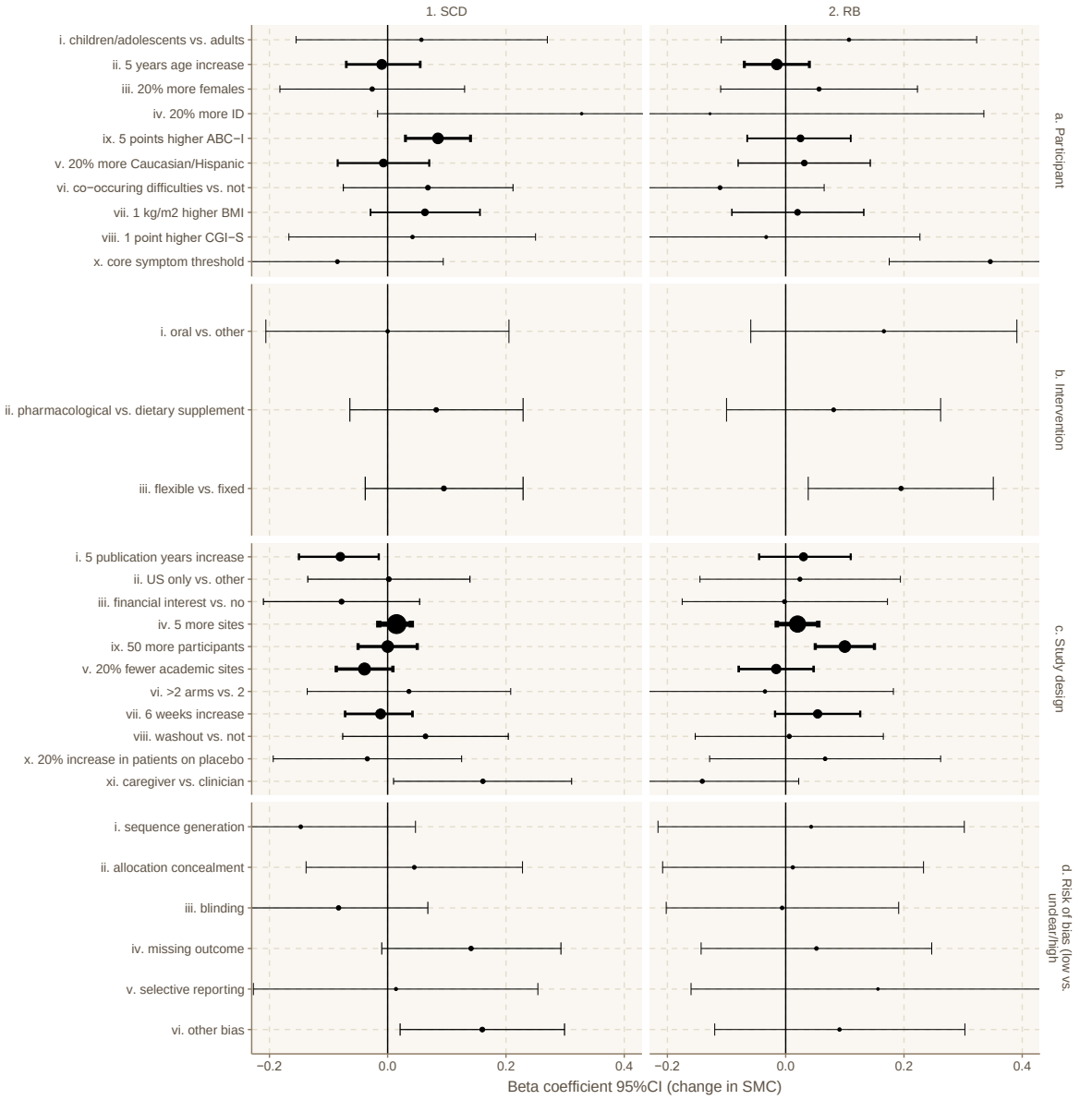
This analysis included data from the placebo arms of k=86 RCTs with n=2360 participants (search up to July 4, 2019), which had similar descriptive characteristics to the sample of the network meta-analysis (see 3.1). Single-group meta-analyses were conducted to quantify the magnitude of the placebo-effects in core symptoms and identify potential predictors.

There were considerable placebo-effects in ASD, and on average about a fifth (19%, 95%CI[16%, 22%]) of the participants had a positive response as defined by at least much improvement in CGI-I (CGI-I score of 1 or 2). The magnitude of the placebo-effects in terms of standardized mean changes (SMC) was on average 0.32 95%CI[0.25, 0.39] for social-communication difficulties, 0.23 95%CI[0.15, 0.32] for repetitive behaviors, and 0.36 95%CI [0.26, 0.46] in scales measuring overall core symptoms in single scores. There was some heterogeneity in these estimates (I^2 ranging from 32% to 55%) and small-study effects in repetitive behaviors (i.e., the trim-and-fill adjusted estimate was 0.33 95%CI [0.25, 0.41]).

Potential predictors of higher placebo-effects in at least one core symptom domain were identified from a list of factors related to the study design, participant characteristics or intervention, i.e., older studies, larger sample sizes, more sites, lower risk of bias (vs. higher), caregiver ratings (vs. clinician), flexible dosing (vs. fixed), higher levels of irritability at baseline, and the use of a threshold of core symptom severity as inclusion criteria (**Figure-3**). However, there were limited and/or narrowly ranged data for some factors (e.g., participant-related factors), and few influential

outlier studies (e.g., three large antidepressant trials with substantial placebo-effects in repetitive behaviors (Herscu et al. 2020; King et al. 2009; Reddihough et al. 2019)).

Figure 3 Meta-regression coefficients of potential predictors of placebo effects in the core symptoms of ASD (social-communication difficulties, SCD and repetitive behaviors, RB). Beta coefficients and their 95% confidence intervals of univariable meta-regressions are presented (beta>0 corresponds to an increase in placebo-effects as measured with standardized mean changes, SMC). The size of the point is proportional to the inverse standard error of the estimate. Scales measuring overall core symptoms in single scores were also evaluated, yet not presented here. The results in the plots may be slightly different from the ones in the publications, because they were produced using an updated version of the package metafor 3.0-2.



According to the findings, it can be recommended that potential predictors of placebo effects should be considered in the design of future trials, i.e., i) adequately powered trials, avoiding extremely large sample sizes, careful selection of sites, rigorous

enrollment of participants, ii) careful selection of measurements, not solely based on caregivers, use of different scales for inclusion and outcome assessment, proper training of the outcome assessors iii) careful selection of study duration and dosing based on the mechanism of action of the investigated medication or dietary-supplement. However, further research is warranted because it was not possible to investigate a full multivariable model (e.g., missing data), participant-related factors (e.g., missing data, study-level data) and the different impact of predictors on placebo and treatment effects due to the diverse experimental interventions (see 3.1).

3.3. Imputing the number of responders from mean and SD of CGI-I

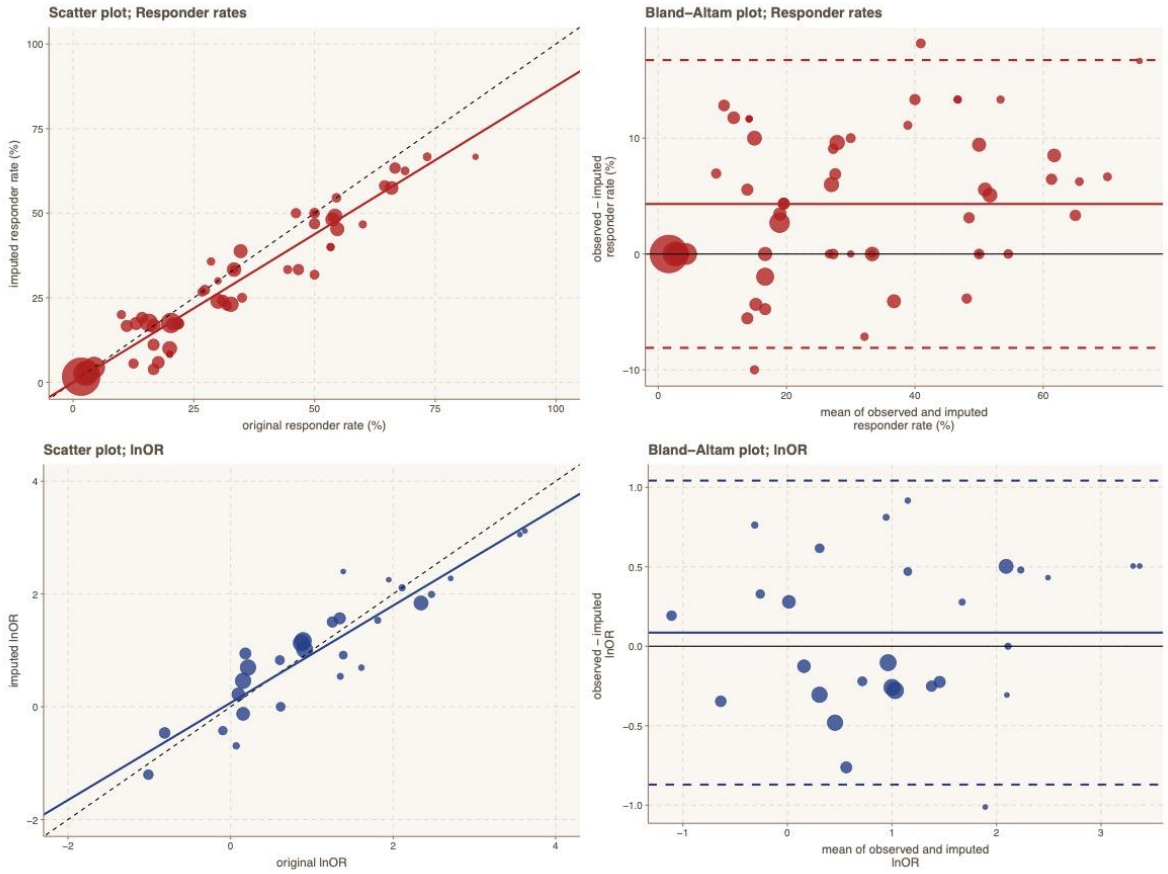
Detailed results of this analysis can be found in the respective publication (Siafis et al. 2021)(see Publications).

This analysis included 27 RCTs with 58 arms and 1428 participants that reported CGI-I data in both dichotomous (i.e., number of participants with a positive response defined by a CGI-I score of 1 or 2) and continuous format (i.e., mean and SD of CGI-I) (search up to August 31, 2020).

An imputation method was applied to estimate responder rates and odds ratios (ORs) from the mean and SD of CGI-I, assuming a normal distribution of a latent continuous variable and taking into consideration the Likert-scale nature of CGI-I (i.e., using ≤2.5 instead of ≤2 as the primary cut-off of the latent continuous variable to estimate the number of participants with a CGI-I of 1 or 2). The performance of the method was examined with concordance correlation coefficients, linear regression models, Bland-Altman plots and subgroup differences of summary estimates from single-group and pairwise meta-analysis.

The method had an overall good performance. In particular, the bias was towards more conservative estimates, i.e., imputed responder rates were on average 4.3% 95%CI[-8.1%, 16.7%] smaller, and imputed odds ratios were on average 1.1 95%CI[0.42, 2.83] (calculated by the exponential bias) times smaller in comparison to the original values (**Figure-4**). However, the limits of agreement of the bias were substantially wide, e.g., in comparison to schizophrenia scales (difference of responder rates 95%CI[-9.8%, 8.4%]; ratio of odds ratios 95%CI [0.79, 1.42]) (Samara et al. 2013). Summary estimates of meta-analyses that used imputed or original values did not differ. In addition, the performance of the method was poorer when a cutoff of ≤2 of was used in a secondary analysis.

Figure 4 Scatter and Bland-Altman plots for responder rates (%, red color) and natural logarithm of odds ratios (lnOR, blue color). In both plots, the points represent the pair of imputed and original values, and their size is proportional to the inverse of their standard error. In the scatter plots, the dashed line represents the identity line and the solid color line the linear regression model. In the Bland-Altman plots, the solid black line represents the line of no difference, while the color line represents the bias and the 95% confidence intervals.



According to the findings, the imputation method that estimates the number of responders from the mean and SD of scale scores (Furukawa et al. 2005; Samara et al. 2013) was further validated with CGI-I in clinical trials of ASD, taking also into consideration the Likert-scale nature of CGI-I. Given the overall good performance of the method, it could be used in meta-analysis in order to incorporate data from more studies and provide a more comprehensive synthesis of the evidence. However, and due to the relatively wide limits of agreement, a sensitivity analysis by excluding imputed values can be recommended in order to test the robustness of the results.

3.4. Own contributions to the publications

For all of the three publications included in the thesis (Siafis et al. 2020; Siafis et al. 2022b; Siafis et al. 2021), I was the first author and primary contributor to all of the procedures, including in the conceptualization and design of the study, writing and registering the protocol, design of the search strategy, evaluation of all records identified by the search, hand-searching of reviews for additional studies, data extraction and risk of bias evaluation of all eligible studies, contacting authors for additional data, conceptualization and conducting the statistical analysis, evaluating confidence in the evidence, interpretation of the data, writing and finalizing the manuscript, submitting for publication, revising the manuscript according to the reviewer comments, and supervising the work of other reviewers and/or contributors (e.g., study selection, data extraction, and risk of bias assessment of the individual studies in duplicate).

Additional information about author contributions can be found in the relevant section of the respective publications (see also Publications).

4. Discussion

4.1. Summary of the findings

In this thesis, I conducted and presented the first comprehensive systematic review of a total 203 RCTs with 12111 participants that investigated pharmacological and dietary-supplement treatments for ASD.

4.1.1. Network meta-analysis on the effects of pharmacological and dietary-supplement treatments for ASD

The first part of the thesis was a network meta-analysis on the efficacy and tolerability of pharmacological and dietary-supplement treatments for people with ASD (see 3.1). This analysis provided a more comprehensive synthesis of the evidence in comparison with previous reviews that had a more limited focus on certain symptoms and specific medications, e.g., (Ameis et al. 2018; Fallah et al. 2019; Fraguas et al. 2019; Rodrigues et al. 2021; Salazar de Pablo et al. 2022; Zhou et al. 2021).

The summary of the main results presented here was produced using a minimally contextualized framework of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach (Brignardello-Petersen et al. 2020). Interventions were classified into two categories based on their efficacy in the core symptoms (among the most effective, and among the least effective) and two groups based on the certainty in the evidence (moderate and low).

4.1.1.1. Children/adolescents

In terms of change in core symptoms (primary outcomes), there was low certainty evidence that aripiprazole, risperidone, atomoxetine, guanfacine, valproate, folinic acid, omega-3-fatty acids, bumetanide and tideglusib might be effective in improving at least one core symptom domain generally with small-to-medium effect-sizes (**Table-2**). Among these possible candidates for the core symptoms, aripiprazole, risperidone,

atomoxetine, and guanfacine could also improve irritability and/or ADHD symptoms, albeit associated with adverse events.

Table 2 Medications that might be effective in improving at least one core symptom domain in children/adolescents with ASD. The mechanism of action of medications was extracted from the Neuroscience-Based-Nomenclature (NbN)(Zohar et al. 2015) and the IUPHAR/BPS (Harding et al. 2022). SMD: standardized mean difference for the comparison between medication and placebo (>0: favors medications; significant SMDs are noted with bold). SCD: social-communication difficulties, RB: repetitive behaviors; D2: dopamine 2 receptor, 5-HT1A/2/2A: serotonin 1A, 2 and 2A receptors, NET: norepinephrine transporter, alpha-2: norepinephrine alpha-2 receptor; ALDH5A1: aldehyde dehydrogenase 5 A1, HDAC: histone deacetylase; FFAR: free fatty acid receptor, RXRa: retinoid X receptor a; GSK-3b: glycogen synthase kinase 3 beta

Pharmacological class	**Mechanism of action**	**Medication (k number of studies)**	**SMD_{SCD} 95%CI**	**SMD_{RB} 95%CI**	**Efficacy in co-occurring difficulties**	**Comment**
Very low or low certainty in the evidence						
Antipsychotic	D2, 5-HT1A partial agonist, 5-HT2A antagonist	Aripiprazole (k=6)	**0.27 (0.09, 0.44)**	**0.48 (0.26, 0.7)**	$SMD_{irritability}$= 0.63 (0.44, 0.82) SMD_{ADHD}= 0.82 (0.59, 1.05)	Aripiprazole had an increased risk of any adverse event, sedation, weight gain, extrapyramidal symptoms. The majority of the trials were conducted in participants with irritability.
Antipsychotic	D2, 5-HT2, alpha-2 antagonist	Risperidone (k=4)	**0.31 (0.06, 0.55)**	**0.6 (0.29, 0.9)**	$SMD_{irritability}$ =1.05 (0.75, 1.33) SMD_{ADHD}= 0.79 (0.47, 1.11)	Risperidone had an increased risk of any adverse event, sedation, weight gain (higher than aripiprazole), extrapyramidal symptoms. The majority of the trials were conducted in participants with irritability.
ADHD medication	NET inhibitor	Atomoxetine (k=3)	0.05 (-0.22, 0.32)	**0.49 (0.18, 0.8)**	SMD_{ADHD}= 0.64 (0.30, 0.99)	The trials were conducted in participants with ADHD symptoms.
ADHD medication	alpha-2 agonist	Guanfacine (k=1)	0.04 (-0.46, 0.54)	0.55 (−0.02, 1.11)	$SMD_{irritability}$ =0.5 (0.00, 1.01) SMD_{ADHD}= 1.39 (0.73, 2.05)	Guanfacine had an increased risk of any adverse event and sedation. The trial was conducted in participants with ADHD symptoms.
Antiepileptic/mood-stabilizer	unclear, ALDH5A1, HDAC inhibition	Valproate (k=0/1)	n.a.	1.33 (−0.03, 2.68)	-	The results were imprecise, based on one small RCTs, and the lower boundary of the 95%CI crossed the line of no difference.
Dietary supplement	co-factor in single-carbon transfer reactions	Folinic acid (k=2/1)	0.44 (-0.05, 0.93)	0.5 (−0.13, 1.13)	-	The results were imprecise, based on two small RCTs, and the lower boundary of the 95%CI crossed the line of no difference.
Dietary supplement	unclear, FFAR, RXRa agonist	Omega-3-fatty acids (k=10/9)	0.21 (0.00, 0.43)	0.15 (−0.09, 0.4)	-	The results were imprecise and the lower boundary of the 95%CI crossed the line of no difference.
Loop diuretic	Na-K-Cl symporter inhibitor	Bumetanide (k=4)	0.14 (−0.08, 0.37)	**0.35 (0.09, 0.62)**	-	There were two negative phase-III trials that did not provide usable data for the meta-analysis (Crutel et al. 2021).
Experimental	GSK-3b inhibitor	Tideglusib (k=1)	0.38 (-0.06, 0.82)	0.33 (−0.18, 0.84)	-	Data from an abstract were used (Anagnostou et al. 2018). The results were imprecise and the lower boundary of the 95%CI crossed the line of no difference.

The rest of medications were classified as among the least effective in improving core symptoms with a moderate (arbaclofen, oxytocin and N-acetylcysteine) or low certainty of the evidence. Nevertheless, some of them might be efficacious in improving co-occurring difficulties, e.g., citalopram and sulforaphane for irritability, olanzapine and naltrexone for ADHD symptoms, melatonin for caregiver stress. These medications had also an increased risk of adverse events, while other medications may also worsen co-occurring difficulties, e.g., irritability by levetiracetam and vitamin-B12.

4.1.1.2. Adults

There was generally sparse data in adults. There was moderate certainty evidence that oxytocin was effective for repetitive behaviors with small-to-medium effect-sizes, and low certainty evidence from single trials that fluoxetine, fluvoxamine and risperidone might be effective for repetitive behaviors with large effect-sizes (**Table-3**).

Table 3 Medications that were found to be or might be effective in improving at least one core symptom domain in adults with ASD. SMD: standardized mean difference for the comparison between medication and placebo (>0: favors the medications; significant SMDs are noted with bold) SCD: social-communication difficulties, RB: repetitive behaviors; D2: dopamine 2 receptor, 5-HT2: serotonin 2 receptor, SERT: serotonin transporter, alpha-2: norepinephrine alpha-2 receptor; OT: oxytocin receptor, VP: vasopressin receptor

Pharmacological class	**Mechanism of action**	**Medication (k number of studies)**	**SMD_{SCD} 95%CI**	**SMD_{RB} 95%CI**	**Efficacy in co-occurring difficulties**	**Comment**
Moderate certainty in the evidence						
Neuropeptide	OT, VP agonist	Oxytocin (k=4/6)	0.01 (-0.43, 0.44)	**0.41 (0.16, 0.66)**	-	The effects on social-communication difficulties were mixed. The studies were mainly conducted in high-functioning participants.
Very low or low certainty in the evidence						
Antidepressant	SERT inhibitor	Fluoxetine (k=0/1)	n.a.	**1.2 (0.45, 1.96)**	-	The results were based on one small RCT.
Antidepressant	SERT inhibitor	Fluvoxamine (k=0/1)	n.a.	**1.04 (0.27, 1.81)**	-	The results were based on one small RCT.
Antipsychotic	D2, 5-HT2, alpha-2 antagonist	Risperidone (k=0/1)	n.a.	**0.97 (0.21, 1.74)**	$SMD_{irritability}$=1.19 (0.34, 2.04)	Risperidone had an increased risk of any adverse event. The results were based on one small RCT.

The rest of medications were classified as among the least effective in improving core symptoms with low certainty of evidence.

4.1.1.3. Limitations

This analysis had certain limitations, which were clearly reported in the respective publication (Siafis et al. 2022b). In terms of the analytical method, the networks were mainly star-shaped and poorly connected, and most of the interventions were investigated in a few trials with small sample sizes. Therefore, heterogeneity and incoherence could not be adequately tested due to the limited statistical power, which raises also concerns about the transitivity assumption and the potential confounding of the effect estimates by study characteristics. For example, the effects of some medications could have been confounded by the presence of co-occurring difficulties, e.g., antipsychotics and ADHD medications. There were also indications for reporting bias, despite the efforts to include data from unpublished studies. These concerns were addressed among others in the evaluation of the evidence using the CINeMA approach (Nikolakopoulou et al. 2020).

In terms of the characteristics of eligible studies, there were limited and scattered data for adult populations, certain commonly used medications (e.g., methylphenidate due to the crossover design of the studies (Rodrigues et al. 2021; Sturman, Deckx, and van Driel 2017)) and some important outcomes (e.g., anxiety despite being a top research priority (Anixt et al. 2020; Autistica 2016)). In addition, the included clinical trials used different scales to measure core symptoms of ASD, which requires further investigation of their conformity (McCracken et al. 2021). There is no agreement on outcome measures of change in core symptoms of ASD (McCracken et al. 2021), yet some scales could be considered at least “appropriate with conditions” and were

generally preferred in the review (Anagnostou et al. 2015; Lecavalier et al. 2014; McCracken et al. 2021; Scahill et al. 2015).

Given that a large number of medications with different mechanisms of action were investigated, a comprehensive investigation of the tolerability and safety was out of the scope of this analysis. Nevertheless, the analysis considered three important side-effects, i.e., sedation, weight gain and extrapyramidal symptoms, that could overlap across psychotropic medications (Solmi et al. 2020), as well as dropouts and any adverse event. Further investigation is warranted, given that there may be a higher risk of side-effects in individuals with ASD (Howes et al. 2018).

4.1.2. Meta-analysis of placebo-effects in the core symptoms of ASD

The second part of the thesis was a meta-analysis of the magnitude and predictors of placebo-effects in the core symptoms of ASD (see 3.2). This analysis provided more definite answers in comparison with previous analyses that were not focused on core symptoms (Masi et al. 2015) or were based on single trials (Arnold et al. 2010; King et al. 2013).

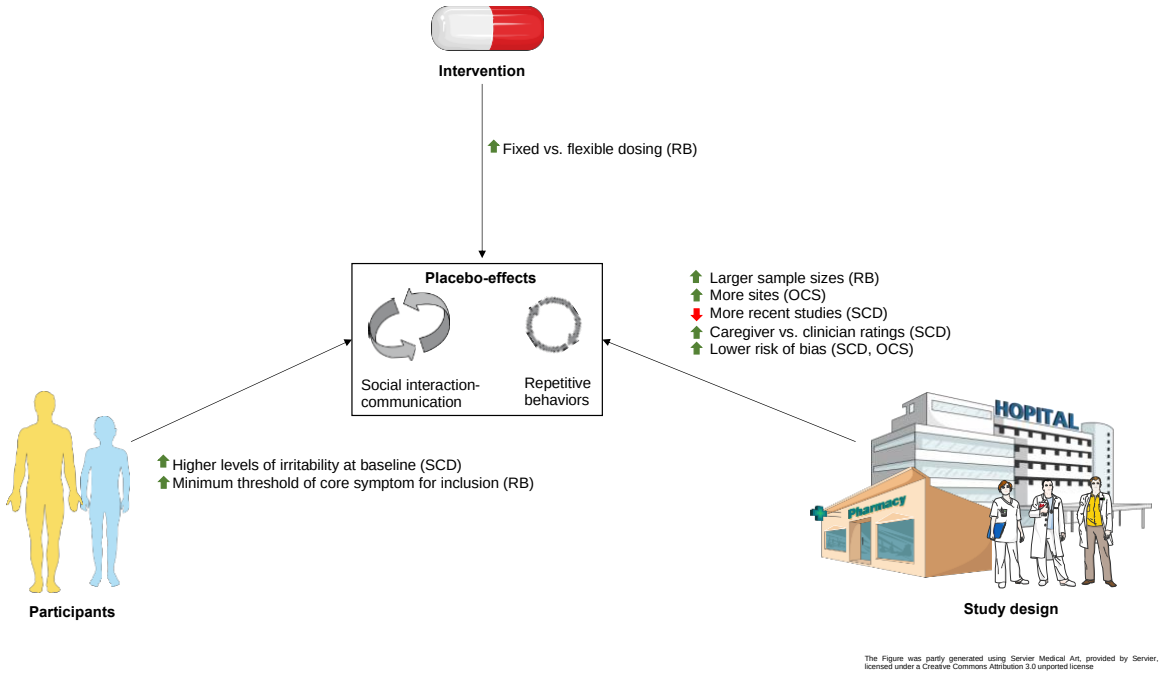
4.1.2.1. Magnitude of placebo-effects

RCTs in ASD may be prone to substantial placebo-effects, given that on average about a fifth of the participants had a clinically important response in terms of global impression. In particular, placebo-effects in the core symptom domains seemed to be of comparable magnitude, and on average (SMC_{SCD}=0.32, SMD_{RB}=0.23 or 0.33 the trim-and-fill adjusted estimated) about more than the half of the minimum clinically important difference ($MCID_{SMC}$=0.5)(Norman, Sloan, and Wyrwich 2003). Therefore, this effect indicates that the core symptoms may be improved in a considerable number of participants receiving treatment with placebo (Johnston et al. 2010).

4.1.2.2. Predictors of placebo-effects

There were some potential predictors of placebo-effects in core symptoms of ASD, in terms of participant (irritability levels at baseline, minimum threshold of core symptom severity for inclusion), intervention (dosing schedule) and study design characteristics (number of sites, sample size, type of informant, risk of bias, publication year) (**Figure-5**). In addition, predictors could have a different impact on the placebo-effects of the two core symptom domains.

Figure 5 Schematic presentation of predictors of placebo-effects in core symptoms of ASD. SCD: Social-communication difficulties, RB: repetitive behaviors, OCS: overall core symptoms. The Figure was partly generated using Servier Medical Art, provided by Servier, licensed under a Creative Commons Attribution 3.0 unported license.



4.1.2.3. Limitations

This analysis had certain limitations, which were clearly discussed in the respective publication (Siafis et al. 2020). In terms of the analytical method, potential predictors of placebo-effects were examined with a series of univariable meta-regression models, and a full multivariable meta-regression model that could adjust the effects of multiple variables could not be conducted due to missing information for many of the predictors across studies. Accordingly, data for many of the investigated factors were missing,

inconsistently reported across studies or narrowly ranged, and especially for participant-related factors (e.g., age, sex, baseline severity of core symptoms and co-occurring difficulties, intellectual disability). Meta-regressions used aggregate data of participant-level factors as independent variables, and thus, there was the risk of ecological fallacy (Geissbühler et al. 2021). Further investigation with individual-participant-data meta-analysis is warranted.

There was also a correlation between placebo- and drug-effects (moderate for social-communication difficulties, and weaker for repetitive behaviors). This may indicate that efforts to reduce placebo-effects may influence also drug-effects. In addition, some factors may have a different impact on placebo- and drug-effects, e.g., as found in RCTs of antipsychotics for schizophrenia (Leucht et al. 2019). Therefore, moderators of drug-placebo differences would be important to guide the design of future trials, e.g., as in schizophrenia (Leucht et al. 2017). However, such an analysis of potential moderators of drug-placebo differences was not possible here, since medications with substantially different mechanisms of action were investigated in the RCTs.

4.1.3. Imputing the number of responders from mean and SD of CGI-I

The third part was a further validation of a method to estimate the number of responders from continuous data of CGI-I in ASD (see 3.3).

The method assumed a normal distribution of the continuous data and it had already been validated with depression (Furukawa et al. 2005) and schizophrenia scales (Samara et al. 2013). The satisfactory performance of the method was replicated in this analysis, and it was better when a higher threshold of an underlying latent continuous variable of CGI-I was considered (e.g., a score of ≤2.5 instead of ≤2 to impute the number of participants with a CGI-I of 1 or 2). The imputation method could further facilitate the combination and/or comparison of findings across RCTs in ASD,

given that there is still no optimal scale to measure change in core symptoms (McCracken et al. 2021), and thus, CGI-I is recommended to be used in all RCTs in ASD (Aman et al. 2004; Provenzani et al. 2020).

Nevertheless, there were some limitations. The limits of agreement between the imputed and original values were wider than the ones found in schizophrenia scales (Samara et al. 2013). There was also skewness of the CGI-I data in about half of the arms. This could cast doubts about the normality assumption, yet the performance of the methods unexpectedly satisfactory. The findings can also not be generalized to other fields of medicine and cutoffs of response.

4.2. Implication of the findings

4.2.1. Implications for clinical practice

The thesis provided data on the efficacy and tolerability of pharmacological and dietary-supplement treatments for ASD. This information could be suitable to guide evidence-based clinical practice, treatment guidelines and a shared-decision making process. Nevertheless, there were only sparse data in adults, and the findings refer mainly to children/adolescents.

4.2.1.1. Evidence-based clinical practice and shared-decision making for choosing medications for ASD

The number of available medications for ASD is increasing, e.g., 41 pharmacological treatments and 17 dietary-supplements were identified in the current analysis. Psychotropic medications are frequently used in individuals with ASD (Jobski et al. 2017). Current medications differ substantially in their mechanism of action, efficacy and side-effect profiles, as well as in their evidence base. There is a paucity of approved medications with a specific indication for ASD, i.e., aripiprazole, haloperidol and risperidone for irritability/aggression and melatonin for sleep disorders (Fuentes,

Hervás, and Howlin 2021). Other medications are used “off-label” or following the established guidelines for the general population, e.g., for depression, anxiety and ADHD (Fuentes, Hervás, and Howlin 2021; Howes et al. 2018; Rodrigues et al. 2021). Some medications are even used despite the lack of empirical support (Matson et al. 2013). Several factors could influence treatment decisions, e.g., challenging behaviors, caregiver stress, beliefs about ASD, current recommendations, specific needs of the individual and side-effects of the medications (Wilson et al. 2018).

In this context, shared-decision making among clinicians, caregivers and individuals with ASD is necessary for treatment decisions about when and which medication to be used.

The thesis provided the best available evidence on the pharmacotherapy of core symptoms and co-occurring difficulties in ASD, which could inform a shared-decision making process. For example, the evidence could be presented with the use of a digital shared-decision making assistant, e.g., as recently designed for choosing antipsychotics in the treatment of schizophrenia (Siafis et al. 2022a).

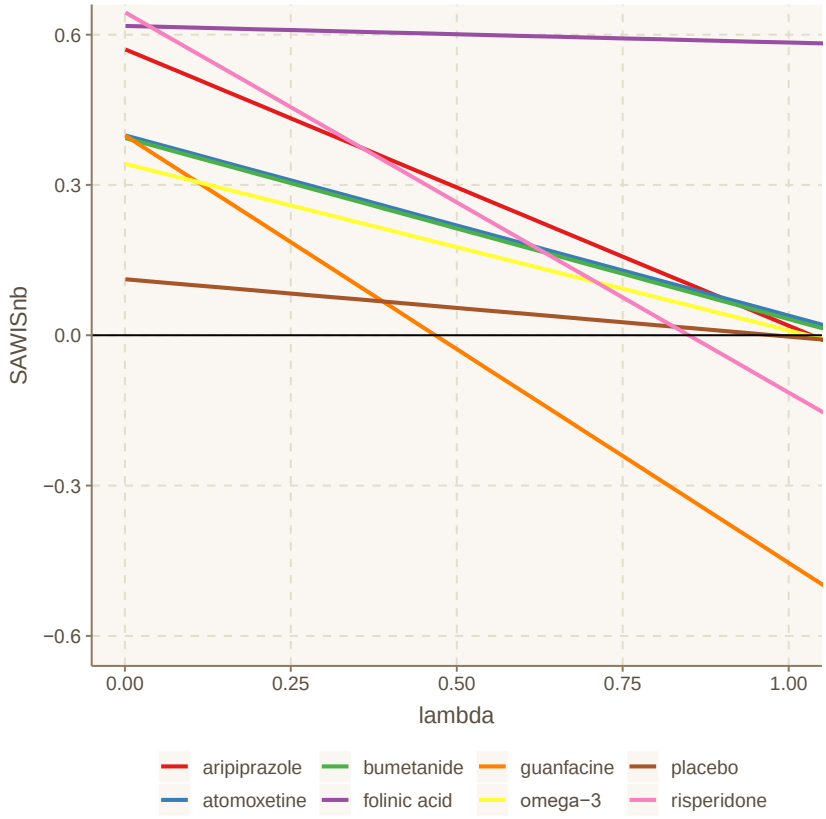
4.2.1.1.1. Core symptoms

The findings further supported current guidelines, e.g., (DGPPN 2021; Fuentes, Hervás, and Howlin 2021; Howes et al. 2018; Hyman et al. 2020; NICE 2021a, 2021b), which recommend against the routine prescription of current medications in individuals with ASD who want medical support for the core symptoms.

In particular, some medications might be effective in improving core symptoms (**Table-2** and **Table-3**). However, the evidence was generally preliminary and of low certainty, as well as the effects were on average small-to-medium, potentially confounded by improvements in co-occurring difficulties and accompanied with a higher risk of adverse events.

Therefore, the advantages of these medications considering their efficacy in the core symptoms and corresponding side-effects seems to be generally small (**Figure-6**), and thus, careful consideration and regular monitoring of the potential benefits and risks is required when these medications are used (Howes et al. 2018).

Figure 6 Trade-off between efficacy in the core symptoms and any adverse events in children/adolescents with ASD. The ranking of treatments was based on a recently developed method (Chiocchia et al. 2022), which is an extension of the spie chart (Daly et al. 2020). Specifically, the standardized area within the spie chart (SAWIS) of efficacy is calculated for each intervention using the P-scores for efficacy in the core symptoms (assuming an equal importance of social-communication difficulties and repetitive behaviors), and similarly for safety using the P-scores of any adverse event. The trade-off of the spie chart ($SAWIS^{NB}$) between efficacy and safety is calculated by the subtraction of the area of the $SAWIS^{efficacy}$ minus λ times the area of $SAWIS^{safety}$. λ is a unit of tolerance of efficacy relatively to safety (λ=0 indicates that efficacy is only important). A positive $SAWIS^{NB}$ indicates a positive trade-off between efficacy and safety given a unit of tolerance to adverse events. Potentially effective medications listed in Table-2 with data for both core symptom domains and any adverse event were presented in this figure.



4.2.1.1.2. Co-occurring difficulties

The findings were also generally in line with the recommendations about the pharmacological treatment of three important co-occurring difficulties, i.e., irritability, ADHD symptoms and anxiety (Anixt et al. 2020), e.g., (Fuentes, Hervás, and Howlin 2021; Howes et al. 2018; Rodrigues et al. 2021)(**Table-4**). Other common co-occurring

conditions, e.g., sleep disorders, gastrointestinal symptoms, and epileptic seizures, were not considered as outcomes of the thesis.

Table 4 Medications for co-occurring difficulties in children/adolescents with ASD.

Co-occurring difficulties	Implications of the findings
Irritability/aggression	Aripiprazole (FDA approved) and risperidone (EMA/FDA approved) showed medium-to-large effect sizes in improving irritability. There were no clear differences between them, but aripiprazole may have a smaller risk for weight gain. Some of their effects may also be maintained in the long-term (Findling et al. 2014; Troost et al. 2005). Due to their adverse events, they could be considered as second-line treatment when behavioral or educational interventions have failed, as well as their use require careful monitoring of their benefit-risk ratio (Howes et al. 2018). On the other hand, the evidence for other antipsychotics was sparse and inconclusive, such as for haloperidol (approved by EMA; evidence from old small and mainly crossover RCTs), lurasidone (one well-powered and negative RCT (Loebel et al. 2016)) etc. Therefore, extreme caution is needed when extrapolating findings from aripiprazole and risperidone to all antipsychotics, e.g., as it has been done in other meta-analysis (D'Alò et al. 2021). Selective serotonin reuptake inhibitors (SSRI), e.g., fluoxetine, citalopram and sertraline, might improve irritability with small-to-medium effect-sizes, yet this finding requires further investigation given that most of the RCTs were not focused on irritability. Some medications may also improve, e.g., sulforaphane and guanfacine, or even worsen irritability, e.g., vitamin-B12 and levetiracetam. Further research is warranted.
ADHD symptoms	Methylphenidate, i.e., the first-line medication for ADHD symptoms after behavioral interventions have failed (Rodrigues et al. 2021), was investigated only in crossover RCTs, which did not provide usable data according to the protocol of the thesis. Nevertheless, previous meta-analysis found a medium-to-large effect for methylphenidate in improving ADHD symptoms (Sturman, Deckx, and van Driel 2017). The second-line medications, atomoxetine and guanfacine showed medium-to-large effect-sizes in improving ADHD symptoms. In addition, guanfacine may be more efficacious based on an indirect comparison and may also improve irritability with a medium effect-size. However, guanfacine was associated with adverse events, e.g., sedation, and could be less tolerable, according also to previous reviews (Cortese et al. 2018; Rodrigues et al. 2021). There was no or limited evidence for other ADHD medications, e.g., amphetamines or clonidine. Aripiprazole and risperidone showed large effect-sizes in improving ADHD symptoms, yet most of the trials focused on irritability. Nevertheless, they could be considered as alternative options when standard interventions have failed (Lamberti et al. 2016). Some medications may also improve ADHD symptoms, e.g., olanzapine, naltrexone, omega-3-fatty acids and sulforaphane. Further research is warranted.
Anxiety or depression	None of the investigated medications were found to improve anxiety or depression.
Other common co-occurring difficulties were not investigated as outcomes in this review, e.g., sleep disorders, epileptic seizures, gastrointestinal issues.	

In order to disseminate the findings to the public and further facilitate evidence-based treatment decisions, lay summaries were written in collaboration with the AIMS-2-

TRIALS communication team and presented in the respective website (AIMS-2-TRIALS 2020, 2022).

4.2.1.2. Nomenclature for psychotropic medications, unclear class effects and the example of antipsychotic drugs

The findings further highlight the need to move from an indication- to a pharmacology-driven nomenclature of psychotropic medications (Zohar et al. 2015).

Current nomenclature for psychotropic medications is rather anachronistic and based on the primary therapeutic indication, e.g., antidepressants and antipsychotics (Zohar et al. 2015). The indication-based nomenclature could be confusing and misleading. In particular, commonly used medications in ASD have another primary indication, e.g., antipsychotics, antidepressants, ADHD medications, antiepileptics. Accordingly, medications that belong to the same pharmacological class could be perceived as having similar properties, i.e., “class effect”, which, however, may not pertain outside the primary therapeutic indication (Fountoulakis et al. 2011). Thus, extra caution is needed when inferring and extrapolating conclusions about medications within an indication-based pharmacological class, especially if the target condition is different from the primary indication of the pharmacological class. To that direction, the adoption and improvement of the Neuroscience-Based Nomenclature (NbN) (Zohar et al. 2015), which aim to classify psychotropic medications based on their pharmacology and mechanism of action, may reduce such misconceptions and facilitate better treatment decisions.

A notable example is antipsychotic drugs for ASD. Antipsychotics act as antagonists or partial agonists to the D_2 receptors and they are the mainstay pharmacological treatment of schizophrenia with small differences in their efficacy (Huhn et al. 2019; Schneider-Thoma et al. 2022). Therefore, they are usually classified into the same

pharmacological class, i.e., “antipsychotics”, a name that have originated from these properties (Carpenter and Davis 2012). This indication-based classification was also frequently used in ASD to draw conclusions, e.g., (D'Alò et al. 2021; Salazar de Pablo et al. 2022). Nevertheless, antipsychotics target also other receptors, e.g., serotonin, acetylcholine, histamine, adrenergic receptors, etc., and differ substantially in their receptor-binding profiles (Siafis, Davis, and Leucht 2021). For this reason, there may be no “class effect” for antipsychotics in ASD.

In particular, two antipsychotics, i.e., aripiprazole and risperidone, were found to be efficacious in ASD by improving irritability (in line with their approved indication), ADHD symptoms, and the two core symptom domains, although the latter effects could be unspecific. On the other hand, there was limited and inconclusive evidence for some antipsychotics, i.e., amisulpride (Dollfus et al. 1992), haloperidol (indicated by EMA for ASD-related aggression, evidence from old small and mainly crossover RCTs), olanzapine (unpublished studies and reporting bias), paliperidone (one RCT from China with concerns in methodology)(Li et al. 2016), and lurasidone (one well-powered and negative RCT)(Loebel et al. 2016), no randomized evidence for other antipsychotics, e.g., quetiapine, ziprasidone and clozapine, and emerging evidence for newer antipsychotics, e.g., cariprazine and brexpiprazole (NCT04174365 2019; NCT05439616 2022). Therefore, caution is needed when extrapolated findings from aripiprazole and risperidone to all antipsychotics.

4.2.2. Implications for clinical research

The findings of the thesis could also have implications for clinical research by identifying medications for further investigation, and insights for future clinical trials and systematic reviews.

4.2.2.1. Medications for further investigation

The quality of the evidence for the majority of the medications was very low or low, and thus, further research is generally required. Nevertheless, some medications could be discussed in more detail (**Table-5**), and given the lack of effective medications for ASD, replication of positive findings should be prioritized (McCracken et al. 2021).

Table 5 Medications that require further investigation. The mechanism of action of medications was extracted from the Neuroscience-Based-Nomenclature (NbN)(Zohar et al. 2015) and the IUPHAR/BPS (Harding et al. 2022). D2: dopamine 2 receptor, 5-HT1A/2/2A: serotonin 1A, 2 and 2A receptors, NET: norepinephrine transporter, alpha-2: norepinephrine alpha-2 receptor; ALDH5A1: aldehyde dehydrogenase 5 A1, HDAC: histone deacetylase; FFAR: free fatty acid receptor, RXRa: retinoid X receptor a; GSK-3b: glycogen synthase kinase 3 beta; CB: cannabinoid receptors; NMDA: N-methyl-D-aspartate receptor TRP: transient receptor potential cation channel; CBD: cannabidiol, CBVD: cannabidivarin; THC: Δ9-tetrahydrocannabinol

Pharmacological class	**Mechanism of action**	**Medication**	**Comments for further investigation**
Antipsychotic	D_2, 5-HT_{1A} partial agonist, 5-HT_{2A} antagonist	Aripiprazole	Approved indication for irritability associated with ASD (FDA). Potential role in the management of ADHD symptoms. Effects on core symptoms can be unspecific and subsequent to the improvement of irritability and/or ADHD symptoms. Further investigation on the benefit-risk ratio.
Antipsychotic	D_2, 5-HT_2, alpha-2 antagonist	Risperidone	Approved indication for irritability associated with ASD (EMA, FDA). Potential role in the management of ADHD symptoms. Effects on core symptoms can be unspecific and subsequent to the improvement of irritability and/or ADHD symptoms. RCT focusing on repetitive behaviors is not yet published (NCT01171937 2010). Further investigation on the benefit-risk ratio, e.g., therapeutic window of 15-25µg/l (Kloosterboer et al. 2021) and a higher response in individuals with a higher severity of symptoms at baseline (Arnold et al. 2010; Levine et al. 2016).
Antipsychotic	D_2 antagonist	Haloperidol	Approved indication for severe aggression associated with ASD (EMA). Evidence was based on small, old and mainly crossover RCTs. Nevertheless, further investigation may not be warranted given that haloperidol is associated with dyskinesia and may be less effective than risperidone (Miral et al. 2008).
ADHD medication	DAT, NET inhibitor	Methylphenidate	Methylphenidate is generally recommended as the first-line medication for ADHD symptoms (Rodrigues et al. 2021), yet there was limited eligible data from five small and crossover RCTs. A previous meta-

			analysis that used a different method to analyze data from crossover studies found that methylphenidate may improve ADHD symptoms with a medium-to-large effect size, yet the effects on other symptom domains were unclear (Sturman, Deckx, and van Driel 2017). A more recent open-label RCT from India found generally similar effects between risperidone and methylphenidate in improving core symptoms and co-occurring difficulties in children/adolescents with ASD (Mahajan, Arun, and Chauhan 2022).
ADHD medication	DAT, NET inhibitor	Amphetamines	Amphetamines can be a second-line medication for ADHD symptoms (Cortese et al. 2018; Rodrigues et al. 2021), yet there was no RCT in ASD.
ADHD medication	NET inhibitor	Atomoxetine	Effects on repetitive behaviors could be unspecific and subsequent to the improvement of ADHD symptoms.
ADHD medication	alpha-2 agonist	Guanfacine	Effects on repetitive behaviors could be unspecific and subsequent to the improvement of irritability and/or ADHD symptoms.
Antidepressant	SERT inhibitor	Citalopram, fluoxetine, fluvoxamine, sertraline etc.	No evidence on improving core symptoms, anxiety or depression in children/adolescents. Potential efficacy for irritability, which was yet not the focus in the majority of the RCTs. Potential efficacy for repetitive behaviors in adults based on the findings of two small RCTs.
Antiepileptic/mood -stabilizer	unclear, ALDH5A1, HDAC inhibition	Valproate	Potential efficacy for repetitive behaviors, and mixed evidence on irritability. Further research is generally warranted about antiepileptics in individuals with ASD and comorbid epilepsy, which is prevalent in about 10% (Liu et al. 2022), given that there was limited evidence, and some of them, e.g., levetiracetam and zonisamide, may have a higher risk of behavioral and psychiatric adverse events (Chen et al. 2017; Wasserman et al. 2006).
Sleep medication	MT agonist	Melatonin	Approved indication for sleep disorders associated with ASD (EMA). Potential effects on other co-occurring difficulties, e.g., irritability and ADHD symptoms, can be unspecific and subsequent to the improvement of sleep problems (not an outcome of this analysis).
Neuropeptide	OT, VP agonist	Oxytocin	No effects on core symptoms and co-occurring difficulties in children/adolescents. Potential efficacy for repetitive behaviors in adults, yet the evidence was mainly based in high-functioning men. Mixed evidence in improving social-communication difficulties in adults.
Neuropeptide	VP, OT agonist	Vasopressin	Potential positive effects based on a small trial in children/adolescents, which was however excluded in this analysis

			due to unconcealed allocation (Parker et al. 2019).
Experimental	VP_{1A} antagonist, 5-HT_{2B} antagonist	Balovaptan	No effects based on large trials in children/adolescents and adults, yet small improvements in quality of life were noted.
Loop diuretic	Na-K-Cl symporter inhibitor	Bumetanide	Potential efficacy for repetitive behaviors. Two recent large RCTs were negative, but no data were available to be included in the analysis (Crutel et al. 2021). A more recent RCT found that bumetanide could improve irritability in children/adolescents with ASD, ADHD and/or epilepsy and sensory abnormalities (van Andel et al. 2022).
ALS medication	High voltage-activated calcium channel inhibitor	Riluzole	Potential effects on irritability, but data were imprecise and based on a presentation of an unpublished trial (NCT01661855 2017). Possibly not efficacious for irritability that did not respond to aripiprazole or risperidone based on the findings of a very small RCT (Wink et al. 2018).
Anti-dementia medication	NMDA antagonist	Memantine	Potential effects on neurocognition (not an outcome of this analysis) (Soorya et al. 2021). No or mixed evidence on the efficacy for core symptoms and co-occurring difficulties. Further RCTs are ongoing, e.g., (NCT03553875 2018).
Experimental	$GABA_B$ agonist	Arbaclofen	No or mixed evidence on the efficacy for core symptoms and co-occurring difficulties. Further RCTs are ongoing, e.g., (Parellada et al. 2021).
Experimental	GSK-3b inhibitor	Tideglusib	Potential effects on core symptoms, but data were imprecise and based on a presentation of an unpublished trial (Anagnostou et al. 2018).
Experimental	Tyrosine hydroxylase inhibitor	L1-79	Potential effects on core symptoms, but data were imprecise and based on a presentation of an unpublished trial (Rothman 2020). Further RCTs are ongoing, e.g., (NCT05067582 2021).
Drug used in addictive disorders	Opioid μ, κ and δ antagonist	Naltrexone	Potential efficacy for ADHD symptoms. No evidence for core symptoms (Roy et al. 2015).
Cannabinoids	$CB_{1/2}$ partial agonist (THC), CB_1 negative allosteric modulator (CBD, CBDV), TRP activation (CBD, CBDV)	Cannabinoids (THC, CBD, CBDV, etc.)	Mixed evidence on the efficacy for core symptoms and co-occurring difficulties. A more recent RCT was also inconclusive (da Silva Junior et al. 2022). Further RCTs are ongoing, etc., (NCT03202303 2017; NCT04745026 2021).
Dietary-supplement	unclear, FFAR, RXRa agonist	Omega-3-fatty acids	Potential efficacy for social-communication difficulties, irritability and ADHD symptoms. Further and larger RCTs are ongoing and necessary, e.g., (NCT01260961 2010)
Dietary-supplement	co-factors and/or substrates in single-carbon transfer reactions concerning four interconnected pathways, i.e., folate, methylation, tetrahydrobiopterin and	Folinic acid, vitamin-B_{12}, sapropterin, cysteine-rich whey protein	There was mixed evidence on the efficacy for core symptoms. Further RCTs are ongoing and necessary, e.g., for folinic acid (NCT02839915 2016). Vitamin-B_{12} may worsen irritability and ADHD symptoms.

	glutathione metabolism (Delhey et al. 2018)		
Dietary-supplement	Unclear, enhance heat shock and antioxidant response via Keap1-Nrf2 pathway (Gan et al. 2010) (Gan et al. 2010)	Sulforaphane	Mixed evidence on the efficacy for core symptoms and co-occurring difficulties. A more recent RCT from China had also inconclusive findings (Smith et al. 2020). Nevertheless, medium-to-large effect-sizes were indicated for some outcomes, and further research is warranted.
Dietary-supplement	Unclear, production of neuroactive substances and restoration of gut-brain axis (Dinan, Stanton, and Cryan 2013)	Probiotics	Potential efficacy for gastrointestinal symptoms (not an outcome of this analysis)(Arnold et al. 2019). Inconclusive evidence for the core symptoms. Further RCTs are ongoing, e.g., (Zhang et al. 2022).
Dietary-supplement	Unclear, antiglycating and antioxidative properties (Reddy et al. 2005)	Carnosine	Potential efficacy for sleep disorders (not an outcome of this analysis)(Mehrazad-Saber, Kheirouri, and Noorazar 2018). No or mixed evidence on the efficacy for core symptoms.
Dietary-supplement	NMDA partial agonist	Dimethylglycine, D-cycloserine	No or mixed evidence in improving core symptoms. No RCT for D-cycloserine was eligible for this analysis. Another RCT found that D-cycloserine could maintain the effects of social skills training, despite there was no difference with placebo in the short-term (Minshawi et al. 2016; Wink et al. 2017).

It should be noted again that medications could only be considered as part of a multidisciplinary support and care for individuals with ASD (Lai et al. 2020). Therefore, there is a tremendous need to evaluate their role within multimodal interventions, e.g., potential synergistic and potentiation effects between pharmacological and non-pharmacological interventions (Díaz-Caneja et al. 2021; McCracken et al. 2021). For example, some medications could theoretically enhance the effects of behavioral or education interventions, such as indirectly by improving challenging behaviors and subsequently allowing more participation in social interaction, e.g., risperidone (McDougle et al. 2005; Scahill et al. 2012), or directly by potentiating learning, e.g., D-cycloserine (Wink et al. 2017).

However, there is a paucity of RCTs on this topic, and the latter were excluded from this analysis, e.g., D-cycloserine and social skills training (Minshawi et al. 2016; Wink et al. 2017), melatonin and cognitive behavioral therapy (CBT) for insomnia (Cortesi

et al. 2012), 3,4-methylenedioxymethamphetamine (MDMA) assisted psychotherapy for social anxiety (Danforth et al. 2018), the combination of atomoxetine (Handen et al. 2015), bumetanide (Du et al. 2015), haloperidol (Campbell et al. 1978), memantine (Karahmadi et al. 2018), risperidone (Aman et al. 2009; Rezaei et al. 2018; Scahill et al. 2012), and levetiracetam (Wang, Jiang, and Tang 2017) with behavioral or educational interventions.

4.2.2.2. Insights for future clinical trials and meta-analysis

The findings of the thesis could give insights in the design of future clinical trials and meta-analysis, especially regarding heterogeneity in ASD, outcome measures, sample size and other design characteristics.

4.2.2.2.1. Heterogeneity

As noted in the introduction (see 1.1.), ASD is characterized by extreme heterogeneity, which could be a major challenge in drug discovery (Díaz-Caneja et al. 2021; McCracken et al. 2021). In particular, ASD may not be a unitary diagnosis, but it could consist of conditions sharing the two core symptom domains, albeit having a different neurobiology, and subsequently a different phenotypic expression and treatment response, etc. (McCracken et al. 2021).

There are efforts to disentangle heterogeneity, facilitate the stratification of individuals and prediction of treatment response across different medications, such as based on the severity of sensory abnormalities and other symptom subtypes, the presence of comorbidities, e.g., monogenetic disorder, epilepsy etc., as well as emerging objective biomarkers, e.g., indices in electroencephalograph (EEG), eye-tracking, neuroimaging, magnetic resonance spectroscopy (MRS), blood levels of serotonin, etc. (Díaz-Caneja et al. 2021; McCracken et al. 2021; McPartland et al. 2020). For example, the efficacy of bumetanide could be higher in individuals with a higher baseline severity of sensory

abnormalities and repetitive behaviors, as well as predicted by EEG and MRS indices of baseline excitatory-inhibitory imbalance (Dai et al. 2021; Juarez-Martinez et al. 2021; van Andel et al. 2022; Zhang et al. 2020). Nevertheless, these findings warrant further replication (Geertjens et al. 2022).

Heterogeneity in ASD could also be reflected by a substantial interindividual variability of treatment response, which warrants further investigation (Baribeau, Vorstman, and Anagnostou 2022). Substantial interindividual variability, e.g., as suggested by a higher variance of symptom change in the medication compared with the placebo group, could lead to a small effect-size on average, yet it could still indicate that there may be a subgroup of individuals that respond to the medication. Thus, meta-analysis of variation could further elucidate the potential presence of interindividual variability of treatment response and facilitate the replicability and generalizability of the findings (Usui et al. 2021). Nevertheless, such interpersonal variability of treatment response has not been found in schizophrenia, another condition with substantial heterogeneity in terms of neurobiology (Winkelbeiner et al. 2019).

Accordingly, there is also a great need to identify mediators and moderators of treatment response (Díaz-Caneja et al. 2021). Thus, further investigation is warranted with individual-participant-data (IPD) meta-analysis, which could allow for a fine-grained analysis of participant-related factors (Debray et al. 2015), e.g., age, sex, baseline severity etc. Apart from exploring heterogeneity, such analyses could further elucidate the benefit-risk ratio of current medications. For example, post-hoc analyses of the NIMH-RUPP trial of risperidone found that a higher baseline severity of symptoms was a moderator of response (Arnold et al. 2010; Levine et al. 2016), while compliance and less weight gain mediated positively the treatment response (Arnold et al. 2010). Thus, a positive benefit-risk ratio could be suggested for risperidone in

participants with a higher severity of illness and within the alleged therapeutic window of 15-25µg/l for acceptable efficacy and weight gain (Kloosterboer et al. 2021).

4.2.2.2.2. Outcome measures

As noted previously (see 2.5.2.1 and 4.1.1.3), there is still no agreement on the selection of outcome measures in clinical trials of ASD (Aman et al. 2004; McCracken et al. 2021). There is a long list of more than 300 outcome measures, from which only a few have been used in at least 5% of current trial (Bolte and Diehl 2013; Provenzani et al. 2020) and some of them have been recommended by expert consensus, e.g., (McCracken et al. 2021). Therefore, diverse, yet validated, outcome measures were identified and used in the thesis, and they were classified into the different outcome domains. This classification could be used as guidance for future meta-analysis, e.g., a systematic review and network meta-analysis on psychological and psychosocial interventions is planned. Nevertheless, it is unclear whether there is a disagreement among the treatment effects measured by different scales, and further research is necessary (McCracken et al. 2021).

Accordingly, commonly used outcome measures should be considered in clinical trials in order to ease the interpretation, comparability and combination of their findings (McCracken et al. 2021). In particular, the CGI scales are recommended to be used in all trials of ASD, given the phenotypic heterogeneity of the condition and the lack of optimal scale (Aman et al. 2004; Provenzani et al. 2020). The thesis by validating an imputation method to estimate the number of responders in CGI-I could further facilitate the comparability and combination of data across studies.

Another issue is that current outcome measures are mainly based on informant reports, e.g., self-, caregiver- and clinician-ratings, which could be prone to expectancy biases

and could vary importantly across the different informants (Jacob et al. 2022; Möricke, Buitelaar, and Rommelse 2016).

In particular, the thesis identified that placebo-effects in social-communication difficulties might be higher in caregiver compared with clinician ratings (Siafis et al. 2020). This could indicate the presence of placebo-by-proxy effects, such as due to parental expectations, satisfaction with care, participation effects and observation within the trial (Jones et al. 2017; Weimer et al. 2013). Participation effects and unspecific fluctuations of symptoms might be reduced after a screening phase of an adequate duration, which could also be combined with a washout from previous psychotropic medications and/or a placebo lead-in phase, yet there was limited data to support these recommendations (Siafis et al. 2020).

Furthermore, placebo-by-proxy effects could be reflected by modification of the caregiver perception of symptoms (i.e., improvements in caregiver ratings), and/or modification of the behavior of the caregiver that may influence the symptoms and behaviors of the child (i.e., improvements in clinician ratings and other outcome measures) (Weimer et al. 2013; Whalley and Hyland 2013). The latter phenomenon might also partially explain the substantial placebo-effects across different outcome measures in the recent balovaptan trials (Baribeau, Vorstman, and Anagnostou 2022). Nevertheless, clinician ratings could also be prone to placebo-effects (Masi et al. 2015), e.g., as observed in CY-BOCS in three large antidepressant trials (Herscu et al. 2020; King et al. 2009; Reddihough et al. 2019).

Therefore, there is a certain need of more objective, developmentally sensitivity, and psychometrically-sound outcome measures (Baribeau, Vorstman, and Anagnostou 2022; Díaz-Caneja et al. 2021; Jacob et al. 2022; McCracken et al. 2021), such as by utilizing digital health technology, e.g., the multimodal data capture system Janssen

Autism Knowledge Engine (JAKE®)(Ness et al. 2019) and smartphone questions (Jones et al. 2018), emerging biomarkers, e.g., eye-tracking-based measures (Frazier et al. 2018), novel rating scales, e.g., Brief Observation of Social Communication Change (BOSCC) (Grzadzinski et al. 2016), and improving currently available scales, e.g., applying item response theory to ABC (Haem et al. 2020) or SRS (Sturm et al. 2017).

New approaches to incorporate the perspectives of the individual and caregivers are also needed, such as by utilizing measures of self-reported outcomes and quality of life, e.g., the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) Autism Battery, and measures of individual- or parent-defined target problems (Arnold et al. 2003; Scahill et al. 2017). This is particularly important given that the priorities of individuals with ASD and their families may differ from those of researchers, e.g., anxiety/mood symptoms are more often highlighted as treatment targets by caregivers in comparison with repetitive behaviors (McCracken et al. 2021).

4.2.2.2.3. Sample size and other trial design characteristics

A major limitation of current RCTs is their small sample sizes (about 20-80 participants) and their considerable lack of statistical power to detect small-to-medium effect-sizes, which are expected for the core symptoms of ASD (McCracken et al. 2021). There is no doubt that larger and adequately powered RCTs are needed, yet some concerns should be considered with larger sample sizes, given that they could increase heterogeneity and may be associated with higher placebo-effects, as also found in this analysis (Siafis et al. 2020).

In particular, recruitment of participants could be challenging in larger RCTs (Carlisle et al. 2015), and even more complicated in ASD (Bent and Hendren 2010). Recruitment pressure necessitates the utilization of more study sites, often non-

academic private sites with less experience in ASD (Siafis et al. 2020), loosening of the eligibility criteria, e.g., (Reddihough et al. 2019), and online recruitment (Brøgger-Mikkelsen et al. 2020; Jacob et al. 2022).

Multicenter trials are generally associated with smaller effect-sizes (Dechartres et al. 2011), and higher placebo-effects in ASD (Siafis et al. 2020). Among others, multicenter trials tend to include more heterogenous populations and less experienced research teams, and have additional practical challenges. For example, increased heterogeneity of the included participants could be associated with smaller effect-sizes as discussed above (see 4.2.2.2.1). In addition, less experienced teams may show higher placebo-effects and lower inter-rater reliability, e.g., for outcome measures that require extensive training such as VABS (Jacob et al. 2022). A low inter-rater reliability could further increase the variability of measurements and lead to smaller effect-sizes. Practical issues and protocol deviations could also for example decrease the inter-rater reliability, e.g., changes of raters in the arbaclofen trial (Veenstra-VanderWeele et al. 2017). Therefore, careful selection, proper training and monitoring of the sites is necessary, and the utilization of centralized experienced raters (Kobak et al. 2010) could be preferred when possible.

Furthermore, non-academic private sites often recruit competitively and utilize time-efficient recruitment methods, such as online recruitment, which was associated with higher placebo-effects in the balovaptan trials (Jacob et al. 2022), potentially due to the proactive research attitudes, higher expectancies and specific characteristics of web-referenced participants (Jacob et al. 2022; Rødgaard et al. 2022). Another issue is that competitive recruitment could lead to inflation of baseline scores of clinician-rating scales when a minimum score of baseline severity is required for inclusion (Mundt et al. 2007; Rutherford and Roose 2013). The use of a minimum threshold of

symptom severity as inclusion criteria could also be associated with regression to the mean and higher placebo-effects, e.g., as noted in ASD (Siafis et al. 2020). These effects could be partially reduced by using different rating scales for inclusion and as primary outcomes (Parellada et al. 2012) or by a blinded analysis excluding participants with baseline scores close to the minimum required threshold (Mancini et al. 2014).

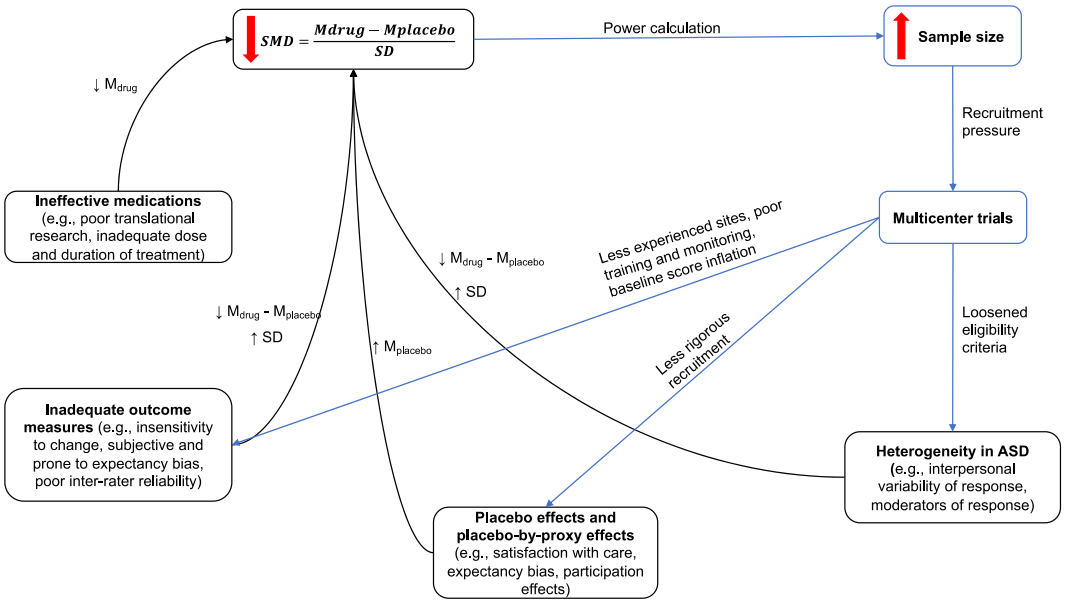
Other design characteristics should also be considered, yet they are beyond the focus of thesis, e.g., younger age of participants to test potentially disease-modifying effects of medications, longer duration of treatment to observe changes in the core symptoms, randomized discontinuation trials and follow-up visits to investigate maintenance of effects, issues of crossover trials, careful selection of dose, appropriate blinding etc. (Díaz-Caneja et al. 2021; EMA 2018; McCracken et al. 2021).

The aforementioned issues could be summarized in a theoretical model for the failure of late-stage clinical trials (**Figure-7**). In particular, challenges in drug development, e.g., ineffective medications, inadequate outcome measures, high placebo-effects and substantial heterogeneity in ASD, could lead to smaller effect-sizes, which would subsequently necessitate larger sample sizes and multicenter trials, which issues could further decrease the observed effect-sizes, and thus, creating a vicious circle between sample sizes and small effect-sizes. A similarly model has been suggested in schizophrenia (Leucht et al. 2017). Nevertheless, further research is necessary to elucidate the mechanisms of this model.

Therefore, there is a need to improve multicenter RCTs in ASD. Accordingly, one of the objectives of the AIMS-2-TRIALS (see 1.5) was to build a network of academia, pharmaceutical industry, autism organizations and regulatory agencies in order to facilitate the execution of multicenter trials, proper training and research experience

acquisition, as well as recruitment of well-characterized individuals with ASD (Díaz-Caneja et al. 2021).

Figure 7 Theoretical model for the failure of late-stage clinical trials in ASD combining the challenges in drug development and the vicious circle between sample size and small effect-sizes. Ineffective medications, inadequate outcome measures, substantial heterogeneity in ASD and high placebo-effects could lead to smaller effect-sizes for the difference between medication and placebo. Smaller effect-sizes would require larger sample sizes in order to be detected. Larger sample sizes could necessitate multicenter trials, which could further decrease the observed effect-size. SMD: standardized mean difference; M_{drug}: mean improvement of symptoms in the medication group; $M_{placebo}$: mean improvement of symptoms in the placebo group; SD: pooled standard deviation of the improvement



4.3. Conclusions

There is still no approved medication for the core symptoms of ASD, and previous late-stage clinical trials failed to identify effective medications. Thus, the aim of the thesis was to inform evidence-based pharmacotherapy and drug development in ASD by providing an up-to-date and comprehensive synthesis of 203 RCTs.

First, some medications, e.g., aripiprazole and risperidone, might be effective for the core symptoms and/or co-occurring difficulties of ASD, albeit associated with adverse events. Nevertheless, the evidence was generally preliminary and with low certainty.

Therefore, routine prescription of medications for the core symptoms cannot be recommended and further investigation is necessary.

Second, the magnitude of placebo-effects was considerable and predictors of higher placebo-effects were identified, e.g., caregiver-ratings and larger trials. Based on these findings and the results of more recent trials, a theoretical model for the failure of late-stage trials was described by combining challenges in drug development and a vicious circle between small effect-sizes and sample sizes.

Third, a method to estimate the number of responders from continuous data of CGI-I was validated, and thus, facilitating the combination and comparability of findings across trials.

The findings and subsequently the implications of the thesis would hopefully provide a stepping stone to improve the support and care for individuals with ASD and their families.

5. References

Abidin, R.R. 1983. 'Parenting Stress Index: Manual, Administration Booklet,[and] Research Update'.

AIMS-2-TRIALS. 2020. 'Understanding placebo response in clinical trials with autistic people', Accessed 14.10.2020. https://www.aims-2-trials.eu/2020/10/14/placebo-response/.

———. 2022. 'Reviewing the evidence for medications and dietary supplements to support autistic people. ', Accessed 10.06.2022. https://www.aims-2-trials.eu/2022/06/20/reviewing-the-evidence-for-medications-and-dietary-supplements-to-support-autistic-people/.

Aman, M.G., McDougle, C.J., Scahill, L., Handen, B., Arnold, L.E., Johnson, C., Stigler, K.A., Bearss, K., Butter, E., Swiezy, N.B., Sukhodolsky, D.D., Ramadan, Y., Pozdol, S.L., Nikolov, R., Lecavalier, L., Kohn, A.E., Koenig, K., Hollway, J.A., Korzekwa, P., Gavaletz, A., Mulick, J.A., Hall, K.L., Dziura, J., Ritz, L., Trollinger, S., Yu, S., Vitiello, B., and Wagner, A. 2009. 'Medication and parent training in children with pervasive developmental disorders and serious behavior problems: results from a randomized clinical trial', *J Am Acad Child Adolesc Psychiatry*, 48: 1143-54.

Aman, M.G., Novotny, S., Samango-Sprouse, C., Lecavalier, L., Leonard, E., Gadow, K.D., King, B.H., Pearson, D.A., Gernsbacher, M.A., and Chez, M. 2004. 'Outcome measures for clinical drug trials in autism', *CNS spectrums*, 9: 36-47.

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

Ameis, S.H., Kassee, C., Corbett-Dick, P., Cole, L., Dadhwal, S., Lai, M.C., Veenstra-VanderWeele, J., and Correll, C.U. 2018. 'Systematic review and guide to management of core and psychiatric symptoms in youth with autism', *Acta Psychiatr Scand*, 138: 379-400.

Anagnostou, E. 2018. 'Clinical trials in autism spectrum disorder: evidence, challenges and future directions', *Curr Opin Neurol*, 31: 119-25.

Anagnostou, E., Bennett, T.A., Thorpe, K., and Nicolson, R. 2018. "A Phase 2 Randomized, Placebo-Controlled Trial of Tideglusib, an Orally Administered GSK-3 Beta Inhibitor, in the Treatment of Adolescents With ASD." In.: AACAP.

Anagnostou, E., Jones, N., Huerta, M., Halladay, A.K., Wang, P., Scahill, L., Horrigan, J.P., Kasari, C., Lord, C., and Choi, D. 2015. 'Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder', *Autism*, 19: 622-36.

Anixt, J.S., Murray, D.S., Coury, D.L., Kuhlthau, K.A., Eskra, D., Seide, J., Kelly, A., Hess, A., Lipkin, P.H., Law, J.K., Fedele, A., and Lannon, C. 2020. 'Improving Behavior Challenges and Quality of Life in the Autism Learning Health Network', *Pediatrics*, 145: S20-s29.

APA. 1980. *Diagnostic and statistical manual of mental disorders: DSM-III* (American psychiatric association Washington, DC).

Arnold, L.E., Farmer, C., Kraemer, H.C., Davies, M., Witwer, A., Chuang, S., DiSilvestro, R., McDougle, C.J., McCracken, J., Vitiello, B., Aman, M.G., Scahill, L., Posey, D.J., and Swiezy, N.B. 2010. 'Moderators, mediators, and other predictors of risperidone response in children with autistic disorder and irritability', *J Child Adolesc Psychopharmacol*, 20: 83-93.

Arnold, L.E., Luna, R.A., Williams, K., Chan, J., Parker, R.A., Wu, Q., Hollway, J.A., Jeffs, A., Lu, F., Coury, D.L., Hayes, C., and Savidge, T. 2019. 'Probiotics for Gastrointestinal Symptoms and Quality of Life in Autism: A Placebo-Controlled Pilot Trial', *J Child Adolesc Psychopharmacol*, 29: 659-69.

Arnold, L.E., Vitiello, B., McDougle, C., Scahill, L., Shah, B., Gonzalez, N.M., Chuang, S., Davies, M., Hollway, J., and Aman, M.G. 2003. 'Parent-defined target symptoms respond to risperidone in RUPP autism study: customer approach to clinical trials', *Journal of the American Academy of Child & Adolescent Psychiatry*, 42: 1443-50.

Autistica. 2016. 'Your questions: shaping future autism research'. https://www.autistica.org.uk/downloads/files/Autism-Top-10-Your-Priorities-for-Autism-Research.pdf.

Bakbergenuly, I., Hoaglin, D.C., and Kulinskaya, E. 2019. 'Pitfalls of using the risk ratio in meta‐analysis', *Research synthesis methods*, 10: 398-419.

Balk, E.M., Earley, A., Patel, K., Trikalinos, T.A., and Dahabreh, I.J. 2013. 'Empirical assessment of within-arm correlation imputation in trials of continuous outcomes'.

Baribeau, D., Vorstman, J., and Anagnostou, E. 2022. 'Novel treatments in autism spectrum disorder', *Curr Opin Psychiatry*, 35: 101-10.

Baxter, A.J., Brugha, T.S., Erskine, H.E., Scheurer, R.W., Vos, T., and Scott, J.G. 2015. 'The epidemiology and global burden of autism spectrum disorders', *Psychol Med*, 45: 601-13.

Becker, B.J. 1988. 'Synthesizing standardized mean‐change measures', *British Journal of Mathematical and Statistical Psychology*, 41: 257-78.

Bent, S., and Hendren, R.L. 2010. 'Improving the prediction of response to therapy in autism', *Neurotherapeutics*, 7: 232-40.

Berry-Kravis, E.M., Lindemann, L., Jønch, A.E., Apostol, G., Bear, M.F., Carpenter, R.L., Crawley, J.N., Curie, A., Des Portes, V., Hossain, F., Gasparini, F., Gomez-Mancilla, B., Hessl, D., Loth, E., Scharf, S.H., Wang, P.P., Von Raison, F., Hagerman, R., Spooren, W., and Jacquemont, S. 2018. 'Drug development for neurodevelopmental disorders: lessons learned from fragile X syndrome', *Nat Rev Drug Discov*, 17: 280-99.

Bland, J.M., and Altman, D. 1986. 'Statistical methods for assessing agreement between two methods of clinical measurement', *The lancet*, 327: 307-10.

Bland, J.M., and Altman, D.G. 1999. 'Measuring agreement in method comparison studies', *Statistical methods in medical research*, 8: 135-60.

Bolte, E.E., and Diehl, J.J. 2013. 'Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder', *Journal of autism and developmental disorders*, 43: 2491-501.

Brignardello-Petersen, R., Florez, I.D., Izcovich, A., Santesso, N., Hazlewood, G., Alhazanni, W., Yepes-Nuñez, J.J., Tomlinson, G., Schünemann, H.J., and Guyatt, G.H. 2020. 'GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework', *Bmj*, 371: m3900.

Brøgger-Mikkelsen, M., Ali, Z., Zibert, J.R., Andersen, A.D., and Thomsen, S.F. 2020. 'Online Patient Recruitment in Clinical Trials: Systematic Review and Meta-Analysis', *J Med Internet Res*, 22: e22179.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., and Munafò, M.R. 2013. 'Power failure: why small sample size undermines the reliability of neuroscience', *Nature Reviews Neuroscience*, 14: 365-76.

Campbell, M., Anderson, L.T., Meier, M., Cohen, I.L., Small, A.M., Samit, C., and Sachar, E.J. 1978. 'A comparison of haloperidol and behavior therapy and their interaction in autistic children', *J Am Acad Child Psychiatry*, 17: 640-55.

Carlisle, B., Kimmelman, J., Ramsay, T., and MacKinnon, N. 2015. 'Unsuccessful trial accrual and human subjects protections: an empirical analysis of recently closed trials', *Clin Trials*, 12: 77-83.

Carpenter, W.T., and Davis, J.M. 2012. 'Another view of the history of antipsychotic drug discovery and development', *Molecular Psychiatry*, 17: 1168-73.

Catalá-López, F., Hutton, B., Page, M.J., Driver, J.A., Ridao, M., Alonso-Arroyo, A., Valencia, A., Macías Saint-Gerons, D., and Tabarés-Seisdedos, R. 2022. 'Mortality in Persons With Autism Spectrum Disorder or Attention-Deficit/Hyperactivity Disorder: A Systematic Review and Meta-analysis', *JAMA Pediatr*, 176: e216401.

Chaimani, A., Higgins, J.P.T., Mavridis, D., Spyridonos, P., and Salanti, G. 2013. 'Graphical tools for network meta-analysis in STATA', *PloS one*, 8: e76654-e54.

Chen, B., Choi, H., Hirsch, L.J., Katz, A., Legge, A., Buchsbaum, R., and Detyniecki, K. 2017. 'Psychiatric and behavioral side effects of antiepileptic drugs in adults with epilepsy', *Epilepsy Behav*, 76: 24-31.

Chiocchia, V., Furukawa, T.A., Schneider-Thoma, J., Siafis, S., Cipriani, A., Leucht, S., and Salanti, G. 2022. 'Ranking competing treatments according to the trade-off between benefit and harms on multiple health outcomes', *under development*.

Cohen, J. 1988. 'The effect size', *Statistical power analysis for the behavioral sciences*: 77-83.

Constantino, J.N., and Gruber, C.P. 2012. *Social responsiveness scale: SRS-2* (Western psychological services Torrance, CA).

Cortese, S., Adamo, N., Del Giovane, C., Mohr-Jensen, C., Hayes, A.J., Carucci, S., Atkinson, L.Z., Tessari, L., Banaschewski, T., Coghill, D., Hollis, C., Simonoff, E., Zuddas, A., Barbui, C., Purgato, M., Steinhausen, H.C., Shokraneh, F., Xia, J., and Cipriani, A. 2018. 'Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis', *Lancet Psychiatry*, 5: 727-38.

Cortesi, F., Giannotti, F., Sebastiani, T., Panunzi, S., and Valente, D. 2012. 'Controlled-release melatonin, singly and combined with cognitive behavioural therapy, for persistent insomnia in children with autism spectrum disorders: a randomized placebo-controlled trial', *J Sleep Res*, 21: 700-9.

Crutel, V., Lambert, E., Penelaud, P.F., Albarrán Severo, C., Fuentes, J., Rosier, A., Hervás, A., Marret, S., Oliveira, G., Parellada, M., Kyaga, S., Gouttefangeas, S., Bertrand, M., Ravel, D., and Falissard, B. 2021. 'Bumetanide Oral Liquid Formulation for the Treatment of Children and Adolescents with Autism Spectrum Disorder: Design of Two Phase III Studies (SIGN Trials)', *J Autism Dev Disord*, 51: 2959-72.

D'Alò, G.L., De Crescenzo, F., Amato, L., Cruciani, F., Davoli, M., Fulceri, F., Minozzi, S., Mitrova, Z., Morgano, G.P., Nardocci, F., Saulle, R., Schünemann, H.J., and Scattoni, M.L. 2021. 'Impact of antipsychotics in children and adolescents with autism spectrum disorder: a systematic review and meta-analysis', *Health Qual Life Outcomes*, 19: 33.

da Silva Junior, E.A., Medeiros, W.M.B., Dos Santos, J.P.M., de Sousa, J.M.M., da Costa, F.B., Pontes, K.M., Borges, T.C., Segundo, C.E.N., e Silva, A.H.A., and Nunes, E.L.G. 2022. 'Evaluation of the efficacy and safety of cannabidiol-rich

cannabis extract in children with autism spectrum disorder: randomized, double-blind and controlled placebo clinical trial', *Trends in Psychiatry and Psychotherapy*: 0-0.

Dai, Y., Zhang, L., Yu, J., Zhou, X., He, H., Ji, Y., Wang, K., Du, X., Liu, X., Tang, Y., Deng, S., Langley, C., Li, W.-G., Zhang, J., Feng, J., Sahakian, B.J., Luo, Q., and Li, F. 2021. 'Improved symptoms following bumetanide treatment in children aged 3−6 years with autism spectrum disorder: a randomized, double-blind, placebo-controlled trial', *Science Bulletin*, 66: 1591-98.

Daly, C.H., Mbuagbaw, L., Thabane, L., Straus, S.E., and Hamid, J.S. 2020. 'Spie charts for quantifying treatment effectiveness and safety in multiple outcome network meta-analysis: a proof-of-concept study', *BMC Med Res Methodol*, 20: 266.

Danforth, A.L., Grob, C.S., Struble, C., Feduccia, A.A., Walker, N., Jerome, L., Yazar-Klosinski, B., and Emerson, A. 2018. 'Reduction in social anxiety after MDMA-assisted psychotherapy with autistic adults: a randomized, double-blind, placebo-controlled pilot study', *Psychopharmacology*, 235: 3137-48.

Debray, T.P., Moons, K.G., van Valkenhoef, G., Efthimiou, O., Hummel, N., Groenwold, R.H., and Reitsma, J.B. 2015. 'Get real in individual participant data (IPD) meta-analysis: a review of the methodology', *Res Synth Methods*, 6: 293-309.

Dechartres, A., Boutron, I., Trinquart, L., Charles, P., and Ravaud, P. 2011. 'Single-Center Trials Show Larger Treatment Effects Than Multicenter Trials: Evidence From a Meta-epidemiologic Study', *Annals of Internal Medicine*, 155: 39-51.

Delhey, L.M., Tippett, M., Rose, S., Bennuri, S.C., Slattery, J.C., Melnyk, S., James, S.J., and Frye, R.E. 2018. 'Comparison of Treatment for Metabolic Disorders Associated with Autism:Reanalysis of Three Clinical Trials', *Front Neurosci*, 12: 19.

DerSimonian, R., and Laird, N. 1986. 'Meta-analysis in clinical trials', *Controlled clinical trials*, 7: 177-88.

DGPPN. 2021. "S3-Leitlinie Autismus-Spektrum-Störungen im Kindes-, Jugend- und Erwachsenenalter, Teil 2: Therapie." In.

Dias, S., Welton, N.J., Caldwell, D.M., and Ades, A.E. 2010. 'Checking consistency in mixed treatment comparison meta-analysis', *Stat Med*, 29: 932-44.

Díaz-Caneja, C.M., State, M.W., Hagerman, R.J., Jacquemont, S., Marín, O., Bagni, C., Umbricht, D., Simonoff, E., de Andrés-Trelles, F., Kaale, A., Pandina, G., Gómez-Mancilla, B., Wang, P.P., Cusak, J., Siafis, S., Leucht, S., Parellada, M., Loth, E., Charman, T., Buitelaar, J.K., Murphy, D., and Arango, C. 2021. 'A white paper on a neurodevelopmental framework for drug discovery in autism and other neurodevelopmental disorders', *Eur Neuropsychopharmacol*, 48: 49-88.

Dinan, T.G., Stanton, C., and Cryan, J.F. 2013. 'Psychobiotics: a novel class of psychotropic', *Biol Psychiatry*, 74: 720-6.

Djulbegovic, B., and Guyatt, G.H. 2017. 'Progress in evidence-based medicine: a quarter century on', *The Lancet*, 390: 415-23.

Dollfus, S., Petit, M., Menard, J.F., and Lesieur, P. 1992. 'Amisulpride versus bromocriptine in infantile autism: A controlled crossover comparative study of two drugs with opposite effects on dopaminergic function', *Journal of Autism and Developmental Disorders*, 22: 47-60.

Du, L., Shan, L., Wang, B., Li, H., Xu, Z., Staal, W.G., and Jia, F. 2015. 'A Pilot Study on the Combination of Applied Behavior Analysis and Bumetanide Treatment

for Children with Autism', *Journal of Child and Adolescent Psychopharmacology*, 25: 585-88.
Efthimiou, O., Debray, T.P., van Valkenhoef, G., Trelle, S., Panayidou, K., Moons, K.G., Reitsma, J.B., Shang, A., and Salanti, G. 2016. 'GetReal in network meta-analysis: a review of the methodology', *Res Synth Methods*, 7: 236-63.
Egger, M., Smith, G.D., Schneider, M., and Minder, C. 1997. 'Bias in meta-analysis detected by a simple, graphical test', *Bmj*, 315: 629-34.
Elbourne, D.R., Altman, D.G., Higgins, J.P.T., Curtin, F., Worthington, H.V., and Vail, A. 2002. 'Meta-analyses involving cross-over trials: methodological issues', *International journal of epidemiology*, 31: 140-49.
EMA. 2018. "Guideline on the clinical development of medicinal products for the treatment of Autism Spectrum Disorder (ASD)." In.
Enck, P., Bingel, U., Schedlowski, M., and Rief, W. 2013. 'The placebo response in medicine: minimize, maximize or personalize?', *Nature reviews Drug discovery*, 12: 191-204.
Fallah, M.S., Shaikh, M.R., Neupane, B., Rusiecki, D., Bennett, T.A., and Beyene, J. 2019. 'Atypical Antipsychotics for Irritability in Pediatric Autism: A Systematic Review and Network Meta-Analysis', *J Child Adolesc Psychopharmacol*, 29: 168-80.
Findling, R.L., Mankoski, R., Timko, K., Lears, K., McCartney, T., McQuade, R.D., Eudicone, J.M., Amatniek, J., Marcus, R.N., and Sheehan, J.J. 2014. 'A randomized controlled trial investigating the safety and efficacy of aripiprazole in the long-term maintenance treatment of pediatric patients with irritability associated with autistic disorder', *J Clin Psychiatry*, 75: 22-30.
Fountoulakis, K.N., Gonda, X., Vieta, E., and Rihmer, Z. 2011. 'Class effect of pharmacotherapy in bipolar disorder: fact or misbelief?', *Ann Gen Psychiatry*, 10: 8.
Fraguas, D., Díaz-Caneja, C.M., Pina-Camacho, L., Moreno, C., Durán-Cutilla, M., Ayora, M., González-Vioque, E., de Matteis, M., Hendren, R.L., Arango, C., and Parellada, M. 2019. 'Dietary Interventions for Autism Spectrum Disorder: A Meta-analysis', *Pediatrics*, 144.
Frazier, T.W., Klingemier, E.W., Parikh, S., Speer, L., Strauss, M.S., Eng, C., Hardan, A.Y., and Youngstrom, E.A. 2018. 'Development and Validation of Objective and Quantitative Eye Tracking-Based Measures of Autism Risk and Symptom Levels', *J Am Acad Child Adolesc Psychiatry*, 57: 858-66.
Fuentes, J., Hervás, A., and Howlin, P. 2021. 'ESCAP practice guidance for autism: a summary of evidence-based recommendations for diagnosis and treatment', *European Child & Adolescent Psychiatry*, 30: 961-84.
Furukawa, T.A., Barbui, C., Cipriani, A., Brambilla, P., and Watanabe, N. 2006. 'Imputing missing standard deviations in meta-analyses can provide accurate results', *J Clin Epidemiol*, 59: 7-10.
Furukawa, T.A., Cipriani, A., Barbui, C., Brambilla, P., and Watanabe, N. 2005. 'Imputing response rates from means and standard deviations in meta-analyses', *International clinical psychopharmacology*, 20: 49-52.
Furukawa, T.A., Salanti, G., Atkinson, L.Z., Leucht, S., Ruhe, H.G., Turner, E.H., Chaimani, A., Ogawa, Y., Takeshima, N., and Hayasaka, Y. 2016. 'Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis', *BMJ open*, 6: e010919.

Gan, N., Wu, Y.C., Brunet, M., Garrido, C., Chung, F.L., Dai, C., and Mi, L. 2010. 'Sulforaphane activates heat shock response and enhances proteasome activity through up-regulation of Hsp27', *J Biol Chem*, 285: 35528-36.

Gaugler, T., Klei, L., Sanders, S.J., Bodea, C.A., Goldberg, A.P., Lee, A.B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J., Ripke, S., Sandin, S., Sklar, P., Svantesson, O., Reichenberg, A., Hultman, C.M., Devlin, B., Roeder, K., and Buxbaum, J.D. 2014. 'Most genetic risk for autism resides with common variation', *Nat Genet*, 46: 881-5.

GBD 2019 Dieases and Injuries Collaborators. 2020. 'Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019', *Lancet*, 396: 1204-22.

Geertjens, L., Cristian, G., Haspels, E., Ramautar, J., van der Wilt, G.J., Verhage, M., and Bruining, H. 2022. 'Single-case experimental designs for bumetanide across neurodevelopmental disorders: BUDDI protocol', *BMC Psychiatry*, 22: 452.

Geissbühler, M., Hincapié, C.A., Aghlmandi, S., Zwahlen, M., Jüni, P., and da Costa, B.R. 2021. 'Most published meta-regression analyses based on aggregate data suffer from methodological pitfalls: a meta-epidemiological study', *BMC Medical Research Methodology*, 21: 123.

Gribkoff, V.K., and Kaczmarek, L.K. 2017. 'The need for new approaches in CNS drug discovery: Why drugs have failed, and what can be done to improve outcomes', *Neuropharmacology*, 120: 11-19.

Grzadzinski, R., Carr, T., Colombi, C., McGuire, K., Dufek, S., Pickles, A., and Lord, C. 2016. 'Measuring Changes in Social Communication Behaviors: Preliminary Development of the Brief Observation of Social Communication Change (BOSCC)', *J Autism Dev Disord*, 46: 2464-79.

Guy, W. 1976. 'Clinical global impression', *Assessment manual for Psychopharmacology*: 217-22.

Haem, E., Doostfatemeh, M., Firouzabadi, N., Ghazanfari, N., and Karlsson, M.O. 2020. 'A longitudinal item response model for Aberrant Behavior Checklist (ABC) data from children with autism', *Journal of Pharmacokinetics and Pharmacodynamics*, 47: 241-53.

Handen, B.L., Aman, M.G., Arnold, L.E., Hyman, S.L., Tumuluru, R.V., Lecavalier, L., Corbett-Dick, P., Pan, X., Hollway, J.A., Buchan-Page, K.A., Silverman, L.B., Brown, N.V., Rice, R.R., Jr., Hellings, J., Mruzek, D.W., McAuliffe-Bellin, S., Hurt, E.A., Ryan, M.M., Levato, L., and Smith, T. 2015. 'Atomoxetine, Parent Training, and Their Combination in Children With Autism Spectrum Disorder and Attention-Deficit/Hyperactivity Disorder', *J Am Acad Child Adolesc Psychiatry*, 54: 905-15.

Harding, S.D., Armstrong, J.F., Faccenda, E., Southan, C., Alexander, S.P.H., Davenport, A.P., Pawson, A.J., Spedding, M., Davies, J.A., and Nc, I. 2022. 'The IUPHAR/BPS guide to PHARMACOLOGY in 2022: curating pharmacology for COVID-19, malaria and antibacterials', *Nucleic acids research*, 50: D1282-D94.

Herscu, P., Handen, B.L., Arnold, L.E., Snape, M.F., Bregman, J.D., Ginsberg, L., Hendren, R., Kolevzon, A., Melmed, R., Mintz, M., Minshew, N., Sikich, L., Attalla, A., King, B., Owley, T., Childress, A., Chugani, H., Frazier, J., Cartwright, C., and Murphy, T. 2020. 'The SOFIA Study: Negative Multi-center Study of Low Dose Fluoxetine on Repetitive Behaviors in Children and Adolescents with Autistic Disorder', *J Autism Dev Disord*, 50: 3233-44.

Higgins, J.P. 2011. 'Cochrane handbook for systematic reviews of interventions. Version 5.1. 0 [updated March 2011]. The Cochrane Collaboration', *www cochrane-handbook org*.

Higgins, J.P., Jackson, D., Barrett, J.K., Lu, G., Ades, A.E., and White, I.R. 2012. 'Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies', *Res Synth Methods*, 3: 98-110.

Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., and Sterne, J.A.C. 2011. 'The Cochrane Collaboration's tool for assessing risk of bias in randomised trials', *BMJ*, 343: d5928.

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., and Welch, V.A. 2019. *Cochrane handbook for systematic reviews of interventions* (John Wiley & Sons).

Höfer, J., Hoffmann, F., and Bachmann, C. 2017. 'Use of complementary and alternative medicine in children and adolescents with autism spectrum disorder: A systematic review', *Autism*, 21: 387-402.

Howes, O.D., Rogdaki, M., Findon, J.L., Wichers, R.H., Charman, T., King, B.H., Loth, E., McAlonan, G.M., McCracken, J.T., Parr, J.R., Povey, C., Santosh, P., Wallace, S., Simonoff, E., and Murphy, D.G. 2018. 'Autism spectrum disorder: Consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology', *J Psychopharmacol*, 32: 3-29.

Hozo, S.P., Djulbegovic, B., and Hozo, I. 2005. 'Estimating the mean and variance from the median, range, and the size of a sample', *BMC Medical Research Methodology*, 5: 13.

Huhn, M., Nikolakopoulou, A., Schneider-Thoma, J., Krause, M., Samara, M., Peter, N., Arndt, T., Bäckers, L., Rothe, P., Cipriani, A., Davis, J., Salanti, G., and Leucht, S. 2019. 'Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis', *Lancet*, 394: 939-51.

Hutton, B., Salanti, G., Caldwell, D.M., Chaimani, A., Schmid, C.H., Cameron, C., Ioannidis, J.P., Straus, S., Thorlund, K., Jansen, J.P., Mulrow, C., Catalá-López, F., Gøtzsche, P.C., Dickersin, K., Boutron, I., Altman, D.G., and Moher, D. 2015. 'The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations', *Ann Intern Med*, 162: 777-84.

Hyman, S.L., Levy, S.E., Myers, S.M., and COUNCIL ON CHILDREN WITH DISABILITIES, S.O.D.A.B.P. 2020. 'Identification, Evaluation, and Management of Children With Autism Spectrum Disorder', *Pediatrics*, 145.

Jacob, S., Anagnostou, E., Hollander, E., Jou, R., McNamara, N., Sikich, L., Tobe, R., Murphy, D., McCracken, J., Ashford, E., Chatham, C., Clinch, S., Smith, J., Sanders, K., Murtagh, L., Noeldeke, J., and Veenstra-VanderWeele, J. 2022. 'Large multicenter randomized trials in autism: key insights gained from the balovaptan clinical development program', *Molecular Autism*, 13: 25.

Jobski, K., Höfer, J., Hoffmann, F., and Bachmann, C. 2017. 'Use of psychotropic drugs in patients with autism spectrum disorders: a systematic review', *Acta Psychiatr Scand*, 135: 8-28.

Johnston, B.C., Thorlund, K., Schünemann, H.J., Xie, F., Murad, M.H., Montori, V.M., and Guyatt, G.H. 2010. 'Improving the interpretation of quality of life evidence in meta-analyses: the application of minimal important difference units', *Health Qual Life Outcomes*, 8: 116.

Jones, R.M., Carberry, C., Hamo, A., and Lord, C. 2017. 'Placebo-like response in absence of treatment in children with Autism', *Autism Res*, 10: 1567-72.

Jones, R.M., Tarpey, T., Hamo, A., Carberry, C., and Lord, C. 2018. 'Smartphone measures of day-to-day behavior changes in children with autism', *NPJ Digit Med*, 1: 34.

Juarez-Martinez, E.L., Sprengers, J.J., Cristian, G., Oranje, B., van Andel, D.M., Avramiea, A.-E., Simpraga, S., Houtman, S.J., Hardstone, R., Gerver, C., Jan van der Wilt, G., Mansvelder, H.D., Eijkemans, M.J.C., Linkenkaer-Hansen, K., and Bruining, H. 2021. 'Prediction of Behavioral Improvement Through Resting-State Electroencephalography and Clinical Severity in a Randomized Controlled Trial Testing Bumetanide in Autism Spectrum Disorder', *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.

Karahmadi, M., Tarrahi, M.J., Vatankhah Ardestani, S.S., Omranifard, V., and Farzaneh, B. 2018. 'Efficacy of Memantine as Adjunct Therapy for Autism Spectrum Disorder in Children Aged <14 Years', *Adv Biomed Res*, 7: 131.

Kelley, K., and Preacher, K.J. 2012. 'On effect size', *Psychol Methods*, 17: 137-52.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., and Pellicano, E. 2016. 'Which terms should be used to describe autism? Perspectives from the UK autism community', *Autism*, 20: 442-62.

Kim, J.Y., Son, M.J., Son, C.Y., Radua, J., Eisenhut, M., Gressier, F., Koyanagi, A., Carvalho, A.F., Stubbs, B., Solmi, M., Rais, T.B., Lee, K.H., Kronbichler, A., Dragioti, E., Shin, J.I., and Fusar-Poli, P. 2019. 'Environmental risk factors and biomarkers for autism spectrum disorder: an umbrella review of the evidence', *Lancet Psychiatry*, 6: 590-600.

King, B.H., Dukes, K., Donnelly, C.L., Sikich, L., McCracken, J.T., Scahill, L., Hollander, E., Bregman, J.D., Anagnostou, E., Robinson, F., Sullivan, L., and Hirtz, D. 2013. 'Baseline factors predicting placebo response to treatment in children and adolescents with autism spectrum disorders: a multisite randomized clinical trial', *JAMA pediatrics*, 167: 1045-52.

King, B.H., Hollander, E., Sikich, L., McCracken, J.T., Scahill, L., Bregman, J.D., Donnelly, C.L., Anagnostou, E., Dukes, K., Sullivan, L., Hirtz, D., Wagner, A., and Ritz, L. 2009. 'Lack of efficacy of citalopram in children with autism spectrum disorders and high levels of repetitive behavior: citalopram ineffective in children with autism', *Arch Gen Psychiatry*, 66: 583-90.

Kloosterboer, S.M., de Winter, B.C.M., Reichart, C.G., Kouijzer, M.E.J., de Kroon, M.M.J., van Daalen, E., Ester, W.A., Rieken, R., Dieleman, G.C., van Altena, D., Bartelds, B., van Schaik, R.H.N., Nasserinejad, K., Hillegers, M.H.J., van Gelder, T., Dierckx, B., and Koch, B.C.P. 2021. 'Risperidone plasma concentrations are associated with side effects and effectiveness in children and adolescents with autism spectrum disorder', *Br J Clin Pharmacol*, 87: 1069-81.

Kobak, K.A., Leuchter, A., DeBrota, D., Engelhardt, N., Williams, J.B., Cook, I.A., Leon, A.C., and Alpert, J. 2010. 'Site versus centralized raters in a clinical depression trial: impact on patient selection and placebo response', *J Clin Psychopharmacol*, 30: 193-7.

Lai, M.C., Anagnostou, E., Wiznitzer, M., Allison, C., and Baron-Cohen, S. 2020. 'Evidence-based support for autistic people across the lifespan: maximising potential, minimising barriers, and optimising the person-environment fit', *Lancet Neurol*, 19: 434-51.

Lamberti, M., Siracusano, R., Italiano, D., Alosi, N., Cucinotta, F., Di Rosa, G., Germanò, E., Spina, E., and Gagliano, A. 2016. 'Head-to-Head Comparison of Aripiprazole and Risperidone in the Treatment of ADHD Symptoms in Children with Autistic Spectrum Disorder and ADHD: A Pilot, Open-Label, Randomized Controlled Study', *Paediatr Drugs*, 18: 319-29.

Lavigne, J.V., Cromley, T., Sprafkin, J., and Gadow, K.D. 2009. 'The Child and Adolescent Symptom Inventory-Progress Monitor: a brief Diagnostic and Statistical Manual of Mental Disorders, -referenced parent-report scale for children and adolescents', *Journal of child and adolescent psychopharmacology*, 19: 241-52.

Lecavalier, L., Wood, J.J., Halladay, A.K., Jones, N.E., Aman, M.G., Cook, E.H., Handen, B.L., King, B.H., Pearson, D.A., and Hallett, V. 2014. 'Measuring anxiety as a treatment endpoint in youth with autism spectrum disorder', *Journal of autism and developmental disorders*, 44: 1128-43.

Leucht, S., Chaimani, A., Cipriani, A.S., Davis, J.M., Furukawa, T.A., and Salanti, G. 2016. 'Network meta-analyses should be the highest level of evidence in treatment guidelines', *Eur Arch Psychiatry Clin Neurosci*, 266: 477-80.

Leucht, S., Chaimani, A., Mavridis, D., Leucht, C., Huhn, M., Helfer, B., Samara, M., Cipriani, A., Geddes, J.R., and Davis, J.M. 2019. 'Disconnection of drug-response and placebo-response in acute-phase antipsychotic drug trials on schizophrenia? Meta-regression analysis', *Neuropsychopharmacology*, 44: 1955-66.

Leucht, S., Leucht, C., Huhn, M., Chaimani, A., Mavridis, D., Helfer, B., Samara, M., Rabaioli, M., Bächer, S., Cipriani, A., Geddes, J.R., Salanti, G., and Davis, J.M. 2017. 'Sixty Years of Placebo-Controlled Antipsychotic Drug Trials in Acute Schizophrenia: Systematic Review, Bayesian Meta-Analysis, and Meta-Regression of Efficacy Predictors', *Am J Psychiatry*, 174: 927-42.

Levine, S.Z., Kodesh, A., Goldberg, Y., Reichenberg, A., Furukawa, T.A., Kolevzon, A., and Leucht, S. 2016. 'Initial severity and efficacy of risperidone in autism: Results from the RUPP trial', *European Psychiatry*, 32: 16-20.

Li, Y.C., Ma, J., Xu, H.M., Yang, G.Y., and Zhang, J.C. 2016. 'EFFICACY AND SAFETY OF PALIPERIDONE AND ARIPIPRAZOLE IN THE TREATMENT OF AUTISM', *CHINESE JOURNAL OF NEW DRUGS*, 25: 1893-97.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., and 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.

Liddell, T.M., and Kruschke, J.K. 2018. 'Analyzing ordinal data with metric models: What could possibly go wrong?', *Journal of Experimental Social Psychology*, 79: 328-48.

Liu, X., Sun, X., Sun, C., Zou, M., Chen, Y., Huang, J., Wu, L., and Chen, W.X. 2022. 'Prevalence of epilepsy in autism spectrum disorders: A systematic review and meta-analysis', *Autism*, 26: 33-50.

Loebel, A., Brams, M., Goldman, R.S., Silva, R., Hernandez, D., Deng, L., Mankoski, R., and Findling, R.L. 2016. 'Lurasidone for the Treatment of Irritability Associated with Autistic Disorder', *J Autism Dev Disord*, 46: 1153-63.

Lord, C., Brugha, T.S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E.J.H., Jones, R.M., Pickles, A., State, M.W., Taylor, J.L., and Veenstra-VanderWeele, J. 2020. 'Autism spectrum disorder', *Nat Rev Dis Primers*, 6: 5.

Lord, C., Charman, T., Havdahl, A., Carbone, P., Anagnostou, E., Boyd, B., Carr, T., de Vries, P.J., Dissanayake, C., Divan, G., Freitag, C.M., Gotelli, M.M., Kasari, C., Knapp, M., Mundy, P., Plank, A., Scahill, L., Servili, C., Shattuck, P., Simonoff, E., Singer, A.T., Slonims, V., Wang, P.P., Ysrraelit, M.C., Jellett, R., Pickles, A., Cusack, J., Howlin, P., Szatmari, P., Holbrook, A., Toolan, C., and McCauley, J.B. 2022. 'The Lancet Commission on the future of care and clinical research in autism', *The Lancet*, 399: 271-334.

Lord, C., Elsabbagh, M., Baird, G., and Veenstra-Vanderweele, J. 2018. 'Autism spectrum disorder', *Lancet*, 392: 508-20.

Luo, Y., Funada, S., Yoshida, K., Noma, H., Sahker, E., and Furukawa, T.A. 2022. 'Large variation existed in standardized mean difference estimates using different calculation methods in clinical trials', *J Clin Epidemiol*.

Mahajan, S., Arun, P., and Chauhan, N. 2022. 'Efficacy of Risperidone and Methylphenidate for Problem Behaviors and Core Symptoms in Autism Spectrum Disorder: A Randomized Trial', *Journal of Indian Association for Child and Adolescent Mental Health*: 09731342221096496.

Mancini, M., Wade, A.G., Perugi, G., Lenox-Smith, A., and Schacht, A. 2014. 'Impact of patient selection and study characteristics on signal detection in placebo-controlled trials with antidepressants', *J Psychiatr Res*, 51: 21-9.

Masi, A., Lampit, A., DeMayo, M.M., Glozier, N., Hickie, I.B., and Guastella, A.J. 2017. 'A comprehensive systematic review and meta-analysis of pharmacological and dietary supplement interventions in paediatric autism: moderators of treatment response and recommendations for future research', *Psychol Med*, 47: 1323-34.

Masi, A., Lampit, A., Glozier, N., Hickie, I.B., and Guastella, A.J. 2015. 'Predictors of placebo response in pharmacological and dietary supplement treatment trials in pediatric autism spectrum disorder: a meta-analysis', *Transl Psychiatry*, 5: e640.

Matson, J.L., Adams, H.L., Williams, L.W., and Rieske, R.D. 2013. 'Why are there so many unsubstantiated treatments in autism?', *Research in Autism Spectrum Disorders*, 7: 466-74.

McCracken, J.T., Anagnostou, E., Arango, C., Dawson, G., Farchione, T., Mantua, V., McPartland, J., Murphy, D., Pandina, G., and Veenstra-VanderWeele, J. 2021. 'Drug development for Autism Spectrum Disorder (ASD): Progress, challenges, and future directions', *Eur Neuropsychopharmacol*, 48: 3-31.

McDougle, C.J., Scahill, L., Aman, M.G., McCracken, J.T., Tierney, E., Davies, M., Arnold, L.E., Posey, D.J., Martin, A., Ghuman, J.K., Shah, B., Chuang, S.Z., Swiezy, N.B., Gonzalez, N.M., Hollway, J., Koenig, K., McGough, J.J., Ritz, L., and Vitiello, B. 2005. 'Risperidone for the core symptom domains of autism: results from the study by the autism network of the research units on pediatric psychopharmacology', *Am J Psychiatry*, 162: 1142-8.

McPartland, J.C., Bernier, R.A., Jeste, S.S., Dawson, G., Nelson, C.A., Chawarska, K., Earl, R., Faja, S., Johnson, S.P., Sikich, L., Brandt, C.A., Dziura, J.D., Rozenblit, L., Hellemann, G., Levin, A.R., Murias, M., Naples, A.J., Platt, M.L., Sabatos-DeVito, M., Shic, F., Senturk, D., Sugar, C.A., and Webb, S.J. 2020. 'The Autism Biomarkers Consortium for Clinical Trials (ABC-CT): Scientific Context, Study Design, and Progress Toward Biomarker Qualification', *Front Integr Neurosci*, 14: 16.

Mehrazad-Saber, Z., Kheirouri, S., and Noorazar, S.-G. 2018. 'Effects of l-Carnosine Supplementation on Sleep Disorders and Disease Severity in Autistic Children:

A Randomized, Controlled Clinical Trial', *Basic & Clinical Pharmacology & Toxicology*, 123: 72-77.
Minshawi, N.F., Wink, L.K., Shaffer, R., Plawecki, M.H., Posey, D.J., Liu, H., Hurwitz, S., McDougle, C.J., Swiezy, N.B., and Erickson, C.A. 2016. 'A randomized, placebo-controlled trial of D-cycloserine for the enhancement of social skills training in autism spectrum disorders', *Molecular Autism*, 7: 1-10.
Miral, S., Gencer, O., Inal-Emiroglu, F.N., Baykara, B., Baykara, A., and Dirik, E. 2008. 'Risperidone versus haloperidol in children and adolescents with AD : a randomized, controlled, double-blind trial', *Eur Child Adolesc Psychiatry*, 17: 1-8.
Möricke, E., Buitelaar, J.K., and Rommelse, N.N.J. 2016. 'Do We Need Multiple Informants When Assessing Autistic Traits? The Degree of Report Bias on Offspring, Self, and Spouse Ratings', *J Autism Dev Disord*, 46: 164-75.
Mundt, J.C., Greist, J.H., Jefferson, J.W., Katzelnick, D.J., DeBrota, D.J., Chappell, P.B., and Modell, J.G. 2007. 'Is it easier to find what you are looking for if you think you know what it looks like?', *Journal of clinical psychopharmacology*, 27: 121-25.
Murphy, D., and Spooren, W. 2012. 'EU-AIMS: a boost to autism research', *Nat Rev Drug Discov*, 11: 815-6.
NCT01171937. 2010. 'Risperidone Treatment In Children With Autism Spectrum Disorder And High Levels Of Repetitive Behavior (ProjectV)'.
NCT01260961. 2010. 'Developing Treatment, Treatment Validation and Treatment Scope in the Setting of an Autism Clinical Trial'.
NCT01661855. 2017. 'A pilot study of riluzole versus placebo in the treatment of children and adolescents with ASD (RILISE)'.
NCT02839915. 2016. 'Leucovorin for the Treatment of Language Impairment in Children With Autism Spectrum Disorder'.
NCT03202303. 2017. 'Cannabidivarin (CBDV) vs. Placebo in Children With Autism Spectrum Disorder (ASD)'.
NCT03553875. 2018. 'Memantine for the Treatment of Social Deficits in Youth With Disorders of Impaired Social Interactions'.
NCT04174365. 2019. 'Brexpiprazole in Treatment of Children and Adolescents With Irritability Associated With Autism Spectrum Disorder (Anchor)'.
NCT04745026. 2021. 'Trial to Investigate the Safety and Efficacy of Cannabidiol Oral Solution (GWP42003-P; CBD-OS) in Children and Adolescents With Autism Spectrum Disorder'.
NCT05067582. 2021. 'A 12-Week Crossover Study to Assess the Efficacy, Safety and Tolerability of L1-79 in Subjects Aged 12-21 Years With Autism Spectrum Disorder'.
NCT05439616. 2022. 'Study of Cariprazine Oral Capsules or Solution to Assess Adverse Events and Change in Irritability Due to Autism Spectrum Disorder (ASD) in Participants Aged 5-17 Years With ASD'.
Ness, S.L., Bangerter, A., Manyakov, N.V., Lewin, D., Boice, M., Skalkin, A., Jagannatha, S., Chatterjee, M., Dawson, G., and Goodwin, M.S. 2019. 'An observational study with the Janssen Autism Knowledge Engine (JAKE®) in individuals with autism spectrum disorder', *Frontiers in neuroscience*, 13: 111.
NICE. 2021a. "Autism spectrum disorder in adults: diagnosis and management [CG142]." In.
———. 2021b. "Autism spectrum disorder in under 19s: support and management (CG170)." In.

Nikolakopoulou, A., Higgins, J.P.T., Papakonstantinou, T., Chaimani, A., Del Giovane, C., Egger, M., and Salanti, G. 2020. 'CINeMA: An approach for assessing confidence in the results of a network meta-analysis', *PLoS Med*, 17: e1003082-e82.

Norman, G.R., Sloan, J.A., and Wyrwich, K.W. 2003. 'Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation', *Med Care*, 41: 582-92.

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

Papakonstantinou, T., Nikolakopoulou, A., Higgins, J.P.T., Egger, M., and Salanti, G. 2020. 'CINeMA: Software for semiautomated assessment of the confidence in the results of network meta-analysis', *Campbell Systematic Reviews*, 16: e1080.

Parellada, M., Moreno, C., Moreno, M., Espliego, A., de Portugal, E., and Arango, C. 2012. 'Placebo effect in child and adolescent psychiatric trials', *Eur Neuropsychopharmacol*, 22: 787-99.

Parellada, M., San José Cáceres, A., Palmer, M., Delorme, R., Jones, E.J.H., Parr, J.R., Anagnostou, E., Murphy, D.G.M., Loth, E., and Wang, P.P. 2021. 'A phase II randomized, double-blind, placebo-controlled study of the efficacy, safety, and tolerability of arbaclofen administered for the treatment of social function in children and adolescents with autism spectrum disorders: study protocol for AIMS-2-TRIALS-CT1', *Frontiers in psychiatry*: 1153.

Parker, K.J., Oztan, O., Libove, R.A., Mohsin, N., Karhson, D.S., Sumiyoshi, R.D., Summers, J.E., Hinman, K.E., Motonaga, K.S., Phillips, J.M., Carson, D.S., Fung, L.K., Garner, J.P., and Hardan, A.Y. 2019. 'A randomized placebo-controlled pilot trial shows that intranasal vasopressin improves social deficits in children with autism', *Sci Transl Med*, 11.

Persico, A.M., Ricciardello, A., Lamberti, M., Turriziani, L., Cucinotta, F., Brogna, C., Vitiello, B., and Arango, C. 2021. 'The pediatric psychopharmacology of autism spectrum disorder: A systematic review - Part I: The past and the present', *Prog Neuropsychopharmacol Biol Psychiatry*, 110: 110326.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R., and Rushton, L. 2008. 'Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry', *Journal of clinical epidemiology*, 61: 991-96.

Posey, D.J., and McDougle, C.J. 2000. 'The pharmacotherapy of target symptoms associated with autistic disorder and other pervasive developmental disorders', *Harvard Review of Psychiatry*, 8: 45-63.

Provenzani, U., Fusar-Poli, L., Brondino, N., Damiani, S., Vercesi, M., Meyer, N., Rocchetti, M., and Politi, P. 2020. 'What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials', *Autism*, 24: 274-84.

R Core Team. 2013. 'R: A language and environment for statistical computing'.

Reddihough, D.S., Marraffa, C., Mouti, A., O'Sullivan, M., Lee, K.J., Orsini, F., Hazell, P., Granich, J., Whitehouse, A.J.O., Wray, J., Dossetor, D., Santosh, P., Silove, N., and Kohn, M. 2019. 'Effect of Fluoxetine on Obsessive-Compulsive Behaviors in Children and Adolescents With Autism Spectrum Disorders: A Randomized Clinical Trial', *Jama*, 322: 1561-69.

Reddy, V.P., Garrett, M.R., Perry, G., and Smith, M.A. 2005. 'Carnosine: a versatile antioxidant and antiglycating agent', *Science of Aging Knowledge Environment*, 2005: pe12-pe12.

Reichow, B., Volkmar, F.R., and Cicchetti, D.V. 2008. 'Development of the evaluative method for evaluating and determining evidence-based practices in autism', *J Autism Dev Disord*, 38: 1311-9.
Rezaei, M., Moradi, A., Tehrani-Doost, M., Hassanabadi, H., and Khosroabadi, R. 2018. 'Effects of Combining Medication and Pivotal Response Treatment on Aberrant Behavior in Children with Autism Spectrum Disorder', *Children (Basel)*, 5.
Rhodes, K.M., Turner, R.M., and Higgins, J.P.T. 2015. 'Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data', *Journal of clinical epidemiology*, 68: 52-60.
Rødgaard, E.M., Jensen, K., Miskowiak, K.W., and Mottron, L. 2022. 'Representativeness of autistic samples in studies recruiting through social media', *Autism Res*.
Rodrigues, F.B., and Ferreira, J.J. 2020. 'Strategies to minimize placebo effects in research investigations', *International review of neurobiology*, 153: 49-70.
Rodrigues, R., Lai, M.C., Beswick, A., Gorman, D.A., Anagnostou, E., Szatmari, P., Anderson, K.K., and Ameis, S.H. 2021. 'Practitioner Review: Pharmacological treatment of attention-deficit/hyperactivity disorder symptoms in children and youth with autism spectrum disorder: a systematic review and meta-analysis', *J Child Psychol Psychiatry*, 62: 680-700.
Rogge, N., and Janssen, J. 2019. 'The Economic Costs of Autism Spectrum Disorder: A Literature Review', *Journal of Autism and Developmental Disorders*, 49: 2873-900.
Rothman, J. 2020. 'L1-79 and the role of catecholamines in autism', *Autism spectrum disorder-profile, heterogeneity, neurobiology and intervention*.
Roy, A., Roy, M., Deb, S., Unwin, G., and Roy, A. 2015. 'Are opioid antagonists effective in attenuating the core symptoms of autism spectrum conditions in children: a systematic review', *Journal of Intellectual Disability Research*, 59: 293-306.
Rücker, G. 2012. 'Network meta-analysis, electrical networks and graph theory', *Res Synth Methods*, 3: 312-24.
Rücker, G., and Schwarzer, G. 2015. 'Ranking treatments in frequentist network meta-analysis works without resampling methods', *BMC Med Res Methodol*, 15: 58.
Rücker, G., Schwarzer, G., Krahn, U., König, J., and Schwarzer, M.G. 2016. "Package 'netmeta': Network Meta-Analysis using Frequentist Methods. 2019." In.
RUPP. 2002. 'Risperidone in children with autism and serious behavioral problems', *N Engl J Med*, 347: 314-21.
Rutherford, B.R., and Roose, S.P. 2013. 'A model of placebo response in antidepressant clinical trials', *Am J Psychiatry*, 170: 723-33.
Rutter, M., Le Couteur, A., and Lord, C. 2003. 'ADI-R', *Autism diagnostic interview revised Manual Los Angeles: Western Psychological Services*.
Salanti, G., Del Giovane, C., Chaimani, A., Caldwell, D.M., and Higgins, J.P. 2014. 'Evaluating the quality of evidence from a network meta-analysis', *PLoS One*, 9: e99682.
Salanti, G., Higgins, J.P.T., Ades, A.E., and Ioannidis, J.P.A. 2008. 'Evaluation of networks of randomized trials', *Statistical methods in medical research*, 17: 279-301.
Salazar de Pablo, G., Jordá, C.P., Vaquerizo-Serrano, J., Moreno, C., Cabras, A., Arango, C., Hernández, P., Veenstra-VanderWeele, J., Simonoff, E., Fusar-Poli, P., Santosh, P., Cortese, S., and Parellada, M. 2022. 'Systematic Review and

Meta-analysis: Efficacy of Pharmacological Interventions for Irritability and Emotional Dysregulation in Autism Spectrum Disorder and Predictors of Response', *J Am Acad Child Adolesc Psychiatry*.

Samara, M.T., Spineli, L.M., Furukawa, T.A., Engel, R.R., Davis, J.M., Salanti, G., and Leucht, S. 2013. 'Imputation of response rates from means and standard deviations in schizophrenia', *Schizophrenia research*, 151: 209-14.

Sandler, A.D., and Bodfish, J.W. 2000. 'Placebo effects in autism: lessons from secretin', *J Dev Behav Pediatr*, 21: 347-50.

Scahill, L., Aman, M.G., Lecavalier, L., Halladay, A.K., Bishop, S.L., Bodfish, J.W., Grondhuis, S., Jones, N., Horrigan, J.P., and Cook, E.H. 2015. 'Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder', *Autism*, 19: 38-52.

Scahill, L., Bearss, K., Sarhangian, R., McDougle, C.J., Arnold, L.E., Aman, M.G., McCracken, J.T., Tierney, E., Gillespie, S., Postorino, V., and Vitiello, B. 2017. 'Using a Patient-Centered Outcome Measure to Test Methylphenidate Versus Placebo in Children with Autism Spectrum Disorder', *J Child Adolesc Psychopharmacol*, 27: 125-31.

Scahill, L., McDougle, C.J., Aman, M.G., Johnson, C., Handen, B., Bearss, K., Dziura, J., Butter, E., Swiezy, N.G., Arnold, L.E., Stigler, K.A., Sukhodolsky, D.D., Lecavalier, L., Pozdol, S.L., Nikolov, R., Hollway, J.A., Korzekwa, P., Gavaletz, A., Kohn, A.E., Koenig, K., Grinnon, S., Mulick, J.A., Yu, S., and Vitiello, B. 2012. 'Effects of risperidone and parent training on adaptive functioning in children with pervasive developmental disorders and serious behavioral problems', *J Am Acad Child Adolesc Psychiatry*, 51: 136-46.

Scahill, L., McDougle, C.J., Williams, S.K., Dimitropoulos, A., Aman, M.G., McCracken, J.T., Tierney, E., Arnold, L.E., Cronin, P., Grados, M., Ghuman, J., Koenig, K., Lam, K.S., McGough, J., Posey, D.J., Ritz, L., Swiezy, N.B., and Vitiello, B. 2006. 'Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders', *J Am Acad Child Adolesc Psychiatry*, 45: 1114-23.

Schedlowski, M., Enck, P., Rief, W., and Bingel, U. 2015. 'Neuro-bio-behavioral mechanisms of placebo and nocebo responses: implications for clinical trials and clinical practice', *Pharmacological reviews*, 67: 697-730.

Schneider-Thoma, J., Chalkou, K., Dörries, C., Bighelli, I., Ceraso, A., Huhn, M., Siafis, S., Davis, J.M., Cipriani, A., Furukawa, T.A., Salanti, G., and Leucht, S. 2022. 'Comparative efficacy and tolerability of 32 oral and long-acting injectable antipsychotics for the maintenance treatment of adults with schizophrenia: a systematic review and network meta-analysis', *The Lancet*, 399: 824-36.

Schopler, E., Reichler, R.J., DeVellis, R.F., and Daly, K. 1980. 'Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS)', *Journal of autism and developmental disorders*.

Schwarzer, G. 2007. 'meta: An R package for meta-analysis', *R news*, 7: 40-45.

Schwarzer, G., Chemaitelly, H., Abu-Raddad, L.J., and Rücker, G. 2019. 'Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions', *Research synthesis methods*, 10: 476-83.

Sharp, W.G., Berry, R.C., McCracken, C., Nuhu, N.N., Marvel, E., Saulnier, C.A., Klin, A., Jones, W., and Jaquess, D.L. 2013. 'Feeding problems and nutrient intake in children with autism spectrum disorders: a meta-analysis and comprehensive review of the literature', *J Autism Dev Disord*, 43: 2159-73.

Siafis, S., Bursch, N., Müller, K., Schmid, L., Schuster, F., Waibel, J., Huynh, T., Matthes, F., Rodolico, A., Brieger, P., Bühner, M., Heres, S., Leucht, S., and Hamann, J. 2022a. 'Evidence-based Shared-Decision-Making Assistant (SDM-assistant) for choosing antipsychotics: protocol of a cluster-randomized trial in hospitalized patients with schizophrenia', *BMC Psychiatry*, 22: 406.

Siafis, S., Çıray, O., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2020. 'Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis', *Mol Autism*, 11: 66.

Siafis, S., Çıray, O., Wu, H., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., San José Cáceres, A., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2022b. 'Pharmacological and dietary-supplement treatments for autism spectrum disorder: a systematic review and network meta-analysis', *Mol Autism*, 13: 10.

Siafis, S., Davis, J.M., and Leucht, S. 2021. 'Antipsychotic drugs: from 'major tranquilizers' to Neuroscience-based-Nomenclature', *Psychol Med*, 51: 522-24.

Siafis, S., Rodolico, A., Çıray, O., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2021. 'Imputing the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder', *Brain Sci*, 11.

Smith, R., Ou, J., Jin, H., Wu, R., Fahey, J., Arriaza, J., Cornblatt, B., Zhao, J., and Davis, J.M. 2020. "Sulforaphane as a Treatment for Autism: A Randomized Double-Blind Study." In, 79-80. SPRINGERNATURE CAMPUS, 4 CRINAN ST, LONDON, N1 9XW, ENGLAND.

Solmi, M., Fornaro, M., Ostinelli, E.G., Zangani, C., Croatto, G., Monaco, F., Krinitski, D., Fusar-Poli, P., and Correll, C.U. 2020. 'Safety of 80 antidepressants, antipsychotics, anti-attention-deficit/hyperactivity medications and mood stabilizers in children and adolescents with psychiatric disorders: a large scale systematic meta-review of 78 adverse effects', *World Psychiatry*, 19: 214-32.

Soorya, L.V., Fogg, L., Ocampo, E., Printen, M., Youngkin, S., Halpern, D., Kolevzon, A., Lee, S., Grodberg, D., and Anagnostou, E. 2021. 'Neurocognitive outcomes from memantine: a pilot, double-blind, placebo-controlled trial in children with autism spectrum disorder', *Journal of Child and Adolescent Psychopharmacology*, 31: 475-84.

Sparrow, S.S. 2011. 'Vineland Adaptive Behavior Scales.' in Jeffrey S. Kreutzer, John DeLuca and Bruce Caplan (eds.), *Encyclopedia of Clinical Neuropsychology* (Springer New York: New York, NY).

Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., Reiczigel, J., Robison-Cox, J., Sebastiani, P., and Solymos, P. 2020. "epiR: Tools for the Analysis of Epidemiological Data. R package version 1.0–14." In.: Published.

Sturm, A., Kuhfeld, M., Kasari, C., and McCracken, J.T. 2017. 'Development and validation of an item response theory-based Social Responsiveness Scale short form', *J Child Psychol Psychiatry*, 58: 1053-61.

Sturman, N., Deckx, L., and van Driel, M.L. 2017. 'Methylphenidate for children and adolescents with autism spectrum disorder', *Cochrane Database of Systematic Reviews*.

Suhail A, D., Furuya-Kanamori, L., Xu, C., Lin, L., Chivese, T., and Thalib, L. 2020. 'Questionable utility of the relative risk in clinical research: A call for change to practice'.

Tromans, S., and Adams, C. 2018. 'Brief Report: Autism Spectrum Disorder: A Comprehensive Survey of Randomized Controlled Trials', *J Autism Dev Disord*, 48: 3228-32.

Troost, P.W., Lahuis, B.E., Steenhuis, M.P., Ketelaars, C.E., Buitelaar, J.K., van Engeland, H., Scahill, L., Minderaa, R.B., and Hoekstra, P.J. 2005. 'Long-term effects of risperidone in children with autism spectrum disorders: a placebo discontinuation study', *J Am Acad Child Adolesc Psychiatry*, 44: 1137-44.

Trudeau, M.S., Madden, R.F., Parnell, J.A., Gibbard, W.B., and Shearer, J. 2019. 'Dietary and Supplement-Based Complementary and Alternative Medicine Use in Pediatric Autism Spectrum Disorder', *Nutrients*, 11: 1783.

Turner, R.M., Davey, J., Clarke, M.J., Thompson, S.G., and Higgins, J.P. 2012. 'Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews', *Int J Epidemiol*, 41: 818-27.

Usui, T., Macleod, M.R., McCann, S.K., Senior, A.M., and Nakagawa, S. 2021. 'Meta-analysis of variation suggests that embracing variability improves both replicability and generalizability in preclinical research', *PLoS Biol*, 19: e3001009.

van Andel, D.M., Sprengers, J.J., Keijzer-Veen, M.G., Schulp, A.J.A., Lillien, M.R., Scheepers, F.E., and Bruining, H. 2022. 'Bumetanide for Irritability in Children With Sensory Processing Problems Across Neurodevelopmental Disorders: A Pilot Randomized Controlled Trial', *Front Psychiatry*, 13: 780281.

Varni, J.W., Burwinkle, T.M., Seid, M., and Skarr, D. 2003. 'The PedsQL™* 4.0 as a pediatric population health measure: feasibility, reliability, and validity', *Ambulatory pediatrics*, 3: 329-41.

Veenstra-VanderWeele, J., Cook, E.H., King, B.H., Zarevics, P., Cherubini, M., Walton-Bowen, K., Bear, M.F., Wang, P.P., and Carpenter, R.L. 2017. 'Arbaclofen in Children and Adolescents with Autism Spectrum Disorder: A Randomized, Controlled, Phase 2 Trial', *Neuropsychopharmacology*, 42: 1390-98.

Veroniki, A.A., Tsokani, S., White, I.R., Schwarzer, G., Rücker, G., Mavridis, D., Higgins, J.P.T., and Salanti, G. 2021. 'Prevalence of evidence of inconsistency and its association with network structural characteristics in 201 published networks of interventions', *BMC Med Res Methodol*, 21: 224.

Veroniki, A.A., Vasiliadis, H.S., Higgins, J.P., and Salanti, G. 2013. 'Evaluation of inconsistency in networks of interventions', *Int J Epidemiol*, 42: 332-45.

Viechtbauer, W. 2010. 'Conducting meta-analyses in R with the metafor package', *Journal of statistical software*, 36: 1-48.

Wagner, A., Lecavalier, L., Arnold, L.E., Aman, M.G., Scahill, L., Stigler, K.A., Johnson, C.R., McDougle, C.J., and Vitiello, B. 2007. 'Developmental disabilities modification of the Children's Global Assessment Scale', *Biological psychiatry*, 61: 504-11.

Wan, X., Wang, W., Liu, J., and Tong, T. 2014. 'Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range', *BMC Med Res Methodol*, 14: 135.

Wang, M., Jiang, L., and Tang, X. 2017. 'Levetiracetam is associated with decrease in subclinical epileptiform discharges and improved cognitive functions in

pediatric patients with autism spectrum disorder', *Neuropsychiatr Dis Treat*, 13: 2321-26.
Wasserman, S., Iyengar, R., Chaplin, W.F., Watner, D., Waldoks, S.E., Anagnostou, E., Soorya, L., and Hollander, E. 2006. 'Levetiracetam versus placebo in childhood and adolescent autism: a double-blind placebo-controlled study', *Int Clin Psychopharmacol*, 21: 363-7.
Weimer, K., Colloca, L., and Enck, P. 2015. 'Placebo eff ects in psychiatry: mediators and moderators', *The lancet Psychiatry*, 2: 246-57.
Weimer, K., Gulewitsch, M.D., Schlarb, A.A., Schwille-Kiuntke, J., Klosterhalfen, S., and Enck, P. 2013. 'Placebo effects in children: a review', *Pediatric research*, 74: 96-102.
Whalley, B., and Hyland, M.E. 2013. 'Placebo by proxy: the effect of parents' beliefs on therapy for children's temper tantrums', *J Behav Med*, 36: 341-6.
WHO. 2018. 'ICD-11 for mortality and morbidity statistics (2018)'.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., and Hester, J. 2019. "Welcome to the Tidyverse. J Open Source Software 4: 1686." In.
Williams, K., Wray, J.A., and Wheeler, D.M. 2012. 'Intravenous secretin for autism spectrum disorders (ASD)', *Cochrane Database of Systematic Reviews*.
Wilson, M., Hamilton, D., Whelan, T., and Pilkington, P. 2018. 'A systematic review of factors related to parents' treatment decisions for their children with autism spectrum disorders', *Research in Autism Spectrum Disorders*, 48: 17-35.
Wink, L.K., Adams, R., Horn, P.S., Tessier, C.R., Bantel, A.P., Hong, M., Shaffer, R.C., Pedapati, E.V., and Erickson, C.A. 2018. 'A Randomized Placebo-Controlled Cross-Over Pilot Study of Riluzole for Drug-Refractory Irritability in Autism Spectrum Disorder', *Journal of Autism and Developmental Disorders*, 48: 3051-60.
Wink, L.K., Minshawi, N.F., Shaffer, R.C., Plawecki, M.H., Posey, D.J., Horn, P.S., Adams, R., Pedapati, E.V., Schaefer, T.L., McDougle, C.J., Swiezy, N.B., and Erickson, C.A. 2017. 'd-Cycloserine enhances durability of social skills training in autism spectrum disorder', *Mol Autism*, 8: 2.
Winkelbeiner, S., Leucht, S., Kane, J.M., and Homan, P. 2019. 'Evaluation of Differences in Individual Treatment Response in Schizophrenia Spectrum Disorders: A Meta-analysis', *JAMA Psychiatry*, 76: 1063-73.
Zhang, L., Huang, C.-C., Dai, Y., Luo, Q., Ji, Y., Wang, K., Deng, S., Yu, J., Xu, M., Du, X., Tang, Y., Shen, C., Feng, J., Sahakian, B.J., Lin, C.-P., and Li, F. 2020. 'Symptom improvement in children with autism spectrum disorder following bumetanide administration is associated with decreased GABA/glutamate ratios', *Translational Psychiatry*, 10: 9.
Zhang, L., Xu, Y., Li, H., Li, B., Duan, G., and Zhu, C. 2022. 'The role of probiotics in children with autism spectrum disorders: A study protocol for a randomised controlled trial', *PLOS ONE*, 17: e0263109.
Zhou, M.S., Nasir, M., Farhat, L.C., Kook, M., Artukoglu, B.B., and Bloch, M.H. 2021. 'Meta-analysis: Pharmacologic Treatment of Restricted and Repetitive Behaviors in Autism Spectrum Disorders', *J Am Acad Child Adolesc Psychiatry*, 60: 35-45.
Zohar, J., Stahl, S., Moller, H.-J., Blier, P., Kupfer, D., Yamawaki, S., Uchida, H., Spedding, M., Goodwin, G.M., and Nutt, D. 2015. 'A review of the current nomenclature for psychotropic agents and an introduction to the Neuroscience-based Nomenclature', *European Neuropsychopharmacology*, 25: 2318-25.

Epilogue and acknowledgments

As I set out on conducting my thesis in 2018, about four years ago, I underestimated the difficulties and challenges that I would face. First, the topic of the thesis was on pharmacotherapy and drug development in autism, an emerging, heterogeneous and rapidly developing field (e.g., >25% of the included clinical trials were published after I started the thesis). Such a primitive, yet expanding, field required often skillful maneuvering and adaption. Second, the world was also rapidly changing during this period, with most notable examples the COVID-19 pandemic and the war in Europe. Such novel global challenges provoked in many occasions feelings of anxiety and insecurity.

But it is because of these challenges that this thesis was a unique journey of academic and personal development, and a beacon that could guide me later in my life.

On this difficult journey, I was fortunate to have caring, supportive, intelligent and trustworthy people with me. I am truly grateful to them, and I would like to thank them by name:

My dear supervisor and scientific father, **Prof. Stefan Leucht**, for giving me the opportunity to work in his team and develop myself further, as well as for his patience, flexibility, valuable advices, and the timely and well-directed guidance in every step. His multidisciplinary knowledge and world-class achievements inspire me to pursue a career in evidence-based medicine and succeed, and his *modus vivedi* and hard work show me the way.

My dear mentor in the thesis and mentor in life, **Prof. Georgios Papazisis**, who first introduced me to research when I was a medical student and ignited my interest in psychiatry and pharmacology, as well as for always encouraging me to pursue my own ideas and spread my wings in order to achieve higher.

My dear friend and role model, **Prof. Myrto Samara**, for being the point of connection between Greece and Germany, for the interesting scientific discussions, and especially for sharing her experiences and for her sharp and direct criticism, which I really trust and appreciate.

My dearest and trusted comrade, **Dr. Hui Wu**, for sharing her unique and insightful perspective into scientific, non-scientific and personal matters, and for her endless support, which could not be described and thanked in a few lines of words.

My dear colleagues and friends, **Dr. Irene Bighelli**, **Alessandro Rodolico**, **Angelika Kapfhammer**, and **Dr. Johannes Schneider**, for sharing their unique skills, for the fruitful collaboration and endless support in scientific, technical and personal matters.

Friends and previous or current members of the lab, **Anna Ceraso**, **Shimeng Dong**, **Lena Feber**, **PD Dr. Maximillian Huhn**, **Giulia Paris**, **Natalie Peter**, **Dr. Yikang Zhu**, and **Dongfang Wang**, with whom, we shared along our journeys, ideas and frustrations, happy and sad times.

Co-authors in the publications included in the thesis, and especially **Prof. Celso Arango**, **Prof. Tony Charman**, **Prof. Dimitris Mavridis**, **Prof. Declan G. Murphy**, **Prof. Mara Parellada**, for their valuable scientific contribution and wonderful collaboration.

My dearest friends **Anthi-Sotiria Giagka**, **Marios-Aristeidis Kirkos**, **Efstathia Vouzaxaki** and **Pantazis Vouzaxakis**, for their friendship and unforgettable moments we shared together, and for their support and long conversations in times of need.

My dearest friend, **Apostolos Zacharis**, who has been like family for me with his unconditional support, his sharp but valuable criticism and the long, multi-layered and recurrent conversations.

Finally, I would like to say a special thank you to my family.

My sister, **Theodora Siafi**, who, despite younger than me, has often be a role model with her determination and strong character, as well as for her objective advices in times of doubts.

My parents, **Konstantinos Siafis** and **Konstantina Anagnostou**, for their unconditional love and support, for believing in me, for encouraging me to pursue my dreams and express myself, for teaching me how to work hard and systematically, as well as for their optimism and appetite for life. I owe every success in my life to them because they provided me with a solid foundation and a guide to build my life.

Munich, 22 July 2022

Spyridon Siafis

Funding statement

The thesis was conducted in part of the AIMS-2-TRIALS. The project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. Any views expressed are those of the author and not necessarily those of the funders.

Publications

The published versions of the articles included in the thesis are listed below (Siafis et al. 2020; Siafis et al. 2022b; Siafis et al. 2021). Their appendices can be found in the respective publications online and were not listed here because of their extensive length.

In particular, two of the articles (Siafis et al. 2020; Siafis et al. 2022b) were published in Molecular Autism, which is currently the leading journal for basic, clinical and translational research on autism spectrum and related neurodevelopmental disorder. Their results were further disseminated to the public with lay summaries written in collaboration with the AIMS-2-TRIALS communication team (AIMS-2-TRIALS 2020, 2022).

The articles were published Open Access and they are licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaption, distribution and reproduction in any medium or format, as long as, appropriate credit is given to the original authors and the source, a link to the Creative Commons license is given and any changes were indicated.

The citations of the articles can be found in the sections "Publications included in the thesis" and "5. References".

The link to the Creative Commons Attribution 4.0 International License is: https://creativecommons.org/licenses/by/4.0/ .

A summary of each publication and my contributions can be found the section "3. Results and publication summaries".

Publication I

Network meta-analysis on the effects of pharmacological and dietary-supplement treatments for autism spectrum disorder

Citation: Siafis, S., Çıray, O., Wu, H., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., San José Cáceres, A., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2022. 'Pharmacological and dietary-supplement treatments for autism spectrum disorder: a systematic review and network meta-analysis', *Mol Autism*, 13: 10. doi: 10.1186/s13229-022-00488-4.

This article is licensed under a Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/

Siafis *et al. Molecular Autism* (2022) 13:10
https://doi.org/10.1186/s13229-022-00488-4

Molecular Autism

RESEARCH Open Access

Pharmacological and dietary-supplement treatments for autism spectrum disorder: a systematic review and network meta-analysis

Spyridon Siafis[1*], Oğulcan Çıray[2], Hui Wu[1], Johannes Schneider-Thoma[1], Irene Bighelli[1], Marc Krause[3], Alessandro Rodolico[4], Anna Ceraso[5], Giacomo Deste[5], Maximilian Huhn[1,6], David Fraguas[7], Antonia San José Cáceres[14,15], Dimitris Mavridis[9,10], Tony Charman[11], Declan G. Murphy[12], Mara Parellada[8,13,14,15], Celso Arango[8,13,14,15] and Stefan Leucht[1]

Abstract

Background: There is still no approved medication for the core symptoms of autism spectrum disorder (ASD). This network meta-analysis investigated pharmacological and dietary-supplement treatments for ASD.

Methods: We searched for randomized-controlled-trials (RCTs) with a minimum duration of seven days in ClinicalTrials.gov, EMBASE, MEDLINE, PsycINFO, WHO-ICTRP (from inception up to July 8, 2018), CENTRAL and PubMed (up to November 3, 2021). The co-primary outcomes were core symptoms (social-communication difficulties-SCD, repetitive behaviors-RB, overall core symptoms-OCS) measured by validated scales and standardized-mean-differences (SMDs). Associated symptoms, e.g., irritability/aggression and attention-deficit/hyperactivity disorder (ADHD) symptoms, dropouts and important side-effects, were investigated as secondary outcomes. Studies in children/adolescents and adults were analyzed separately in random-effects pairwise and network meta-analyses.

Results: We analyzed data for 41 drugs and 17 dietary-supplements, from 125 RCTs ($n = 7450$ participants) in children/adolescents and 18 RCTs ($n = 1104$) in adults. The following medications could improve at least one core symptom domain in comparison with placebo: aripiprazole ($k = 6$ studies in analysis, SCD: SMD = 0.27 95% CI [0.09, 0.44], RB: 0.48 [0.26, 0.70]), atomoxetine ($k = 3$, RB:0.49 [0.18, 0.80]), bumetanide ($k = 4$, RB: 0.35 [0.09, 0.62], OCS: 0.61 [0.31, 0.91]), and risperidone ($k = 4$, SCM: 0.31 [0.06, 0.55], RB: 0.60 [0.29, 0.90]; $k = 3$, OCS: 1.18 [0.75, 1.61]) in children/adolescents; fluoxetine ($k = 1$, RB: 1.20 [0.45, 1.96]), fluvoxamine ($k = 1$, RB: 1.04 [0.27, 1.81]), oxytocin ($k = 6$, RB:0.41 [0.16, 0.66]) and risperidone ($k = 1$, RB: 0.97 [0.21,1.74]) in adults. There were some indications of improvement by carnosine, haloperidol, folinic acid, guanfacine, omega-3-fatty-acids, probiotics, sulforaphane, tideglusib and valproate, yet imprecise and not robust. Confidence in these estimates was very low or low, except moderate for oxytocin. Medications differed substantially in improving associated symptoms, and in their side-effect profiles.

Limitations: Most of the studies were inadequately powered (sample sizes of 20–80 participants), with short duration (8–13 weeks), and about a third focused on associated symptoms. Networks were mainly star-shaped, and there were indications of reporting bias. There was no optimal rating scale measuring change in core symptoms.

*Correspondence: spyridon.siafis@tum.de
[1] Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Ismaningerstr. 22, 81675 Munich, Germany
Full list of author information is available at the end of the article

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Conclusions: Some medications could improve core symptoms, although this could be likely secondary to the improvement of associated symptoms. Evidence on their efficacy and safety is preliminary; therefore, routine prescription of medications for the core symptoms cannot be recommended.

Trial registration PROSPERO-ID CRD42019125317.

Keywords: Autism, Meta-analysis, Treatment, Response, Social communication, Restricted and repetitive behaviors, Irritability, ADHD, Anxiety, Caregiver stress

Background

Autism spectrum disorder (ASD) consists of heterogeneous conditions, which are characterized by social-communication difficulties, restricted interests/repetitive behaviors and sensory abnormalities [1]. Behavioral interventions are the mainstay treatment [2]. Medications with different mechanisms of action have been examined in randomized controlled trials (RCTs) [3–5], and some of them have been found efficacious for associated symptoms, such as aripiprazole, risperidone and haloperidol for irritability, methylphenidate, atomoxetine, clonidine and guanfacine for attention-deficit/hyperactivity disorder (ADHD) symptoms and melatonin for sleep disorders [2, 6]. However, prior late-stage clinical trials failed to identify efficacious treatments for the core symptoms of neurodevelopmental disorders [5, 7]. Despite lack of clear evidence in efficacy, about half of the individuals with ASD receive psychotropic drugs [8]. The current synthesis of literature is restricted to medication classes or target symptoms [9–19], hence failing to combine the huge amount of recently conducted RCTs [3–5]. In order to better inform clinical practice and identify medications potentially efficacious for ASD, we combined evidence from pharmacological and dietary-supplement ASD trials in a network meta-analysis.

Methods

Search strategy and selection criteria

This network meta-analysis analyzed placebo-controlled and head-to-head RCTs on pharmacological/dietary-supplement interventions for ASD according to the PRISMA-NMA (Additional file 1: eAppendix-1) [20], and with PROSPERO-ID: CRD42019125317 (Additional file 1: eAppendix-2).

We searched ClinicalTrials.gov, EMBASE, MEDLINE, PsycINFO, WHO-ICTRP (from inception to July 8, 2018), CENTRAL and PubMed (last update on November 3, 2021), without restrictions in terms of language, document type, date/time and publication status (Additional file 1: eAppendix-3). Reference lists of included studies and reviews [2, 9–17, 19, 21–24] were inspected.

Participants should have a diagnosis of ASD according to standardized diagnostic criteria (e.g., DSM-III or newer versions) and/or validated diagnostic tools (e.g., ADI-R) [2], without restrictions in terms of age, sex, baseline severity and presence of genetic syndromes or other associated conditions (e.g., irritability, ADHD symptoms).

Any drug, dietary-supplement or placebo was eligible. We excluded augmentation and multimodal interventions (e.g., medications combined with risperidone or behavioral interventions) as well as other types of interventions (e.g., behavioral, elimination diets). The minimum duration of treatment was seven days, and there was no restriction in terms of dosing-schedule and route of administration. Multiple doses of the same intervention were combined [25] (Additional file 1: eAppendix-2.2).

Blinded and open RCTs were eligible. RCTs with a low or unclear risk of bias in random sequence generation and allocation concealment were eligible, yet we excluded trials with a high risk of bias in these domains [26]. Trials stated to be randomized but did not report the exact randomization methods (unclear risk of bias) were eligible, since poor reporting does not necessarily reflect the actual conducted methods [27–30] (Additional file 1: eAppendix-2.2). However, such trials were excluded in a *post hoc* sensitivity analysis. We included data only from the first phase of crossover studies in order to avoid carry-over effects [31], and we excluded discontinuation studies, studies published before 1980, or with a randomized sample smaller than ten participants [32].

At least two independent reviewers/contributors selected relevant records (SS, OC, HW, IB, MK, YZ, AC, GD and TF), extracted data from eligible studies into an Access database as well as evaluated risk of bias using the Cochrane risk-of-bias tool (SS, OC, AR, HW) [26]. Studies were rated as having a low, moderate or high overall risk of bias [33]. Differences were resolved with discussion, and if needed, a third reviewer was involved (SL, JST). Study authors were contacted for additional data by e-mail (with a reminder in case of no response) (Additional file 1: eAppendix-4).

Outcomes

The co-primary outcomes were the change in core symptoms measured with validated rating scales: (1) social-communication difficulties (SCD, e.g., ABC-L/SW [34]

or VABS-Socialization [35]), (2) repetitive behaviors (RB, e.g., ABC-S [34] or CYBOCS-PDD [36]), and (3) overall core symptoms (OCS, e.g., SRS [37] or CARS [38]). There is yet no optimal outcome measure [39], and we accepted a wide range of validated scales, giving preference to clinician-ratings and to the commonly used scales mentioned above, similar to our previous review [4] (Additional file 1: eAppendix-5.3).

Secondary outcomes were premature discontinuation (dropout) due to any reason and due to adverse events, number of participants with a positive response (preferably defined with a CGI-Improvement score ≤ 2 or at least "much improved" [40]), change in irritability/aggression, ADHD symptoms and anxiety/depression, quality of life, global functioning and caregiver stress (Additional file 1: eAppendix-5.3). We also examined the number of participants with adverse events, sedation, weight gain (preferably defined as $\geq 7\%$ increase) and extrapyramidal symptoms.

Data analysis

Random-effects pairwise and network meta-analyses were conducted within a frequentist framework using meta v4.15-1 [41] and netmeta v1.2-1 [42] in R statistical software v4.0.3 [43]. The certainty of evidence of comparisons with placebo for the co-primary outcomes was evaluated using CINeMA (Confidence in Network Meta-Analysis) [44, 45] (Additional file 1: eAppendix-6.9).

The effect-sizes for continuous outcomes were standardized mean differences (SMD, Hedge's g) and for dichotomous outcomes were odds ratios (OR), presented with their 95% confidence intervals (95% CI). We post hoc used ORs instead of relative risks, due to their preferred mathematical properties in meta-analysis [46, 47]. In order to present both continuous and dichotomous outcomes in figures, ORs were also converted to SMDs [25]. Treatments were ranked with *P*-scores [48]. Intention-to-treat data were used, whenever available, and methods that handle missing data were preferred to completers' data, giving preference to mixed-models for repeated measures (MMRM) and multiple imputation over last-observation carried forward (LOCF). For dichotomous outcomes, we assumed that participants lost to follow-up did not have a response. The number of participants with a positive response (CGI-Improvement ≤ 2) [40] and weight gain ($\geq 7\%$ increase) was imputed from means and standard deviations (SD) using a validated method, when dichotomous data were not reported [49, 50]. Missing SDs were calculated from available statistics [25], pooling subscales (e.g., SRS subscales, assuming a correlation of 0.5) [51] or using the mean SD of included studies [25]. Change scores were preferred to follow-up scores, and the former were estimated post hoc when both baseline and follow-up scores were available using a correlation of 0.5 [25], since baseline imbalances could have inflated treatment effects (Additional file 1: eAppendix-6.1).

RCTs in children/adolescents and adults (or mixed populations) were analyzed separately, since extrapolation between age groups is discouraged [52]. Transitivity was further assessed by comparing the distribution of clinical and methodological variables (i.e., study duration, type of rater, associated conditions at baseline, baseline scores of CGI-Severity (ranging 1–7) [40], ABC-Irritability (ranging 0–45) [34] and mean age). Trials focused on subgroups, i.e., intellectual disability/high-functioning, genetic syndrome or another associated condition, were classified in CINeMA with moderate indirectness [44].

A common heterogeneity variance (τ^2) was assumed for all comparisons per network, and heterogeneity was quantified as low, moderate or high by comparing τ^2 with its empirical distributions [53, 54]. Incoherence was examined globally with a design-by-treatment interaction test and locally with separating indirect from direct evidence [55].

We aimed to include unpublished trials (e.g., contacting authors, using data reported in trial registries, abstracts and reviews), and eligible studies with no usable data were considered in the assessment of reporting bias [44]. Additionally, small-study effects were examined with comparison-adjusted [56] (assuming the direction of bias towards newer medications) and contour-enhanced funnel plots, when there were more than ten studies per comparison [25].

The robustness of the results was investigated in sensitivity analyses using (a) fixed-effects models, excluding studies with (b) implied randomization, (c) genetic syndrome or (d) associated symptoms as inclusion criteria, (e) using only diagnostic evaluation tools, (f) with non-clinician-ratings, (g) from less developed countries (post hoc) [57], (h) with imputed SDs, (i) overall high risk of bias, (j) unclear risk of bias in random sequence generation or allocation concealment (post hoc), (k) open or single-blind, (l) shorter than four weeks, (m) presenting only completers' data, (n) using a correlation of 0.25 and 0.75 to calculate the SD of change scores, and (o) using ABC-L/SW or ABC-S (post hoc). In a post hoc sensitivity analysis, relative risks were used for dichotomous outcomes. Baseline severity could not be assessed in a subgroup or sensitivity analysis, due to inconsistent reporting and diversity of scales (Additional file 1: eAppendix-2.2).

Alpha was set at two-sided 0.05, except for heterogeneity, incoherence and funnel plot tests at 0.1 due to their small statistical power.

Results

Description of included studies

Study selection is presented with a PRISMA flow diagram (Additional file 1: eAppendix-4.1), and the list of included/excluded full-texts in Additional file 1: eAppendix-4.2/4.3. From 203 eligible trials, 125 trials in children/adolescents ($n=7450$ participants) and 18 in adults ($n=1104$) were included in the quantitative analysis.

Study characteristics are presented in Additional file 1: eAppendix-5.1 and the distribution of potential effect-modifiers in Additional file 1: eAppendix-6.1. The majority of trials were double-blind ($k=138$ studies), placebo-controlled ($k=137$) with a parallel-design ($k=110$) and two-arms ($k=125$). They were recently published (median publication year of 2015, interquartile range [2008–2019]), had a short duration (12 [8–13] weeks), small sample sizes (40 [23–76]) and few sites (1 [1–3]), which were mainly academic ($k=102$ trials had only academic sites).

The median age of participants was 8.2 [6.3–9.5] years in children/adolescents and 24.6 [21.9–27.9] years in adults. The overall male-to-female ratio was 5.3 [3.9–8.2]. Standardized diagnostic criteria were used in most of the studies (95%), and seven studies used only diagnostic evaluation tools. Associated symptoms were required as an inclusion criterion in about a third of the studies, mainly irritability and ADHD symptoms (in 30 trials), and a genetic syndrome (neurofibromatosis-type-I) in one trial [58]. At baseline, the sample was moderately to markedly ill with a CGI-S score of 4.8 [4.4–5.1], and ABC-Irritability of 16.9 [13.3–22.3], and about half of the participants had intellectual disability (50% [0–73.5%]). Nevertheless, reporting of participant characteristics was poor in about two thirds of the studies.

Risk of bias assessment is presented in Additional file 1: eAppendix-5.2. About 25% of the studies had an overall low risk of bias, 55% had moderate and 17% high. About half adequately reported methods of random sequence and allocation concealment, and blinding was adequately addressed in about 65%. High risk of bias was assigned in about 26% studies for incomplete outcome data, 36% for selective reporting and about 12% for other biases, mainly due to baseline imbalance or early trial termination. Finally, about 30% of the studies were funded by industry or their investigators applied for a patent.

Forty-one drugs were investigated in 100 trials (antipsychotics and antidepressants in about a third) and 17 dietary-supplements in 43 trials (Additional file 1: eAppendix-5.1). Interventions were connected in mainly star-shaped networks with placebo as the main node (Additional file 2: Fig. S1). Therefore, we focused on comparisons with placebo (Fig. 1, Additional file 3: Fig. S2),

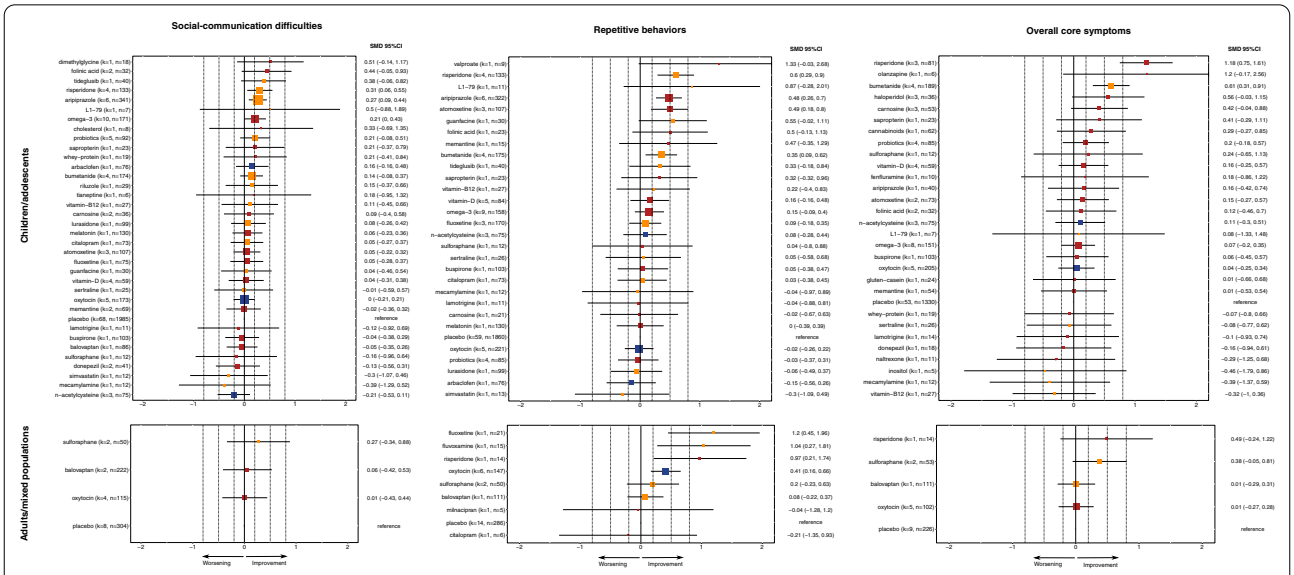


Fig. 1 Forest plots of network meta-analysis for the primary outcomes, i.e., social-communication difficulties (SCD), repetitive behaviors (RB), and overall core symptoms (OCS), in children/adolescents and adults. Placebo was used as reference. The squares and bars represent the effect-sizes (standardized mean differences-SMD) along with their 95% confidence intervals. The size of the square is proportional to the inverse standard error of the effect size. The color represents confidence in the estimates as evaluated with the CINeMA framework, i.e., blue = moderate, yellow = low, and red = very low. SMDs > 0 indicate more improvement with the medication in comparison with placebo, SMDs = 0 indicate no difference between medication and placebo, and SMDs < 0 indicate less improvement with the medication in comparison with placebo. SMDs could be interpreted as small (SMD = |0.2|), medium (SMD = |0.5|) and large (SMD = |0.8|), and these thresholds are presented with dashed lines. k = total number of studies for the intervention; n = total number of participants on the intervention

and league tables with all comparisons are presented in Additional file 4: Table S1. The results of pairwise meta-analyses and individual studies are presented in Additional file 5: Fig. S3. In addition, incoherence could not be evaluated when there were no closed loops (i.e., networks for anxiety/depression, quality of life, caregiver stress and all networks in adults). There was no clear indication of incoherence for the rest of the networks, except for irritability, response, weight gain and sedation in children/adolescents for which pairwise meta-analyses were conducted (Additional file 1: eAppendix-6.8).

Primary outcomes

Social-communication difficulties (SCD)

Social-communication difficulties were measured mainly with ABC-L/SW (55%) and VABS-S (18%).

In children/adolescents, social-communication difficulties were improved by risperidone ($k=4$ studies in the analysis, $n=133$ participants treated with risperidone; SMD = 0.31 95%CI [0.06, 0.55]; *low* quality of evidence) and aripiprazole ($k=6$, $n=341$; SMD = 0.27 [0.09, 0.44]; *low*). Some trends of improvement were noted for folinic acid ($k=2$, $n=32$, SMD = 0.44 [−0.05, 0.93]; *very low*), tideglusib ($k=1$, $n=40$; SMD = 0.38 [−0.06, 0.82]; *low*), omega-3-fatty-acids ($k=10$, $n=171$; SMD = 0.21 [0.00, 0.43], *very low*), probiotics ($k=5$, $n=92$; SMD = 0.21 [−0.08, 0.51]; *low*) and bumetanide ($k=4$, $n=174$; SMD = 0.14 [−0.08, 0.37]; *low*). There were no clear differences between other medications and placebo with very low-to-moderate confidence. Heterogeneity was low ($\tau^2=0$).

In adults, none of the investigated medications (sulforaphane, balovaptan, oxytocin) improved social-communication difficulties with very-low- or low-quality evidence. There were high levels of heterogeneity ($\tau^2=0.096$).

Repetitive behaviors (RB)

Repetitive behaviors were measured mainly with ABC-S (47%) and YBOCS-versions (27%).

In children/adolescents, repetitive behaviors were improved by risperidone ($k=4$, n = 133; SMD = 0.60 [0.29, 0.90]; *low*), aripiprazole ($k=6$, $n=322$; SMD = 0.48 [0.26, 0.70]; *very low*), atomoxetine ($k=3$, $n=107$; SMD = 0.49 [0.18, 0.80]; *very low*) and bumetanide ($k=4$, $n=175$; SMD = 0.35 [0.09, 0.62], *low*). There were trends for valproate ($k=1$, $n=9$; SMD = 1.33 [−0.03, 2.68]; *very low*) and guanfacine ($k=1$, $n=30$; SMD = 0.55 [−0.02, 1.11]; *low*), and no clear differences for other medications with very low-to-moderate confidence. Heterogeneity was low-to-moderate ($\tau^2=0.017$).

In adults, repetitive behaviors were improved by fluoxetine ($k=1$, $n=21$; SMD = 1.20 [0.45, 1.96]; *low*), fluvoxamine ($k=1$, $n=15$; SMD = 1.04 [0.27, 1.81]; *low*), risperidone ($k=1$, $n=14$; SMD = 0.97 [0.21, 1.74]; *very low*), and oxytocin ($k=6$, $n=147$; SMD = 0.41 [0.16, 0.66]; *moderate*). Sulforaphane, balovaptan, milnacipran and citalopram were not found efficacious with very low or low confidence. Heterogeneity was low ($\tau^2=0$).

Overall core symptoms (OCS)

Overall core symptoms were measured mainly with SRS (47%) and CARS (22%).

In children/adolescents, overall core symptoms were improved by risperidone ($k=3$, $n=81$; SMD = 1.18 [0.75, 1.61]; *very low*), and bumetanide ($k=4$, $n=189$; SMD = 0.61 [0.31, 0.91]; *low*). There were some trends for haloperidol ($k=3$, $n=36$; SMD = 0.56 [−0.03, 1.15]; *very low*) and carnosine ($k=3$, $n=53$; SMD = 0.42 [−0.04, 0.88]; *very low*), and no clear differences for other medications with very low-to-moderate confidence. There were moderate levels of heterogeneity ($\tau^2=0.038$) and no indication of incoherence. Nevertheless, a small study ($n=30$) [59] that found no difference between risperidone and memantine (SMD = 0.00 [−0.71, 0.72]) introduced incoherence and was excluded from the primary analysis of this outcome (Additional file 1: eAppendix-6.8), and the results were robust after inclusion of this study (Additional file 6: Fig. S4).

In adults, none of the investigated medications (risperidone, sulforaphane, balovaptan and oxytocin) found to be more efficacious than placebo in reducing overall core symptoms, though a trend was noted for sulforaphane ($k=2$, $n=53$; SMD = 0.38 [−0.05, 0.81]; *low*). Confidence in evidence was very low or low. Heterogeneity was low ($\tau^2=0$).

Sensitivity analysis

The results did not materially change in sensitivity analyses (Additional file 1: eAppendix-6.6, Additional file 6: Fig. S4). There were some potential differences in omega-3-fatty-acids. Omega-3-fatty-acids did not reduce social-communication difficulties in children/adolescents when studies on associated symptoms were excluded ($k=6$, $n=112$, SMD = 0.05 [−0.21, 0.32]) or when clinician-ratings were used ($k=3$, $n=53$, SMD = 0.03 [−0.36, 0.42]). Yet, their effect-size was larger when ABC-L/SW was used ($k=6$, $n=79$, SMD = 0.45 [0.13, 0.77]). In addition, the results for some interventions, i.e., folinic acid, carnosine, vitamin-D, were not robust in sensitivity analyses, which were based on one or two small trials with potentially inflated effect-sizes.

Small-study effects and publications

There was asymmetry in funnel plots for social-communication difficulties in children/adolescents, indicating

small-study effects (Additional file 1: eAppendix-6.8). Funnel plots for the other co-primary outcomes were inconclusive. Reporting bias was suspected for some medications, and quality of evidence was downgraded accordingly (Additional file 1: eAppendix-6.9).

Secondary outcomes

Irritability

Irritability was measured mainly with ABC-I (83%).

In children/adolescents, there was evidence of incoherence (none of the closed loops were incoherent, but $p\text{-}_{\text{design-by-treatment}} = 0.014$) and pairwise meta-analysis were conducted. Irritability was improved by risperidone ($k = 4$ studies in the analysis, $n = 138$ participants treated with risperidone; SMD = 1.05 [0.76, 1.33], $\tau^2 = 0.02$), sulforaphane ($k = 1$, $n = 12$; SMD = 0.97 [0.12, 1.83]), aripiprazole ($k = 5$, $n = 312$; SMD = 0.63 [0.44, 0.82], $\tau^2 = 0$), and citalopram ($k = 1$, $n = 73$; SMD = 0.37 [0.04, 0.69]), as well as there was a trend for guanfacine ($k = 1$, $n = 30$; SMD = 0.50 [0.00, 1.01]) and riluzole ($k = 1$, $n = 29$; SMD = 0.43 [−0.09, 0.95]). On the other hand, irritability was worsened by vitamin-B12 ($k = 1$, $n = 27$; SMD = −0.62 [−1.19, −0.05]) and levetiracetam ($k = 1$, $n = 10$; SMD = -1.47 [−2.48, −0.46]).

In adults, risperidone was found efficacious ($k = 1$, $n = 14$; SMD = 1.19 [0.34, 2.04]), and heterogeneity was moderate ($\tau^2 = 0.028$).

ADHD symptoms

ADHD symptoms were measured in the majority of the studies with ABC-H (79%).

In children/adolescents, ADHD symptoms were improved by olanzapine ($k = 1$, $n = 6$; SMD = 2.08 [0.48, 3.68], *based only on indirect evidence*), guanfacine ($k = 1$, $n = 30$; SMD = 1.39 [0.73, 2.05]), aripiprazole ($k = 7$, $n = 363$; SMD = 0.82 [0.59, 1.05]), risperidone ($k = 5$, $n = 155$; SMD = 0.79 [0.47, 1.11]), naltrexone ($k = 1$, $n = 23$; SMD = 0.85 [0.12, 1.59]), and atomoxetine ($k = 3$, $n = 107$; SMD = 0.64 [0.30, 0.99]), as well as a trend was noted for sulforaphane ($k = 1$, $n = 12$; SMD = 0.88 [−0.03, 1.80]). Heterogeneity was moderate ($\tau^2 = 0.032$).

In adults, none of the investigated medications were found efficacious for ADHD symptoms, and heterogeneity was low ($\tau^2 = 0$).

Anxiety/depressive symptoms

Different scales measured anxiety/depression in children/adolescents (e.g., CBCL-I, BASC-I, CASI, DBC-Anxiety), and STAI-state was used in half of the studies in adults. None of the investigated medications found to improve anxiety or depressive symptoms, except for a trend about risperidone in adults ($n = 1$, $k = 14$; SMD = 0.67 [−0.07, 1.41]). There were moderate-to-high levels of heterogeneity in children/adolescents ($\tau^2 = 0.041$) and low in adults ($\tau^2 = 0$).

Caregiver stress

Caregiver stress was measured mainly with PSI (36%), CSQ (22%) and CGSQ (14%) in children/adolescents, and with PedsQL-Family Impact in adults. In children/adolescents, it was reduced by melatonin ($k = 1$, $n = 54$; SMD = 0.51 [0.12, 0.91]), and there were trends of small improvements by cannabinoids ($k = 1$, $n = 80$; SMD = 0.32 [−0.06, 0.69]) and atomoxetine ($k = 3$, $n = 104$; SMD = 0.21 [−0.06, 0.48]). There were no clear differences between other medications and placebo in both age groups, and heterogeneity was low ($\tau^2 = 0$).

Global functioning

Global functioning was measured with GAF or CGAS. In children/adolescents, it was improved by risperidone ($k = 3$, $n = 62$, SMD = 0.83 [0.40, 1.26]) and aripiprazole ($k = 2$, $n = 69$, SMD = 0.75 [0.33, 1.17]). No clear differences between other investigated medications and placebo were found in both age groups. Heterogeneity was moderate in children/adolescents ($\tau^2 = 0.016$) and low in adults ($\tau^2 = 0$).

Quality of life

Quality of life was measured with PedsQL in children/adolescents, and with PedsQL (40%) and WHO-QOL (60%) in adults. There were no clear differences between medications and placebo in children/adolescents. In adults, quality of life was improved by balovaptan ($k = 2$, $n = 217$; SMD = 0.22 [0.02, 0.43]), and potentially by oxytocin ($k = 3$, $n = 41$; SMD = 0.44 [−0.02, 0.90]). Heterogeneity was low in both age groups ($\tau^2 = 0$).

Response

Pairwise meta-analyses were conducted in children/adolescents due to incoherence (50% of the closed loops were incoherent; $p\text{-}_{\text{design-by-treatment}} = 0.068$). In comparison with placebo, more participants responded with risperidone ($k = 5$, $n = 161$; OR = 11.33 [4.99, 25.70]; $\tau^2 = 0.294$), guanfacine ($k = 1$, $n = 30$; OR = 9.67 [2.41, 38.71]), whey-protein ($k = 1$, $n = 22$; OR = 4.56 [1.25, 16.63]), aripiprazole ($k = 5$, $n = 317$; OR = 4.26 [2.32, 7.83]; $\tau^2 = 0.212$), vitamin-B12 ($k = 1$, $n = 28$; OR = 3.83 [1.20, 12.28]), atomoxetine ($k = 3$, $n = 109$; OR = 3.18 [1.56, 6.48]; $\tau^2 = 0$), melatonin ($k = 1$, $n = 60$; OR = 3.06 [1.38, 6.77]), bumetanide ($k = 3$, $n = 155$; OR = 2.78 [1.48, 5.21]; $\tau^2 = 0$), and cannabinoids ($k = 1$, $n = 100$; OR = 2.56 [1.15, 5.70]), while fewer with oral human immunoglobulins (IGOH) ($k = 1$, $n = 94$; OR = 0.40 [0.16, 0.99]). There were no clear differences for other medications.

In adults, there were more responders with risperidone ($k=1$, $n=15$; OR = 37.40 [1.62, 865.22]) and fluvoxamine ($k=1$, $n=15$; OR = 35.13 [1.52, 814.72]. There were high levels of heterogeneity ($\tau^2=0.257$).

Dropouts due to any cause

In children/adolescents, fewer overall dropouts were noted with risperidone ($k=10$, $n=274$; OR = 0.38 [0.22, 0.65]), lurasidone ($k=1$, $n=100$; OR = 0.35 [0.14, 0.88]) and aripiprazole ($k=8$, $n=399$; OR = 0.46 [0.29, 0.75]), as well as potentially with melatonin ($k=4$, $n=239$; OR = 0.52 [0.26, 1.03]). More dropouts were observed with arbaclofen ($k=1$, $n=76$; OR = 3.39 [1.16, 9.88]), and a trend was noted for fluoxetine ($k=3$, $n=161$; OR = 1.59 [0.97, 2.58]). There were no clear differences for other medications, and there were some indications of incoherence (12.5% of the loops were incoherent; $p\text{-}_{\text{design-by-treatment}}=0.334$). In adults, there were no clear differences for the investigated medications. Heterogeneity was low in both age groups ($\tau^2=0.006$ and $\tau^2=0$).

Dropouts due to adverse events

There were no clear differences between investigated medications and placebo in both age groups, and heterogeneity was low ($\tau^2=0$).

Any adverse event

In children/adolescents, more participants had adverse events with risperidone ($k=4$, $n=123$; OR = 4.74 [2.24, 10.04]), citalopram ($k=1$, $n=73$; OR = 5.38 [1.14, 25.46]), fluvoxamine ($k=1$, $n=18$; OR = 4.50 [1.02, 19.90]) and aripiprazole ($k=6$, $n=348$; OR = 2.62 [1.65, 4.15]), as well as potentially with guanfacine ($k=1$, $n=30$; OR = 17.94 [0.98, 329.56]) and lurasidone ($k=1$, $n=100$; OR = 1.92 [0.95, 3.90]). In adults, more participants had adverse events with risperidone ($k=1$, $n=15$; OR = 14.30 [2.19, 93.37]). There were no clear differences between other medications and placebo. Heterogeneity was low in children/adolescents ($\tau^2=0$) and moderate in adults ($\tau^2=0.049$).

Sedation

In children/adolescents, pairwise meta-analyses were conducted due to incoherence (75% of the closed loops were incoherent; $p\text{-}_{\text{design-by-treatment}}=0.051$). More participants had sedation with guanfacine ($n=1$, $k=30$; OR = 62.83 [12.84, 307.45]), haloperidol ($n=1$, $k=20$; OR = 44.33 [4.78, 410.96]), risperidone ($n=4$; $k=142$, OR = 11.95 [5.86, 24.36], $\tau^2=0$), aripiprazole ($n=5$, $k=317$; OR = 3.56 [1.62, 7.86]; $\tau^2=0$) and melatonin ($n=1$, $k=60$; OR = 3.28 [1.25, 8.59]).

In adults, there were no clear differences, and heterogeneity was low ($\tau^2=0$).

Weight gain

In children/adolescents, there was evidence of incoherence (50% of the closed loops were incoherent; $p\text{-}_{\text{design-by-treatment}}=0.032$) and pairwise meta-analyses were conducted. More participants had weight gain with aripiprazole ($n=5$, $k=317$; OR = 3.78 [2.09, 6.84], $\tau^2=0$) and risperidone ($n=5$, $k=161$; OR = 3.39 [1.80, 6.38], $\tau^2=0$) in comparison with placebo, while aripiprazole caused less weight gain in comparison with risperidone ($n=2$, $k=104$; OR = 0.22 [0.09, 0.55], $\tau^2=0.045$). There were no clear differences between other medications.

In adults, none of the investigated medications (sulforaphane, oxytocin and balovaptan) was associated with weight gain, and heterogeneity was low ($\tau^2=0$).

Extrapyramidal symptoms

The network of children/adolescents was disconnected; therefore, pairwise meta-analyses were conducted. In comparison with placebo, more participants had extrapyramidal symptoms with risperidone ($n=4$, $k=142$; OR = 3.02 [1.22, 7.48]; $\tau^2=0$) and aripiprazole ($n=4$, $k=300$; OR = 2.38 [1.18, 4.77]; $\tau^2=0$).

There were no data available for adults.

Discussion

This is the first comprehensive network meta-analysis on pharmacological and dietary-supplement interventions for ASD. Pediatric and adult populations were analyzed separately, in order to avoid misleading extrapolations [52]. Core symptom domains (SCD and RB) were also examined separately as co-primary outcomes, since differential treatment responses can be expected [52]. In addition, scales that measure overall core symptoms (OCS) in single scores were considered as a distinct outcome. Associated symptoms and side-effects were also investigated as secondary outcomes. Therefore, our analysis provides a more comprehensive synthesis of evidence in comparison with previous reviews that were mainly focused on pediatric populations, certain symptoms or specific medications, or did not utilize a network meta-analysis [9–17, 19, 21, 23, 24].

Our review identified the following medications that could improve at least one core symptom domain: aripiprazole (SCD, RB), atomoxetine (RB), bumetanide (RB, OCS) and risperidone (SCD, RB, OCS) in children/adolescents; fluoxetine (RB), fluvoxamine (RB), oxytocin (RB) and risperidone (RB) in adults. In addition, there were some indications of improvement by carnosine, haloperidol, folinic acid, guanfacine, omega-3-fatty-acids, probiotics, sulforaphane, tideglusib and valproate,

yet they were imprecise based on limited data and not formally statistically significant, as well as not robust in sensitivity analysis.

Summary of evidence

Commonly used medications

Currently, no medication is approved for the core symptoms of ASD [39]. However, about half of the individuals with ASD receive psychotropic drugs, mainly for associated symptoms, such as antipsychotics (median prevalence of 18.1%), ADHD medications (16.6%), antidepressants (17.2%), antiepileptics/mood-stabilizers and sleep medication [8]. Findings of our analysis on these medications are summarized in Fig. 2, facilitating intuitive understanding of the current evidence.

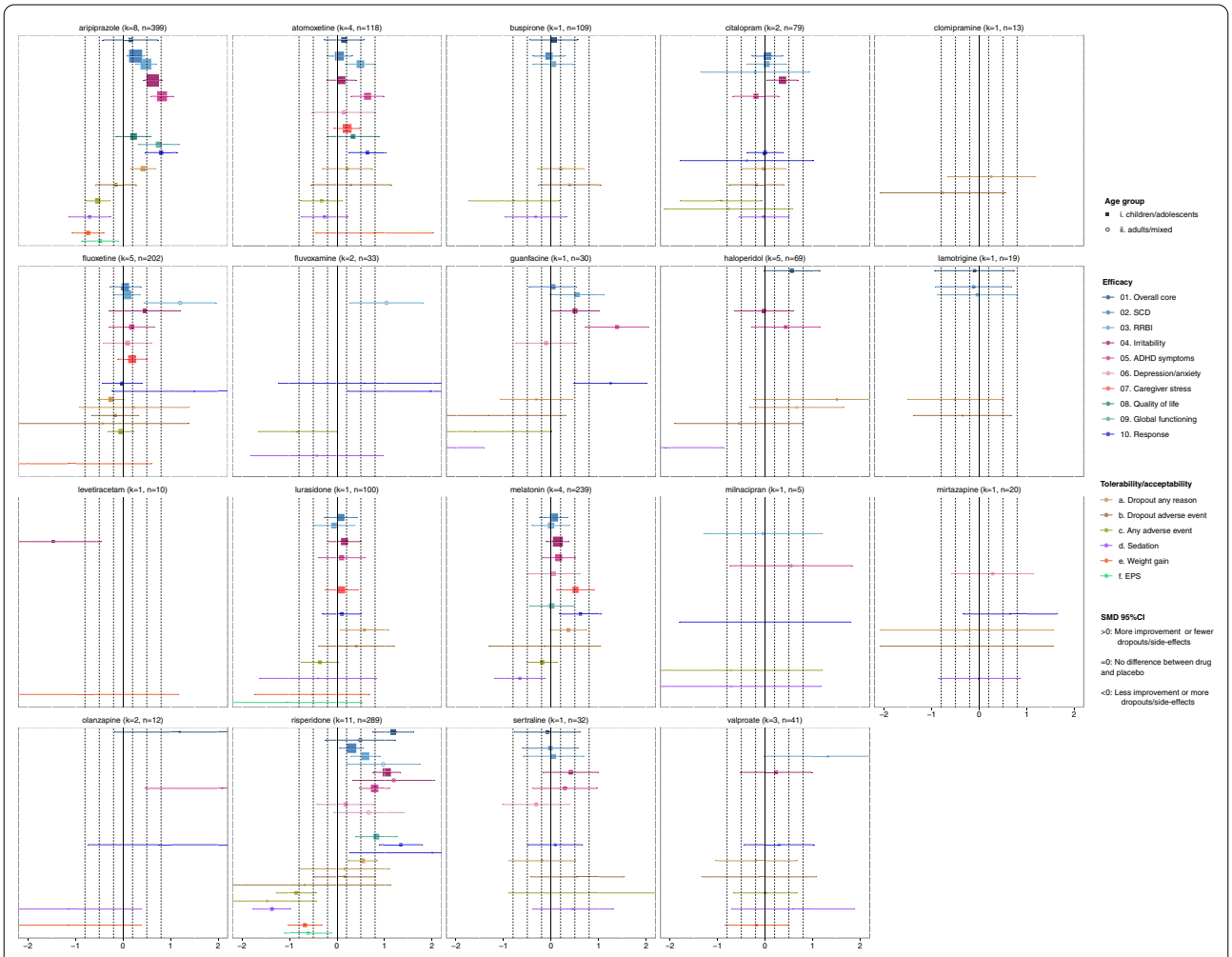


Fig. 2 Summary forest plots for commonly used medications, i.e., antipsychotics, ADHD medications, antiepileptic/mood-stabilizers, sleep medications. Effect-sizes (standardized mean differences—SMDs and their 95% confidence intervals) of comparisons with placebo are presented for each medication, outcome and age group. SMDs are presented with squares in children/adolescents and circles in adults, and their size is proportional to the inverse standard error of the effect size. For dichotomous outcomes (response, dropouts due to any cause or adverse event, any adverse event, sedation, weight gain, extrapyramidal symptoms), odds ratios were converted to SMDs. The results are based on network meta-analysis, except for irritability, response, sedation, weight gain and extrapyramidal symptoms (EPS) in children/adolescents, since pairwise meta-analyses were conducted due to incoherence or disconnected networks. SMDs > 0 indicate more improvement or fewer dropouts/adverse events with the medication in comparison with placebo, SMDs = 0 indicate no difference between medication and placebo, and SMDs < 0 indicate less improvement or more dropouts/adverse events with the medication in comparison with placebo. SMDs could be interpreted as small (SMD = |0.2|), medium (SMD = |0.5|) and large (SMD = |0.8|), and these thresholds are presented with dashed lines. There were no usable data for methylphenidate, and effect-sizes for this drug are not presented. *k* = total number of studies for the intervention with data for at least an outcome and age group, *n* = total number of participants on the intervention with data for at least an outcome and age group. *EPS* extrapyramidal symptoms, *RB* repetitive behaviors, *SCD* social-communication difficulties

Among antipsychotics, aripiprazole and risperidone demonstrated medium-to-large effect-sizes in reducing irritability and ADHD symptoms, while smaller improvements were found in social-communication difficulties and repetitive behaviors. On the other hand, lurasidone was in general not efficacious, and there were only a few data available for olanzapine and haloperidol, and for adults. Antipsychotics were also associated with more adverse events, sedation, weight gain and extrapyramidal symptoms. Nevertheless, reporting bias was suspected (Additional file 1: eAppendix-6.8), e.g., two pediatric studies found that risperidone did not improve social-communication difficulties as measured with ABC-L/SW, yet there were no usable data for this analysis [60, 61]. In addition, trials on antipsychotics were conducted mainly in participants with irritability. As a result, improvements in core symptoms could be collateral to the reduction in interfering challenging behaviors that can subsequently allow participation in social interactions [62]. In other words, antipsychotics may not have direct effects on core symptoms, but rather secondary to the reduction in irritability. Trials focusing on core symptoms are sparse, and data from a small trial ($n=41$) investigating risperidone for repetitive behaviors are not yet reported [63]. Therefore, evidence was downrated due to indirectness and reporting bias (Additional file 1: eAppendix-6.9).

Among ADHD medications, atomoxetine and guanfacine improved ADHD symptoms and potentially repetitive behaviors, but not social-communication difficulties. Guanfacine was also associated with more adverse events and sedation. A causal-mediation analysis suggested a causal link from hyperactivity to repetitive behaviors and from impulsivity/inattention to social-communication difficulties in ASD [64]. Therefore, and since these drugs were investigated in participants with ADHD symptoms, improvements in repetitive behaviors could be indirect and subsequent to the reduction in hyperactivity. Of note, there were no usable data for methylphenidate, since none of the five crossover trials reported usable data from the first phase (Additional file 1: eAppendix-6.8), and none of the ADHD medications were investigated in adults.

Antidepressants and buspirone were not found efficacious for core or associated symptoms in children/adolescents, except citalopram that improved irritability with a small-to-medium effect-size. Citalopram, fluvoxamine and fluoxetine were also associated with more adverse events or dropouts. In adults, however, fluoxetine and fluvoxamine improved repetitive behaviors with large effect-sizes, yet based on single small ($n=30$–37) studies [65, 66]. Apart from the limited data for adults, such differences might be explained by different study designs, participant characteristics and age-dependent variability in treatment response [67, 68].

Antiepileptics/mood-stabilizers were in general not efficacious based on limited and very low-quality data. A single small study ($n=13$) suggested efficacy for valproate [69], yet there was reporting bias and two additional studies did not report appropriate data [70, 71] (Additional file 1: eAppendix-6.8). Of note, levetiracetam worsened irritability with a large effect-size in a small study ($n=12$) [72], in accordance with the well-documented behavioral side-effects of this drug [73]. Last, melatonin was not efficacious for core or associated symptoms, yet it decreased caregiver stress and increased the number of responders. Such beneficial effects could be collateral to the reduction in sleep problems [2, 74, 75]. Sleep outcomes were not investigated in this review, but our findings support its sedative effects.

Experimental medications

Our review identified a considerable number of experimental medications (Fig. 3) with diverse mechanisms of action, which discussion is out of the scope of this review (e.g., see [39, 76–79]). The majority of them were investigated exclusively in children/adolescents, except for oxytocin and balovaptan.

Oxytocin and balovaptan (vasopressin-V_{1A} receptor antagonist) were not efficacious in children/adolescents, based on substantial evidence from large trials, e.g., ($n=290$–339) [80, 81]. In adults, however, oxytocin improved repetitive behaviors with small-to-medium effect-sizes and moderate-quality evidence. This finding needs replication, since studies were mainly focused on high-functioning participants and variability in treatment response due to age-dependent differences in the oxytocin system cannot be excluded [82, 83]. Balovaptan was not found efficacious in adults based on two large studies ($n=223$–322) [84, 85], yet small improvements in quality of life were noted. Of note, intranasal vasopressin was efficacious in a small trial ($n=30$) [86], which was, however, excluded from our analysis due to unconcealed allocation (Additional file 1: eAppendix-4.2.).

Bumetanide (loop-diuretic that may enhance GABAergic inhibition) was found to improve repetitive behaviors and overall core symptoms with small-to-medium effect-sizes, but not social-communication difficulties. However, two large phase-III trials ($n=422$ in total) [87] were negative and prematurely terminated [88], yet they did not report usable data, and therefore, evidence was downrated due to reporting bias. Other experimental medications were not found efficacious based on current data. There were some indications for cannabinoids (more participants had a positive response), and naltrexone (improvement of ADHD symptoms),

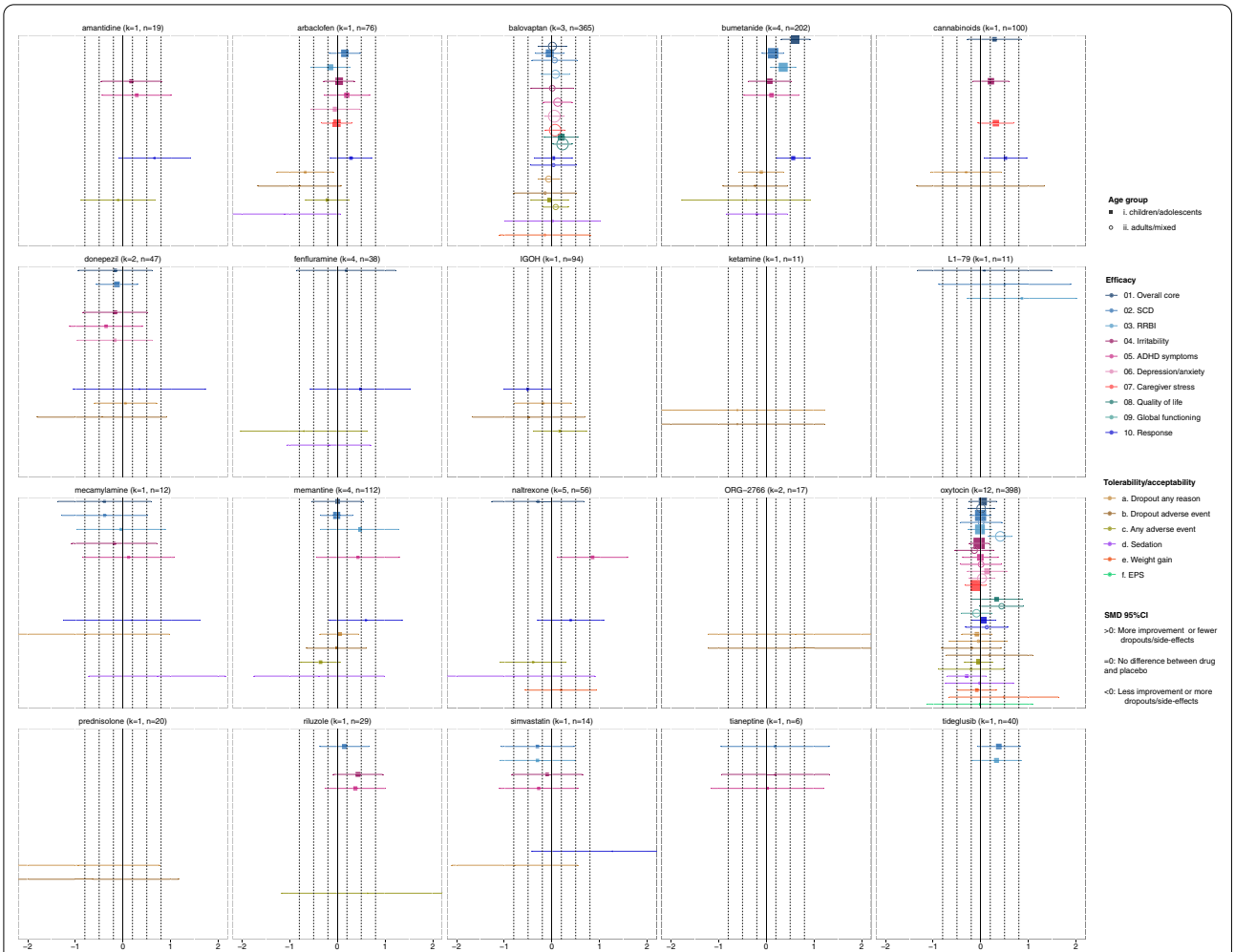


Fig. 3 Summary forest plots for experimental medications. Effect-sizes (standardized mean differences—SMDs and their 95% confidence intervals) of comparisons with placebo are presented for each medication, outcome and age group. SMDs are presented with squares in children/adolescents and circles in adults, and their size is proportional to the inverse standard error of the effect size. For dichotomous outcomes (response, dropouts due to any cause or adverse event, any adverse event, sedation, weight gain, extrapyramidal symptoms), odds ratios were converted to SMDs. The results are based on network meta-analysis, except for irritability, response, sedation, weight gain and extrapyramidal symptoms (EPS) in children/adolescents, since pairwise meta-analyses were conducted due to incoherence or disconnected networks. SMDs > 0 indicate more improvement or fewer dropouts/adverse events with the medication in comparison with placebo, SMDs = 0 indicate no difference between medication and placebo, and SMDs < 0 indicate less improvement or more dropouts/adverse events with the medication in comparison with placebo. SMDs could be interpreted as small (SMD = |0.2|), medium (SMD = |0.5|) and large (SMD = |0.8|), and these thresholds are presented with dashed lines. There were no usable data for dextromethorphan/quinidine and effect-sizes for this drug are not presented. *k* = total number of studies for the intervention with data for at least an outcome and age group; *n* = total number of participants on the intervention with data for at least an outcome and age group; EPS: extrapyramidal symptoms, IGOH: oral human immunoglobulin; RB: repetitive behaviors; SCD: social-communication difficulties

yet they were based on single studies [89, 90] and there was also reporting bias for naltrexone (Additional file 1: eAppendix-6.8). On the other hand, arbaclofen ($GABA_B$ agonist) was associated with more dropouts and IGOH (oral human immunoglobulin) with fewer responders. Nevertheless, several trials are ongoing, e.g., for arbaclofen [91, 92], memantine [93] and cannabinoids [94, 95]. In addition, the findings on tideglusib (GSK-3β inhibitor), L1-79 (tyrosine hydroxylase inhibitor) and riluzole could be imprecise, since data from abstracts were used [96–98].

Dietary-supplements

The efficacy of dietary-supplements was inconclusive (Fig. 4). Omega-3-fatty-acids could potentially improve social-communication difficulties with small effect-sizes, based on very low-quality evidence from ten studies in children/adolescents. Similarly, there were some trends

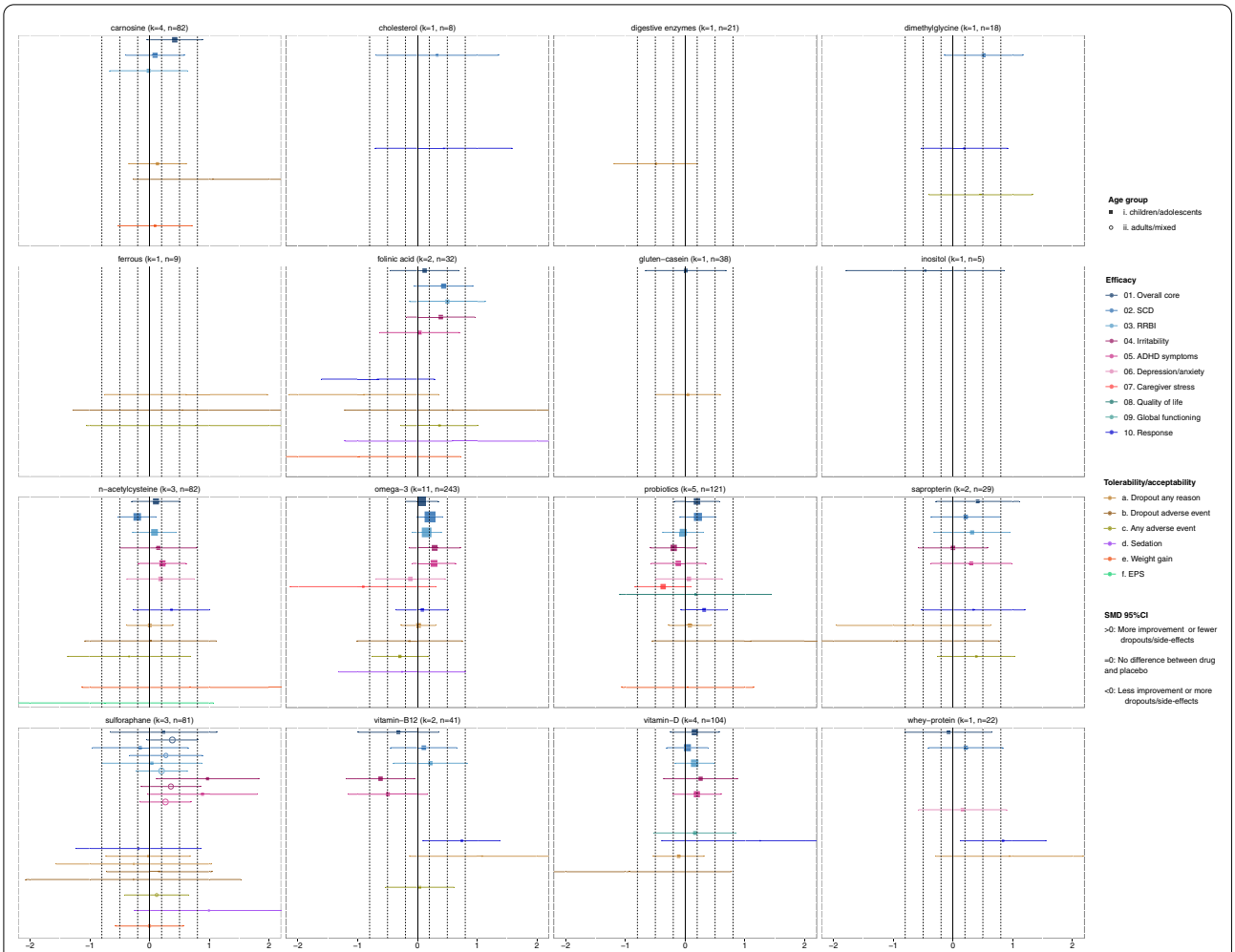


Fig. 4 Summary plots for dietary-supplements. Effect-sizes (standardized mean differences—SMDs and their 95% confidence intervals) of comparisons with placebo are presented for each medication, outcome and age group. SMDs are presented with squares in children/adolescents and circles in adults, and their size is proportional to the inverse standard error of the effect size. For dichotomous outcomes (response, dropouts due to any cause or adverse event, any adverse event, sedation, weight gain, extrapyramidal symptoms), odds ratios were converted to SMDs. The results are based on network meta-analysis, except for irritability, response, sedation, weight gain and extrapyramidal symptoms (EPS) in children/adolescents, since pairwise meta-analyses were conducted due to incoherence or disconnected networks. SMDs > 0 indicate more improvement or fewer dropouts/adverse events with the medication in comparison with placebo, SMDs = 0 indicate no difference between medication and placebo, and SMDs < 0 indicate less improvement or more dropouts/adverse events with the medication in comparison with placebo. SMDs could be interpreted as small (SMD = |0.2|), medium (SMD = |0.5|) and large (SMD = |0.8|), and these thresholds are presented with dashed lines. There were no usable data for pyridoxine, and effect-sizes for this dietary-supplement are not presented. *k* = total number of studies for the intervention with data for at least an outcome and age group; *n* = total number of participants on the intervention with data for at least an outcome and age group. *EPS* extrapyramidal symptoms, *IGOH* oral human immunoglobulin, *RB* repetitive behaviors, *SCD* social-communication difficulties

for carnosine, folinic acid and probiotics, based on fewer data. Nevertheless, these findings were highly heterogeneous (for carnosine and folinic acid), imprecise and not statistically significant (at two-sided alpha 0.05), and not robust in sensitivity analyses. Therefore, results from larger trials are warranted, e.g. [99, 100]. There was also mixed evidence about sulforaphane (broccoli sprout extract), since findings were based on one inconclusive trial ($n = 45$) in children/adolescents [101], and two contradicting trials ($n = 44$–48) in adults [102, 103], while usable data from a larger trial ($n = 110$) are not yet reported [104]. In addition, there were some indications from single studies for cysteine-rich whey-protein [105] and vitamin-B12 [106], since both increased the number of responders but were not found to be efficacious for core or associated symptoms. On the contrary, vitamin-B12 worsened irritability with a medium effect-size, which is in line with a meta-analysis of prevalence

that identified its potential behavioral side-effects [107]. Therefore, the safety of dietary-supplements should not be overlooked.

Limitations

There are certain limitations. First, and in contrast with other fields of psychopharmacology, evidence base of ASD is flooded by small trials focusing on associated symptoms and investigating a plethora of medication classes, for which adequate dosing or duration of treatment is still unclear, and some of them have not yet investigated in RCTs. This reflects the two main approaches that guide psychopharmacology in ASD, by re-purposing psychotropics for associated symptoms or by targeting neurobiological processes [2]. Nevertheless, ASD may not be a unitary diagnosis in terms of neurobiology, according to its heterogeneity and lack of biomarkers. Therefore, it is likely that there is substantial interpersonal variability of treatment response across medications. Individual-participant-data meta-analyses could further explore this issue and investigate the potential impact of participant-level covariates [108], e.g., age, sex, baseline severity of core and associated symptoms. In that direction, there are also efforts to disentangle the neurobiology of subgroups within ASD in order to facilitate biomarker stratification and more targeted treatments [39, 77].

Second, clinical trials in ASD could be prone to substantial placebo responses and a lower ability to detect efficacy, which may be increased with adequately powered trials, rigorous selection of participants and careful selection of outcome measures [4]. In line with this, there is lack of consensus on outcome measures [39], and different scales are often used. We accepted a wide range of validated scales in order to incorporate more evidence, yet we preferred recommended and commonly used scales in order to obtain comparable measures (Additional file 1: eAppendix-5.3). As a result, data for most of the outcomes were derived mainly from one or two scales (Additional file 1: eAppendix-5.1), which treatment effects might agree in most cases, e.g., as suggested between CYBOCS and ABC-S [39]. The results were also generally robust in sensitivity analyses when clinician-ratings or when ABC subscales were used, except for some potential differences in omega-3-fatty-acids (Additional file 1: eAppendix-6.6, Additional file 6: Fig. S4). Nevertheless, further research is needed, since scales with different psychometric properties, e.g., sensitivity to change or susceptibility to placebo effects, could demonstrate discordant treatment effects. For example, a trial found low-dose buspirone to improve repetitive behaviors as measured with ADOS-RRB and RBS, but not with CYBOCS [109], which was preferred in our analysis according to our hierarchy (Additional file 1: eAppendix-5.3).

Third, there were limited data for adults, some medications, e.g., methylphenidate, and secondary outcomes, e.g., anxiety/depressive symptoms, which are, however, considered one of the top research priorities [110, 111]. Fourth, our analysis was mainly based on star-shaped networks of placebo-controlled comparisons and only a few medications were investigated in more than one or two trials, often with small sample sizes. Therefore, heterogeneity and incoherence could be masked, due to the low statistical power of their tests. Small-study effects could also be masked, since comparison-adjusted funnel plots should be interpreted with great caution when there are a few trials per comparison. Fifth, transitivity assumption could not be adequately assessed, since effect-modifiers are still unclear and insufficiently reported in clinical trials. Therefore, and despite of ordering treatments by their ranking in forest plots, indirect evidence, treatment hierarchies and league tables should be interpreted with great caution. There was also evidence of incoherence in irritability, response, sedation and weight gain in children/adolescents; therefore, pairwise meta-analyses were conducted for these outcomes. In addition, about half of the studies stated to be randomized without an exact description of the randomization method, yet the results did not materially change in sensitivity analyses when studies with an unclear risk of bias in random sequence generation or allocation concealment were excluded (Additional file 1: eAppendix-6.6, Additional file 6: Fig. S4).

Last, a comprehensive review of tolerability was beyond the scope of the manuscript, yet we examined dropouts and important side-effects that overlap among drug classes, i.e., sedation, weight gain and extrapyramidal symptoms, and our findings are in line with the literature [112]. Nevertheless, medications with different mechanisms of action can have unique side-effect profiles, e.g., bumetanide as a loop-diuretic can cause diuresis and hypokalemia [113]. Individuals with ASD may also be more sensitive to side-effects in comparison with neurotypical individuals [2]. Therefore, medications should be used after careful consideration and monitoring of their safety [2], as well as at low doses, since a therapeutic window could be expected, e.g., for risperidone [114].

Conclusions

In conclusion, there was evidence that some medications could improve social-communication difficulties and/or repetitive behaviors in children/adolescents: aripiprazole, atomoxetine, bumetanide, and risperidone; while some medications could improve repetitive behaviors in adults: fluoxetine, fluvoxamine, oxytocin and risperidone.

A large part of the evidence consisted of small RCTs (median 40 participants) with a short duration (median 12 weeks) and limited generalizability. Therefore, current commonly used medications, i.e., antipsychotics and ADHD medications, can be used for associated symptoms as indicated, and smaller improvements in core symptoms could also be expected, at least collaterally to the improvement of challenging behaviors. These medications are associated with side-effects, and therefore, they should be prescribed only after careful consideration and monitoring of their benefit-risk ratio. Evidence on the efficacy and safety for other medications, including bumetanide, oxytocin and some dietary-supplements, is at best preliminary and warrants further investigation. In line with the limitations of our review, there are current efforts to advance clinical psychopharmacology in ASD (e.g., within the AIMS-2-Trials consortium or the ISCTM/ECNP ASD working group), first with the elucidation of its neurobiology and the development of more targeted medications, second with the use of appropriate scales for measuring core symptoms, and third with well-designed and adequately powered clinical trials [39, 77].

Abbreviations

95% CI: 95% Confidence intervals; ASD: Autism spectrum disorder; ABC-H: Aberrant Behavior Checklist—Hyperactivity/noncompliance; ABC-I: Aberrant Behavior Checklist—Irritability; ABC-L/SW: Aberrant Behavior Checklist—Lethargy/Social-Withdrawal; ABC-S: Aberrant Behavior Checklist—Stereotypic behavior; ADHD: Attention Deficit Hyperactivity Disorder; ADI-R: Autism Diagnostic Interview-Revised; ADOS-RRB: Autism Diagnostic Observation Schedule—Restricted and Repetitive Behaviors; BASC-I: Behavior Assessment System for Children—Internalizing; CARS: Childhood Autism Rating Scale; CASI: Childhood Anxiety Sensitivity Index; CBCL-I: Child Behavior Checklist—Internalizing; CGAS: Children Global Assessment Scale; CGI: Clinical Global Impression; CGSQ: Caregiver Strain Questionnaire; CINeMA: Confidence in Network Meta-Analysis; CSQ: Client Satisfaction Questionnaire; CYBOCS(-PDD): Children's Yale-Brown Obsessive Compulsive Scale (Modified for Pervasive Developmental Disorders); DBC-Anxiety: Developmental Behavior Checklist—Anxiety; DSM: Diagnostic and Statistical Manual of Mental Disorders; GAF: Global Assessment of Functioning; FDA: Food and Drug Administration; ICD: International Classification of Diseases; IGOH: Oral human immunoglobulin; IQR: Interquartile range; OCS: Overall core symptoms; OR: Odds ratio; LOCF: Last-observation carried forward; MMRM: Mixed-models for repeated measures; PedsQL: Pediatric Quality of Life Inventory; PSI: Parental Stress Index; RB: Repetitive Behaviors; RBS: Repetitive Behavior Scale; RCT: Randomized Controlled Trials; SCD: Social Communication difficulties; SD: Standard deviation; SMD: Standardized Mean Differences; SRS: Social Responsiveness Scale; STAI: State-Trait Anxiety Inventory; VABS: Vineland Adaptive Behavior Scale; WHO-QoL: World Health Organization-Quality of Life.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13229-022-00488-4.

Additional file 1. eAppendix.

Additional file 2. Fig. S1. Network plots.

Additional file 3. Fig. S2. Forest plots for secondary outcomes.

Additional file 4. Table S1. League tables.

Additional file 5. Fig. S3. Forest plots for pairwise meta-analysis and individual studies.

Additional file 6. Fig. S4. Sensitivity analyses.

Additional file 7. Dataset.

Acknowledgements

We would like to thank Farhad Shokraneh, former information specialist of the Cochrane Schizophrenia Group, who conducted the first search in electronic databases, Dr. Yikang Zhu for the translation of a Chinese study and Prof. Toshi Furukawa for the translation of a Japanese study. We would like to thank the following authors that contributed to the review by kindly providing additional data and/or clarifications regarding their studies: Adi Aran, Kaat Alaerts, Nadir Aliyev, Eugene Arnold, Haim Belmaker, Yéhézkel Ben-Ari, Leventhal Bennet, Stephen Bent, Helena Brentani, Jan Buitelaar, Ana Maria Castejon, Michael Chez, Torsten Danfors, Paulo Fontoura, Robert Grimaldi, Paul Gringras, Alexander Häge, Randi J. Hagerman, Antonio Hardan, Robert Hendren, Janet Kern, Bruno Leheup, Wenn Liu, Raquel Martinez, James McCracken, Tali Nir, Deborah Pearson, Laura Politte, Jeanette Ramer, Dan Rossignol, Kevin Sanders, Elisa Santocchi, Renato Scifo, Sarah Shea, Lawrence Scahill, Jeremy Veenstra-Vanderweele Paul Wang, David Wilensky, Hidenori Yamasue, and Lingli Zhang.

Authors' contributions

SS contributed to study design, study selection, data extraction, contacting authors for additional data, statistical analysis, interpretation of the data, drafting the first version of the manuscript, and study supervision; OC contributed to study selection, data extraction, and contacting authors for additional data; HW contributed to study selection, data extraction, and contacting authors for additional data; JST contributed to data extraction and technical support with Access database; IB contributed to study selection; MK contributed to study selection; AR contributed to data extraction; AC contributed to study selection; GD contributed to study selection; MH contributed to technical support with Access database; DF contributed to study design and interpretation of the data; AJCS contributed to interpretation of the data; DM contributed to statistical analysis and interpretation of the data; TC contributed to interpretation of the data; DGM contributed to interpretation of the data; MP contributed to study design and interpretation of the data; CA contributed to study design and interpretation of the data; SL contributed to study design, data extraction, statistical analysis, interpretation of the data, drafting the first version of the manuscript, and study supervision. All authors critically reviewed the manuscript for important intellectual content. The authors read and approved the final manuscript.

Funding

This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under Grant Agreement No. 777394 for the project AIMS-2-TRIALS. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. CA, MP and DF were supported by the Spanish Ministry of Science, Innovation and Universities. Instituto de Salud Carlos III, co-financed by ERDF Funds from the European Commission, "A way of making Europe", CIBERSAM, Madrid Regional Government (B2017/BMD-3740 AGES-CM-2), European Union Structural Funds and European Union Seventh Framework Program and H2020 Program; Fundación Familia Alonso, Fundación Alicia Koplowitz and Fundación Mutua Madrileña. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. Any views expressed are those of the authors and not necessarily those of the funders.

Availability of data and materials

The dataset used in the current study is available in this published article and its Additional file 7: Dataset.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication
Not applicable.

Competing interests
In the past 3 years Stefan Leucht has received honoraria as a consultant and/or advisor and/or for lectures from Alkermes, Angelini, Eisai, Gedeon Richter, Janssen, Lundbeck, Lundbeck Institute, Merck Sharpp and Dome, Otsuka, Recordati, Rovi, Sanofi Aventis, TEVA, Medichem, Mitshubishi. David Fraguas has been a consultant and/or has received fees from Angelini, Casen, Janssen, Lundbeck, and Otsuka. He has also received grant support from Instituto de Salud Carlos III (Spanish Ministry of Science and Innovation) and from Fundación Alicia Koplowitz.Mara Parellada has received educational honoraria from Otsuka, research grants from FAK and Fundación Mutua Madrileña (FMM), Instituto de Salud Carlos III (Spanish Ministry of Science, Innovation and Universities) and European ERANET and H2020 calls, travel grants from Otsuka and Janssen. Consultant for Exeltis and Servier. Celso Arango has been a consultant to or has received honoraria or grants from Acadia, Angelini, Gedeon Richter, Janssen Cilag, Lundbeck, Otsuka, Roche, Sage, Sanofi, Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion and Takeda. Maximilian Huhn has received speakers honoraria from Janssen. Declan Murphy has received consulting fees from Roche. Antonia San José Cáceres has been a consultant for Roche and is currently involved in clinical trials with Servier. The other authors have nothing to disclose.

Author details
[1]Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Ismaningerstr. 22, 81675 Munich, Germany. [2]Department of Child and Adolescent Psychiatry, Mardin State Hospital, Artuklu, Mardin, Turkey. [3]Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Munich, Germany. [4]Department of Experimental and Clinical Medicine, Psychiatric Clinic University Hospital 'Gaspare Rodolico', University of Catania, Catania, Italy. [5]Department of Psychiatry, Spedali Civili Hospital, Brescia, Italy. [6]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, Social Foundation Bamberg, Teaching Hospital of the University of Erlangen, Bamberg, Germany. [7]Institute of Psychiatry and Mental Health, Hospital Clínico San Carlos, IdISSC CIBERSAM, School of Medicine, Universidad Complutense, Madrid, Spain. [8]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, Madrid, Spain. [9]Department of Primary Education, University of Ioannina, Ioannina, Greece. [10]Faculté de Médecine, Université Paris Descartes, Paris, France. [11]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [12]Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [13]School of Medicine, Universidad Complutense, Madrid, Spain. [14]Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain. [15]Centro Investigación Biomédica en Red Salud Mental (CIBERSAM), Madrid, Spain.

Received: 15 April 2021 Accepted: 2 February 2022
Published online: 04 March 2022

References

1. American Psychiatric A. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub; 2013.
2. Howes OD, Rogdaki M, Findon JL, Wichers RH, Charman T, King BH, et al. Autism spectrum disorder: consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. J Psychopharmacol. 2018;32(1):3–29. https://doi.org/10.1177/0269881117741766.
3. Tromans S, Adams C. Brief report: Autism spectrum disorder: a comprehensive survey of randomized controlled trials. J Autism Dev Disord. 2018;48(9):3228–32.
4. Siafis S. Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis. Mol Autism. 2020;11:66.
5. Díaz-Caneja CM, State MW, Hagerman RJ, Jacquemont S, Marín O, Bagni C, et al. A white paper on a neurodevelopmental framework for drug discovery in autism and other neurodevelopmental disorders. Eur Neuropsychopharmacol. 2021. https://doi.org/10.1016/j.euroneuro.2021.02.020.
6. Fuentes J, Hervás A, Howlin P. ESCAP practice guidance for autism: a summary of evidence-based recommendations for diagnosis and treatment. Eur Child Adolesc Psychiatry. 2021;30(6):961–84. https://doi.org/10.1007/s00787-020-01587-4.
7. Berry-Kravis EM, Lindemann L, Jonch AE, Apostol G, Bear MF, Carpenter RL, et al. Drug development for neurodevelopmental disorders: lessons learned from fragile X syndrome. Nat Rev Drug Discov. 2018;17(4):280–99. https://doi.org/10.1038/nrd.2017.221.
8. Jobski K, Hofer J, Hoffmann F, Bachmann C. Use of psychotropic drugs in patients with autism spectrum disorders: a systematic review. Acta Psychiatr Scand. 2017;135(1):8–28. https://doi.org/10.1111/acps.12644.
9. Carrasco M, Volkmar FR, Bloch MH. Pharmacologic treatment of repetitive behaviors in autism spectrum disorders: evidence of publication bias. Pediatrics. 2012;129(5):e1301–10. https://doi.org/10.1542/peds.2011-3285.
10. Sturman N, Deckx L, van Driel ML. Methylphenidate for children and adolescents with autism spectrum disorder. Cochrane Database Syst Rev. 2017;11:011144. https://doi.org/10.1002/14651858.CD011144.pub2.
11. Fung LK, Mahajan R, Nozzolillo A, Bernal P, Krasner A, Jo B, et al. Pharmacologic treatment of severe irritability and problem behaviors in autism: a systematic review and meta-analysis. Pediatrics. 2016;137(Supplement 2):S124–35.
12. Hirsch LE, Pringsheim T. Aripiprazole for autism spectrum disorders (ASD). Cochrane Database Syst Rev. 2016;2016(6):CD009043. https://doi.org/10.1002/14651858.CD009043.pub3.
13. Williams K, Brignell A, Randall M, Silove N, Hazell P. Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). Cochrane Database Syst Rev. 2013;2013(8):CD004677. https://doi.org/10.1002/14651858.CD004677.pub3.
14. Hurwitz R, Blackmore R, Hazell P, Williams K, Woolfenden S. Tricyclic antidepressants for autism spectrum disorders (ASD) in children and adolescents. Cochrane Database Syst Rev. 2012;2012(3):CD008372. https://doi.org/10.1002/14651858.CD008372.pub2.
15. Nye C, Brice A. Combined vitamin B6-magnesium treatment in autism spectrum disorder. Cochrane Database Syst Rev. 2005;2005(4):CD003497. https://doi.org/10.1002/14651858.CD003497.pub2.
16. Fraguas D, Díaz-Caneja CM, Pina-Camacho L, Moreno C, Durán-Cutilla M, Ayora M, et al. Dietary interventions for autism spectrum disorder: a meta-analysis. Pediatrics. 2019;144:e20183218.
17. Zhou MS, Nasir M, Farhat LC, Kook M, Artukoglu BB, Bloch MH. Meta-analysis: pharmacologic treatment of restricted and repetitive behaviors in autism spectrum disorders. J Am Acad Child Adolesc Psychiatry. 2020. https://doi.org/10.1016/j.jaac.2020.03.007.
18. Fallah MS, Shaikh MR, Neupane B, Rusiecki D, Bennett TA, Beyene J. Atypical antipsychotics for irritability in pediatric autism: a systematic review and network meta-analysis. J Child Adolesc Psychopharmacol. 2019;29(3):168–80. https://doi.org/10.1089/cap.2018.0115.
19. Rodrigues R, Lai MC, Beswick A, Gorman DA, Anagnostou E, Szatmari P, et al. Practitioner review: pharmacological treatment of attention-deficit/hyperactivity disorder symptoms in children and youth with autism spectrum disorder: a systematic review and meta-analysis. J Child Psychol Psychiatry. 2020. https://doi.org/10.1111/jcpp.13305.
20. Hutton B, Catala-Lopez F, Moher D. The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA. Med Clin. 2016;147(6):262–6.
21. Masi A, Lampit A, DeMayo MM, Glozier N, Hickie IB, Guastella AJ. A comprehensive systematic review and meta-analysis of pharmacological and dietary supplement interventions in paediatric autism: moderators of treatment response and recommendations for future research. Psychol Med. 2017;47(7):1323–34. https://doi.org/10.1017/S0033291716003457.
22. Masi A, Lampit A, Glozier N, Hickie IB, Guastella AJ. Predictors of placebo response in pharmacological and dietary supplement treatment trials in pediatric autism spectrum disorder: a meta-analysis. Transl Psychiatry. 2015;5(9): e640. https://doi.org/10.1038/tp.2015.143.
23. Williamson E, Sathe NA, Andrews JC, Krishnaswami S, McPheeters ML, Fonnesbeck C, et al. AHRQ comparative effectiveness reviews. Medical

therapies for children with autism spectrum disorder—an update. Agency for Healthcare Research and Quality; 2017.
24. Song L, Luo X, Jiang Q, Chen Z, Zhou L, Wang D, et al. Vitamin D supplementation is beneficial for children with autism spectrum disorder: a meta-analysis. Clin Psychopharmacol Neurosci. 2020;18(2):203–13. https://doi.org/10.9758/cpn.2020.18.2.203.
25. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions. Wiley; 2019.
26. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
27. Devereaux PJ, Choi PT, El-Dika S, Bhandari M, Montori VM, Schünemann HJ, et al. An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. J Clin Epidemiol. 2004;57(12):1232–6. https://doi.org/10.1016/j.jclinepi.2004.03.017.
28. Hill CL, LaValley MP, Felson DT. Discrepancy between published report and actual conduct of randomized clinical trials. J Clin Epidemiol. 2002;55(8):783–6. https://doi.org/10.1016/s0895-4356(02)00440-7.
29. Soares HP, Daniels S, Kumar A, Clarke M, Scott C, Swann S, et al. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. BMJ. 2004;328(7430):22–4. https://doi.org/10.1136/bmj.328.7430.22.
30. Mhaskar R, Djulbegovic B, Magazin A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. J Clin Epidemiol. 2012;65(6):602–9. https://doi.org/10.1016/j.jclinepi.2011.10.016.
31. Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol. 2002;31(1):140–9.
32. Reichow B, Volkmar FR, Cicchetti DV. Development of the evaluative method for evaluating and determining evidence-based practices in autism. J Autism Dev Disord. 2008;38(7):1311–9.
33. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open. 2016;6(7): e010919. https://doi.org/10.1136/bmjopen-2015-010919.
34. Aman MG, Singh NN, Stewart AW, Field CJ. Psychometric characteristics of the aberrant behavior checklist. Am J Ment Defic. 1985;89(5):492–502.
35. Sparrow SS. Vineland adaptive behavior scales. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of clinical neuropsychology. Springer; 2011. p. 2618–21.
36. Scahill L, McDougle CJ, Williams SK, Dimitropoulos A, Aman MG, McCracken JT, et al. Children's Yale-Brown Obsessive Compulsive Scale modified for pervasive developmental disorders. J Am Acad Child Adolesc Psychiatry. 2006;45(9):1114–23. https://doi.org/10.1097/01.chi.0000220854.79144.e7.
37. Constantino JN, Gruber CP. Social Responsiveness Scale: SRS-2 Software Kit. Western Psychological Services 2012.
38. Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: childhood Autism Rating Scale (CARS). J Autism Dev Disord. 1980;10(1):91–103. https://doi.org/10.1007/BF02408436.
39. McCracken JT, Anagnostou E, Arango C, Dawson G, Farchione T, Mantua V, et al. Drug development for autism spectrum disorder (ASD): progress, challenges, and future directions. Eur Neuropsychopharmacol. 2021;48:3–31. https://doi.org/10.1016/j.euroneuro.2021.05.010.
40. Guy W. CGI. Clinical global impressions. ECDEU assessment manual for psychopharmacology. 1976.
41. Schwarzer G. Meta: an R package for meta-analysis. R News. 2007;7(3):40–5.
42. Rücker G, Schwarzer G, Krahn U, König J, Schwarzer MG. Package 'netmeta'. network meta-analysis using frequentist methods (version 07-0). 2015.
43. R Core Team. R: a language and environment for statistical computing. R Core Team; 2019.
44. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, et al. CINeMA: an approach for assessing confidence in the results of a network meta-analysis. PLoS Med. 2020;17(4): e1003082. https://doi.org/10.1371/journal.pmed.1003082.
45. Papakonstantinou T, Nikolakopoulou A, Higgins JPT, Egger M, Salanti G. CINeMA: software for semiautomated assessment of the confidence in the results of network meta-analysis. Campbell Syst Rev. 2020;16(1): e1080. https://doi.org/10.1002/cl2.1080.
46. Bakbergenuly I, Hoaglin DC, Kulinskaya E. Pitfalls of using the risk ratio in meta-analysis. Res Synth Methods. 2019;10(3):398–419. https://doi.org/10.1002/jrsm.1347.
47. Doi SA, Furuya-Kanamori L, Xu C, Lin L, Chivese T, Thalib L. Questionable utility of the relative risk in clinical research: a call for change to practice. J Clin Epidemiol. 2020. https://doi.org/10.1016/j.jclinepi.2020.08.019.
48. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. BMC Med Res Methodol. 2015;15(1):58. https://doi.org/10.1186/s12874-015-0060-8.
49. Samara MT, Spineli LM, Furukawa TA, Engel RR, Davis JM, Salanti G, et al. Imputation of response rates from means and standard deviations in schizophrenia. Schizophr Res. 2013;151(1–3):209–14. https://doi.org/10.1016/j.schres.2013.10.029.
50. Siafis S, Rodolico A, Çıray O, Murphy DG, Parellada M, Arango C, et al. Imputing the number of responders from the mean and standard deviation of CGI-improvement in clinical trials investigating medications for autism spectrum disorder. Brain Sci. 2021;11(7):908. https://doi.org/10.3390/brainsci11070908.
51. Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ. Empirical assessment of within-arm correlation imputation in trials of continuous outcomes. 2012.
52. European Medicines Agency. Guideline on the clinical development of medicinal products for the treatment of Autism Spectrum Disorder (ASD). 2018. p. 1–13.
53. Rhodes KM, Turner RM, Higgins JPT. Predictive distributions were developed for the extent of heterogeneity in meta-analyses of continuous outcome data. J Clin Epidemiol. 2015;68(1):52–60. https://doi.org/10.1016/j.jclinepi.2014.08.012.
54. Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. Int J Epidemiol. 2012;41(3):818–27. https://doi.org/10.1093/ije/dys041.
55. Efthimiou O, Debray TPA, van Valkenhoef G, Trelle S, Panayidou K, Moons KGM, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods. 2016;7(3):236–63. https://doi.org/10.1002/jrsm.1195.
56. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS ONE. 2013;8(10): e76654. https://doi.org/10.1371/journal.pone.0076654.
57. Panagiotou OA, Contopoulos-Ioannidis DG, Ioannidis JP. Comparative effect sizes in randomised trials from less developed and more developed countries: meta-epidemiological assessment. BMJ. 2013;346: f707. https://doi.org/10.1136/bmj.f707.
58. Stivaros S, Garg S, Tziraki M, Cai Y, Thomas O, Mellor J, et al. Randomised controlled trial of simvastatin treatment for autism in young children with neurofibromatosis type 1 (SANTA). Mol Autism. 2018;9:12. https://doi.org/10.1186/s13229-018-0190-z.
59. Nikvarz N, Alaghband-Rad J, Tehrani-Doost M, Alimadadi A, Ghaeli P. Comparing efficacy and side effects of memantine vs. risperidone in the treatment of autistic disorder. Pharmacopsychiatry. 2017;50(1):19–25. https://doi.org/10.1055/s-0042-108449.
60. Kent JM, Kushner S, Ning X, Karcher K, Ness S, Aman M, et al. Risperidone dosing in children and adolescents with autistic disorder: a double-blind, placebo-controlled study. J Autism Dev Disord. 2013;43(8):1773–83.
61. NCT01624675. A study to evaluate the efficacy and safety of risperidone (R064766) in children and adolescents with irritability associated with autistic disorder. 2015.
62. McDougle CJ, Scahill L, Aman MG, McCracken JT, Tierney E, Davies M, et al. Risperidone for the core symptom domains of autism: results from the study by the autism network of the research units on pediatric psychopharmacology. Am J Psychiatry. 2005;162(6):1142–8. https://doi.org/10.1176/appi.ajp.162.6.1142.

63. NCT01171937. Risperidone treatment in children with autism spectrum disorder and high levels of repetitive behavior (ProjectV). 2017.
64. Sokolova E, Oerlemans AM, Rommelse NN, Groot P, Hartman CA, Glennon JC, et al. A causal and mediation analysis of the comorbidity between attention deficit hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). J Autism Dev Disord. 2017;47(6):1595–604. https://doi.org/10.1007/s10803-017-3083-7.
65. Hollander E, Soorya L, Chaplin W, Anagnostou E, Taylor BP, Ferretti CJ, et al. A double-blind placebo-controlled trial of fluoxetine for repetitive behaviors and global severity in adult autism spectrum disorders. Am J Psychiatry. 2012;169(3):292–9. https://doi.org/10.1176/appi.ajp.2011.10050764.
66. McDougle CJ, Naylor ST, Cohen DJ, Volkmar FR, Heninger GR, Price LH. A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. Arch Gen Psychiatry. 1996;53(11):1001–8. https://doi.org/10.1001/archpsyc.1996.01830110037005.
67. Brandenburg C, Blatt GJ. Differential serotonin transporter (5-HTT) and 5-HT(2) receptor density in limbic and neocortical areas of adults and children with autism spectrum disorders: implications for selective serotonin reuptake inhibitor efficacy. J Neurochem. 2019;151(5):642–55. https://doi.org/10.1111/jnc.14832.
68. Pandina G, Ring RH, Bangerter A, Ness S. Current approaches to the pharmacologic treatment of core symptoms across the lifespan of autism spectrum disorder. Psychiatr Clin. 2020;43(4):629–45. https://doi.org/10.1016/j.psc.2020.08.003.
69. Hollander E, Soorya L, Wasserman S, Esposito K, Chaplin W, Anagnostou E. Divalproex sodium vs. placebo in the treatment of repetitive behaviours in autism spectrum disorder. Int J Neuropsychopharmacol. 2006;9(2):209–13.
70. Hollander E, Chaplin W, Soorya L, Wasserman S, Novotny S, Rusoff J, Feirsen N, Pepa L, Anagnostou E. Divalproex sodium vs placebo for the treatment of irritability in children and adolescents with autism spectrum disorders. Neuropsychopharmacology. 2010;35(4):990–8.
71. Hellings JA, Weckbaugh M, Nickel EJ, Cain SE, Zarcone JR, Reese RM, Hall S, Ermer DJ, Tsai LY, Schroeder SR, Cook EH. A Double-Blind, Placebo-Controlled Study of Valproate for Aggression in Youth with Pervasive Developmental Disorders. J Child Adolesc Psychopharmacol. 2005;15(4):682–92.
72. Wasserman S, Iyengar R, Chaplin WF, Watner D, Waldoks SE, Anagnostou E, et al. Levetiracetam versus placebo in childhood and adolescent autism: a double-blind placebo-controlled study. Int Clin Psychopharmacol. 2006;21(6):363–7. https://doi.org/10.1097/01.yic.0000224787.13782.0f.
73. Chen B, Choi H, Hirsch LJ, Katz A, Legge A, Buchsbaum R, et al. Psychiatric and behavioral side effects of antiepileptic drugs in adults with epilepsy. Epilepsy Behav. 2017;76:24–31. https://doi.org/10.1016/j.yebeh.2017.08.039.
74. McLay LL, France KG, Blampied NM, Hunter JE, van Deurs JR, Woodford EC, et al. Collateral child and parent effects of function-based behavioral interventions for sleep problems in children and adolescents with autism. J Autism Dev Disord. 2021. https://doi.org/10.1007/s10803-021-05116-3.
75. Hunter JE, McLay LK, France KG, Blampied NM. Systematic review of the collateral effects of behavioral sleep interventions in children and adolescents with autism spectrum disorder. Res Autism Spectrum Disord. 2020;79: 101677. https://doi.org/10.1016/j.rasd.2020.101677.
76. Lacivita E, Perrone R, Margari L, Leopoldo M. Targets for drug therapy for autism spectrum disorder: challenges and future directions. J Med Chem. 2017;60(22):9114–41. https://doi.org/10.1021/acs.jmedchem.7b00965.
77. Díaz-Caneja CM, State MW, Hagerman RJ, Jacquemont S, Marín O, Bagni C, et al. A white paper on a neurodevelopmental framework for drug discovery in autism and other neurodevelopmental disorders. Eur Neuropsychopharmacol. 2021;48:49–88. https://doi.org/10.1016/j.euroneuro.2021.02.020.
78. Baribeau D, Anagnostou E. Novel treatments for autism spectrum disorder based on genomics and systems biology. Pharmacol Ther. 2021;230:107939. https://doi.org/10.1016/j.pharmthera.2021.107939.
79. Persico AM, Ricciardello A, Lamberti M, Turriziani L, Cucinotta F, Brogna C, et al. The pediatric psychopharmacology of autism spectrum disorder: a systematic review—part I: the past and the present. Prog Neuropsychopharmacol Biol Psychiatry. 2021;110: 110326. https://doi.org/10.1016/j.pnpbp.2021.110326.
80. NCT02901431. A Study to investigate the efficacy and safety of Balovaptan (ro5285119) in Participants With Autism Spectrum Disorder (ASD). 2020.
81. Sikich L, Kolevzon A, King BH, McDougle CJ, Sanders KB, Kim SJ, et al. Intranasal oxytocin in children and adolescents with autism spectrum disorder. N Engl J Med. 2021;385(16):1462–73. https://doi.org/10.1056/NEJMoa2103583.
82. Freeman SM, Palumbo MC, Lawrence RH, Smith AL, Goodman MM, Bales KL. Effect of age and autism spectrum disorder on oxytocin receptor density in the human basal forebrain and midbrain. Transl Psychiatry. 2018;8(1):257. https://doi.org/10.1038/s41398-018-0315-3.
83. John S, Jaeggi AV. Oxytocin levels tend to be lower in autistic children: a meta-analysis of 31 studies. Autism. 2021;25(8):2152–61. https://doi.org/10.1177/13623613211034375.
84. Bolognani F, Del Valle RM, Squassante L, Wandel C, Derks M, Murtagh L, et al. A phase 2 clinical trial of a vasopressin V1a receptor antagonist shows improved adaptive behaviors in men with autism spectrum disorder. Sci Transl Med. 2019;11(491):eaat7838. https://doi.org/10.1126/scitranslmed.aat7838.
85. NCT03504917. A study of balovaptan in adults with autism spectrum disorder with a 2-year open-label extension. 2020.
86. Parker KJ, Oztan O, Libove RA, Mohsin N, Karhson DS, Sumiyoshi RD, et al. A randomized placebo-controlled pilot trial shows that intranasal vasopressin improves social deficits in children with autism. Sci Transl Med. 2019;11(491):eaau356. https://doi.org/10.1126/scitranslmed.aau7356.
87. Crutel V, Lambert E, Penelaud PF, Albarrán Severo C, Fuentes J, Rosier A, et al. Bumetanide oral liquid formulation for the treatment of children and adolescents with autism spectrum disorder: design of two phase III studies (SIGN trials). J Autism Dev Disord. 2020. https://doi.org/10.1007/s10803-020-04709-8.
88. Neurochlore S. Servier and Neurochlore announce the main results of the two phase 3 clinical studies assessing bumetanide in the treatment of autism spectrum disorders in children and adolescents. https://servier.com/en/communique/servier-and-neurochlore-announce-the-main-results-of-the-two-phase-3-clinical-studies-assessing-bumetanide-in-the-treatment-of-autism-spectrum-disorders-in-children-and-adolescents/. https://servier.com/en/communique/servier-and-neurochlore-announce-the-main-results-of-the-two-phase-3-clinical-studies-assessing-bumetanide-in-the-treatment-of-autism-spectrum-disorders-in-children-and-adolescents/ (2021). Accessed 07.09.2021.
89. Aran A, Harel M, Cassuto H, Polyansky L, Schnapp A, Wattad N, et al. Cannabinoid treatment for autism: a proof-of-concept randomized trial. Mol Autism. 2021;12(1):6. https://doi.org/10.1186/s13229-021-00420-2.
90. Campbell M, Anderson LT, Small AM, Adams P, Gonzalez NM, Ernst M. Naltrexone in autistic children: behavioral symptoms and attentional learning. J Am Acad Child Adolesc Psychiatry. 1993;32(6):1283–91. https://doi.org/10.1097/00004583-199311000-00024.
91. Anagnostou E. NCT03887676 Arbaclofen vs. placebo in the treatment of children and adolescents with ASD (ARBA) (2019).
92. Parellada M, San José Cáceres A, Palmer M, Delorme R, Jones EJH, Parr JR, et al. A phase II randomized, double-blind, placebo-controlled study of the efficacy, safety, and tolerability of arbaclofen administered for the treatment of social function in children and adolescents with autism spectrum disorders: study protocol for AIMS-2-TRIALS-CT1. Front Psychiatry. 2021;12:701729. https://doi.org/10.3389/fpsyt.2021.701729.
93. Gagan J. NCT03553875. Memantine for the treatment of social deficits in youth with disorders of impaired social interactions. 2018.
94. GW Research Ltd. NCT04745026. Trial to investigate the safety and efficacy of cannabidiol oral solution (GWP42003-P; CBD-OS) in children and adolescents with autism spectrum disorder. 2021.
95. Hollander E. NCT03202303. Cannabidivarin (CBDV) vs. placebo in children with autism spectrum disorder (ASD). 2017.
96. Anagnostou E, Bennett TA, Thorpe K, Nicolson R. 5.16 A phase 2 randomized, placebo-controlled trial of Tideglusib, an orally administered GSK-3 beta inhibitor, in the treatment of adolescents with ASD. J Am Acad Child Adolesc Psychiatry. 2018;57(10):S232. https://doi.org/10.1016/j.jaac.2018.09.311.

97. Anagnostou E. NCT01661855. A pilot study of riluzole versus placebo in the treatment of children and adolescents with ASD (RILISE). 2017.
98. Rothman J. L1-79 and the role of catecholamines in autism. Autism spectrum disorder-profile, heterogeneity, neurobiology and intervention. 2020.
99. Frye R. NCT02839915. Leucovorin for the treatment of language impairment in children with autism spectrum disorder. 2016.
100. Novotny S. NCT01260961. Developing treatment, treatment validation and treatment scope in the setting of an autism clinical trial. 2010.
101. Zimmerman AW, Singh K, Connors SL, Liu H, Panjwani AA, Lee LC, et al. Randomized controlled trial of sulforaphane and metabolite discovery in children with autism spectrum disorder. Mol Autism. 2021;12(1):38. https://doi.org/10.1186/s13229-021-00447-5.
102. Singh K, Connors SL, Macklin EA, Smith KD, Fahey JW, Talalay P, et al. Sulforaphane treatment of autism spectrum disorder (ASD). Proc Natl Acad Sci. 2014;111(43):15550. https://doi.org/10.1073/pnas.1416940111.
103. NCT02909959. Sulforaphane for the treatment of young men with autism spectrum disorder. 2020.
104. Smith R, Ou J, Jin H, Wu R, Fahey JW, Arriaza J, et al. M22. Sulforaphane as a treatment for autism: a randomized double-blind study. ACNP 59th annual meeting: poster session I. Neuropsychopharmacology. 2020;45(1):79–80. https://doi.org/10.1038/s41386-020-00890-7
105. Castejon AM, Spaw JA, Rozenfeld I, Sheinberg N, Kabot S, Shaw A, et al. Improving antioxidant capacity in children with autism: a randomized, double-blind controlled study with cysteine-rich whey protein. Front Psychiatry. 2021;12: 669089. https://doi.org/10.3389/fpsyt.2021.669089.
106. Hendren RL, James SJ, Widjaja F, Lawton B, Rosenblatt A, Bent S. Randomized, placebo-controlled trial of methyl B12 for children with autism. J Child Adolesc Psychopharmacol. 2016;26(9):774–83. https://doi.org/10.1089/cap.2015.0159.
107. Rossignol DA, Frye RE. The effectiveness of cobalamin (B12) treatment for autism spectrum disorder: a systematic review and meta-analysis. J Pers Med. 2021;11(8):784. https://doi.org/10.3390/jpm11080784.
108. Debray TP, Moons KG, van Valkenhoef G, Efthimiou O, Hummel N, Groenwold RH, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods. 2015;6(4):293–309. https://doi.org/10.1002/jrsm.1160.
109. Chugani DC, Chugani HT, Wiznitzer M, Parikh S, Evans PA, Hansen RL, et al. Efficacy of low-dose buspirone for restricted and repetitive behavior in young children with autism spectrum disorder: a randomized trial. J Pediatr. 2016;170:45-53.e1-4. https://doi.org/10.1016/j.jpeds.2015.11.033.
110. Autistica. Top ten questions for autism research. 2016.
111. Anixt JS, Murray DS, Coury DL, Kuhlthau KA, Eskra D, Seide J, et al. Improving behavior challenges and quality of life in the autism learning health network. Pediatrics. 2020;145(Suppl 1):S20–9. https://doi.org/10.1542/peds.2019-1895E.
112. Solmi M, Fornaro M, Ostinelli EG, Zangani C, Croatto G, Monaco F, et al. Safety of 80 antidepressants, antipsychotics, anti-attention-deficit/hyperactivity medications and mood stabilizers in children and adolescents with psychiatric disorders: a large scale systematic meta-review of 78 adverse effects. World Psychiatry. 2020;19(2):214–32. https://doi.org/10.1002/wps.20765.
113. Lemonnier E, Villeneuve N, Sonie S, Serret S, Rosier A, Roue M, et al. Effects of bumetanide on neurobehavioral function in children and adolescents with autism spectrum disorders. Transl Psychiatry. 2017;7(5): e1124. https://doi.org/10.1038/tp.2017.101.
114. Kloosterboer SM, de Winter BCM, Reichart CG, Kouijzer MEJ, de Kroon MMJ, van Daalen E, et al. Risperidone plasma concentrations are associated with side effects and effectiveness in children and adolescents with autism spectrum disorder. Br J Clin Pharmacol. 2020. https://doi.org/10.1111/bcp.14465.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC

Publication II

Meta-analysis of placebo-effects in the core symptoms of autism spectrum disorder

Citation: Siafis, S., Çıray, O., Schneider-Thoma, J., Bighelli, I., Krause, M., Rodolico, A., Ceraso, A., Deste, G., Huhn, M., Fraguas, D., Mavridis, D., Charman, T., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2020. 'Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis', *Mol Autism*, 11: 66. doi: 10.1186/s13229-020-00372-z.

This article is licensed under a Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/

Siafis et al. Molecular Autism (2020) 11:66
https://doi.org/10.1186/s13229-020-00372-z

Molecular Autism

RESEARCH Open Access

Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): systematic review and meta-regression analysis

Spyridon Siafis[1*], Oğulcan Çıray[2], Johannes Schneider-Thoma[1], Irene Bighelli[1], Marc Krause[1], Alessandro Rodolico[3], Anna Ceraso[4], Giacomo Deste[4], Maximilian Huhn[1,5], David Fraguas[6,7], Dimitris Mavridis[8,9], Tony Charman[10], Declan G. Murphy[11], Mara Parellada[8,9], Celso Arango[8,9] and Stefan Leucht[1]

Abstract

Background: Placebo response in autism spectrum disorder (ASD) might dilute drug-placebo differences and hinder drug development. Therefore, this meta-analysis investigated placebo response in core symptoms.

Methods: We searched ClinicalTrials.gov, CENTRAL, EMBASE, MEDLINE, PsycINFO, WHO-ICTRP (up to July 8, 2018), and PubMed (up to July 4, 2019) for randomized pharmacological and dietary supplement placebo-controlled trials (RCTs) with a minimum of seven days of treatment. Single-group meta-analyses were conducted using a random-effects model. Standardized mean changes (SMC) of core symptoms in placebo arms were the primary outcomes and placebo positive response rates were a secondary outcome. Predictors of placebo response were investigated with meta-regression analyses. The protocol was registered with PROSPERO ID CRD42019125317.

Results: Eighty-six RCTs with 2360 participants on placebo were included in our analysis (87% in children/adolescents). The majority of trials were small, single-center with a duration of 8–12 weeks and published after 2009. Placebo response in social-communication difficulties was SMC = − 0.32, 95% CI [− 0.39, − 0.25], in repetitive behaviors − 0.23[− 0.32, − 0.15] and in scales measuring overall core symptoms − 0.36 [− 0.46, − 0.26]. Overall, 19%, 95% CI [16–22%] of participants were at least much improved with placebo. Caregiver (vs. clinician) ratings, lower risk of bias, flexible-dosing, larger sample sizes and number of sites, less recent publication year, baseline levels of irritability, and the use of a threshold of core symptoms at inclusion were associated with larger placebo response in at least a core symptom domain.

Limitations: About 40% of the trials had an apparent focus on core symptoms. Investigation of the differential impact of predictors on placebo and drug response was impeded by the use of diverse experimental interventions with essentially different mechanisms of action. An individual-participant-data meta-analysis could allow for a more fine-grained analysis and provide more informative answers.

(Continued on next page)

* Correspondence: spyridon.siafis@tum.de
[1]Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Ismaningerstr. 22, 81675 Munich, Germany
Full list of author information is available at the end of the article

© The Author(s). 2020, corrected publication 2020. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)
Conclusions: Placebo response in ASD was substantial and predicted by design- and participant-related factors, which could inform the design of future trials in order to improve the detection of efficacy in core symptoms. Potential solutions could be the minimization and careful selection of study sites as well as rigorous participant enrollment and the use of measurements of change not solely dependent on caregivers.

Keywords: Autism spectrum disorder, Placebo, Trials

Background

Autism spectrum disorder (ASD) is a group of heterogeneous neurodevelopmental conditions, characterized by social-communication difficulties as well as repetitive-restricted behaviors and sensory abnormalities [1]. The prevalence is about 1–2% [2, 3], and lifetime costs are substantial (at US $1.4–2.44 million per individual) [4]. Behavioral interventions are the cornerstone of treatment and there is still no approved medication for the core symptoms [5]. Despite that, about half of the individuals with ASD, who might be more susceptible to side effects than neurotypical populations [5], receive psychotropic drugs [6]. Currently approved medications target associated symptoms, e.g., aripiprazole and risperidone for irritability [5]. Therefore, there is an unmet need to develop effective and safe treatments that target causal pathophysiological pathways, improve core symptoms and quality of life.

In spite of the recent advances in "translational" research, late-stage clinical trials for neurodevelopmental disorders have failed [7]. The low success rate could be explained by several factors, such as poor translational validity of preclinical models, true lack of drug efficacy, and suboptimal trial design [8]. One concern is also that placebo effects might dilute effect sizes. However, the magnitude and predictors of placebo response in core symptoms of ASD are still unknown; only investigated in post-hoc analyses of single trials [9, 10] and meta-analyses using aggregated outcome measures, potentially confounded by associated symptoms [11, 12]. In summary, placebo response may play an important role in the failure of clinical trials and the subsequent lack of approved medications for core symptoms. In order to improve the design and sensitivity of future trials, we meta-analyzed placebo response of core symptoms in pharmacological and dietary supplement ASD trials.

Methods

This systematic review and meta-regression analysis was conducted according to PRISMA [13] (Additional file 3: eAppendix-1) with PROSPERO registration ID CRD42019125317 (Additional file 3: eAppendix-2).

Participants and interventions

Participants

Participants with a diagnosis of ASD using standardized diagnostic criteria (e.g., DSM-III, ICD-10, or more recent versions) and/or validated diagnostic tools (e.g., ADI-R) [5]. There were no restrictions in terms of age, sex, ethnicity, setting, severity, or the presence of co-occurring conditions.

Interventions

Any pharmacological treatment or dietary supplement was eligible, when compared with placebo. We excluded psychological/behavioral and combination interventions (since placebo response might be confounded by the active component of the combination) as well as other interventions (e.g., elimination diets, milk formulations, or homeopathy). The minimum duration of treatment was 7 days, since we aimed to investigate a broad range of data but to exclude trials with a clearly very short duration, e.g., single-dose interventions. There was no restriction in terms of route of administration and dosing-schedule.

Type of studies

Blinded and unblinded randomized placebo-controlled trials (RCTs) were eligible. In case of cross-over studies, we used only data from the first phase of the crossover to avoid carry-over effects [14]. We excluded studies with placebo-controlled discontinuation or cluster randomization [15], published before 1980 or smaller than ten participants [16]. Risk of bias of included studies was evaluated by at least two independent reviewers (SS, OC, AR) using the Cochrane Collaboration risk-of-bias tool [17]. Disagreements were resolved by discussion, and if needed, a third author was involved (SL, JST). Studies with a high risk of bias in sequence generation or allocation concealment were excluded (e.g., allocation by alternation or by an unblinded investigator). Studies were further classified as having an overall low, moderate, or high risk of bias [18].

Search strategy and selection criteria

We searched (July 8, 2018) ClinicalTrials.gov, CENTRAL, EMBASE, MEDLINE, PsycINFO, PubMed (update on July 4, 2019), and WHO ICTRP. There was no date/time, language, document type, and publication status limitations (Additional file 3: eAppendix-3). Reference lists of included studies and previous reviews [5, 11, 12, 19–27] were also inspected.

Outcome measures and data extraction

We investigated placebo response in core symptoms. The following primary outcomes, as measured by published scales, were analyzed: (1) social-communication difficulties (e.g., ABC-L/SW [28] or VABS-Socialization [29]), (2) repetitive-restricted behaviors (e.g., ABC-S [28] or CYBOC-PDD [30]), and (3) overall measures of core symptoms (e.g., SRS [31] or CARS [32]). There is no agreement on the optimal outcome measures to use in clinical trials of ASD and so preference was given to the aforementioned most frequently used scales (Additional file 3: eAppendix-5.3) [5, 33–36]. A higher score indicated more difficulties and when necessary, scores were minus-transformed. In the primary analysis, we pooled all studies by preferring ratings by clinicians (observations or interviews) to caregivers/teachers. Separate analyses by type of raters and positive response to treatment defined as at least much improvement in CGI-I, preferably anchored to global autism or core symptoms (when more than one CGI-I evaluations were reported), were analyzed as secondary outcomes. When the number of participants with a positive response was not reported, it was imputed from mean and standard deviation (SD) of CGI-I using a validated method (Additional file 3: eAppendix-2.2) [37, 38].

At least two independent reviewers/contributors selected relevant records and extracted data from eligible studies in an Access database (SS, OC, IB, AR, AC, GD, MK, YZ, and TF). Intention-to-treat data were preferred when available, and for a positive response to treatment, if the original authors presented only the results of completer population, we assumed that participants lost to follow-up did not have a positive response to treatment. Missing SDs were calculated according to the following hierarchy from available statistics (e.g., SE, p values, t tests) [39], median/range [40], pooling subscales (e.g., SRS subscales, assuming a correlation of 0.5) [41], or using a validated imputation method [39, 42]. Corresponding authors were contacted by e-mail for additional data, with a reminder e-mail in case of no response (complete list in Additional file 3: eAppendix-4).

Statistical analysis

Synthesis of the results

Single-group meta-analyses of placebo arms were conducted using a random effects model [43]. The effect size for continuous outcomes (core symptoms) was the standardized mean change (SMC) with raw score standardization using the baseline SD of the placebo arm [44, 45]. When baseline SDs were not reported, change or follow-up SDs were used. In the primary analysis, a common pre-post correlation of 0.5 [41] was used for the calculation of variance of SMC [44]. Positive response rates were logit transformed, and back-transformed for presentation [46]. Heterogeneity was evaluated by visual inspection of forest plots and with the χ^2 (p value < 0.1) and I^2 statistics (considerable heterogeneity when > 50%); χ^2 might detect small amounts of clinically unimportant heterogeneity; therefore, we based our evaluation on I^2 [17].

Sensitivity analyses and publication bias

Predefined sensitivity analyses of the primary outcomes were conducted using a fixed-effects model or by exclusion of studies with genetic syndrome as inclusion criteria, using only diagnostic tools, single-blind, shorter than 4 weeks, presenting only completers data, with at least moderate overall risk of bias, with estimated SD (imputed, from medians/range, or pooled subscales). Post-hoc, we excluded studies without baseline SDs and we used the correlations of 0.25/0.75 for the calculation of variance of SMC [41]. Regarding responder rates, we post-hoc excluded studies with imputed responder rates [38]. We explored small study effects as proxy for publication bias with contour-enhanced funnel plots, Egger's test [47], and trim-and-fill [48].

Meta-regression analyses

The dependent variable was SMC and the independent variable was selected from a list of covariates from the literature [9, 11, 12, 49–51]. First, we conducted univariable and then multivariable meta-regressions similar to our previous analyses in schizophrenia [51]: we used the factors that were significant in the univariable analysis and then a formal backward stepwise algorithm with a removal criterion of $p = 0.15$. Meta-regressions were not performed for categorical covariates with less than five data points per level. Spearman's ρ were calculated post-hoc between SMCs of placebo and experimental intervention as well as between covariates.

Intervention-related factors Intervention-related factors were route of administration (oral versus others) [52], type of experimental intervention (pharmacological versus dietary supplement), and dose-schedule (fixed versus flexible).

Study-related factors Study-related factors were duration of treatment (weeks), publication year, washout from psychotropic medications (coded post-hoc as the presence of washout or not, because definitions varied), placebo lead-in with exclusion of those showing a positive response, type of rater (clinicians versus caregivers), total sample size, number of sites, %academic sites, number of arms and medications, %participants on placebo, sponsorship (industry-funded/patent application

versus industry-independent), country of origin (US versus not only US), and risk of bias domains.

Participant-related factors Participant-related factors were the presence of any associated conditions by inclusion criteria (i.e., irritability, ADHD, and other conditions apart from intellectual disability or genetic syndrome), mean age and age group (children/adolescent versus adults/mixed, post-hoc), %participants with intellectual disability (at least mild or IQ < 70), %female (post-hoc), ethnicity (%Caucasian/Hispanic, post-hoc), and baseline BMI (post-hoc) [9]. Due to inconsistent reporting of baseline severity [11, 12], we used CGI-Severity (ranging 1–7) as a measure of global severity and ABC-Irritability (ranging 0–45) as a measure of serious behavioral problems [53]. Baseline severity in core symptoms could not be investigated as a potential predictor due to the large diversity of scales and standardization methods (such as using the lower and upper limits of the measurement scale [54]) could not be utilized (trials reported raw and standard scores such as of VABS or T-scores of SRS). We also examined the use of a threshold of core symptom severity for inclusion (not only for the confirmation of diagnosis).

Analyses were conducted using metafor (v2.1-0) [45] and meta (v4.9-9) [55] in R (v3.6.2) [56]. Statistical threshold was set at two-sided alpha 5%. Due to the limited statistical power and exploratory (observational) nature of meta-regression analyses, alpha was not adjusted for multiple testing. Correction for multiple testing is not generally recommended by the Cochrane Handbook [39].

Results

Description of included studies

The PRISMA flow diagram is presented in Fig. 1. In this analysis, 86 (k) studies were included, 71% comparing pharmacological treatments and 29% dietary supplements with placebo (eAppendix-5.1 and Table-S). Of the 86 studies, 75 were conducted in children/adolescents, eight in adults, and three included both age groups. The overall sample size (n) was 5365, 44% on placebo. The majority of studies were parallel (85%), single-center (60%, indicated in k = 78), and double-blind (only one single-blind [57] and none was open) with two arms (88%) and small sample sizes (median 45, interquartile-range [30–91]). About half of the studies (48%) had a duration of 12 weeks or more and three less than 4 weeks [57–59] (median 10 weeks [8–12]) as well as half used a fixed dose schedule (51%, k = 84) and had a washout from psychotropic drugs (55%, k = 75), yet definitions and duration varied. Placebo lead-in with exclusion of those with a positive response was used in five studies.

All of the studies used standardized diagnostic criteria, except five that used only diagnostic tools [60–64]. Associated conditions were the focus of and they were

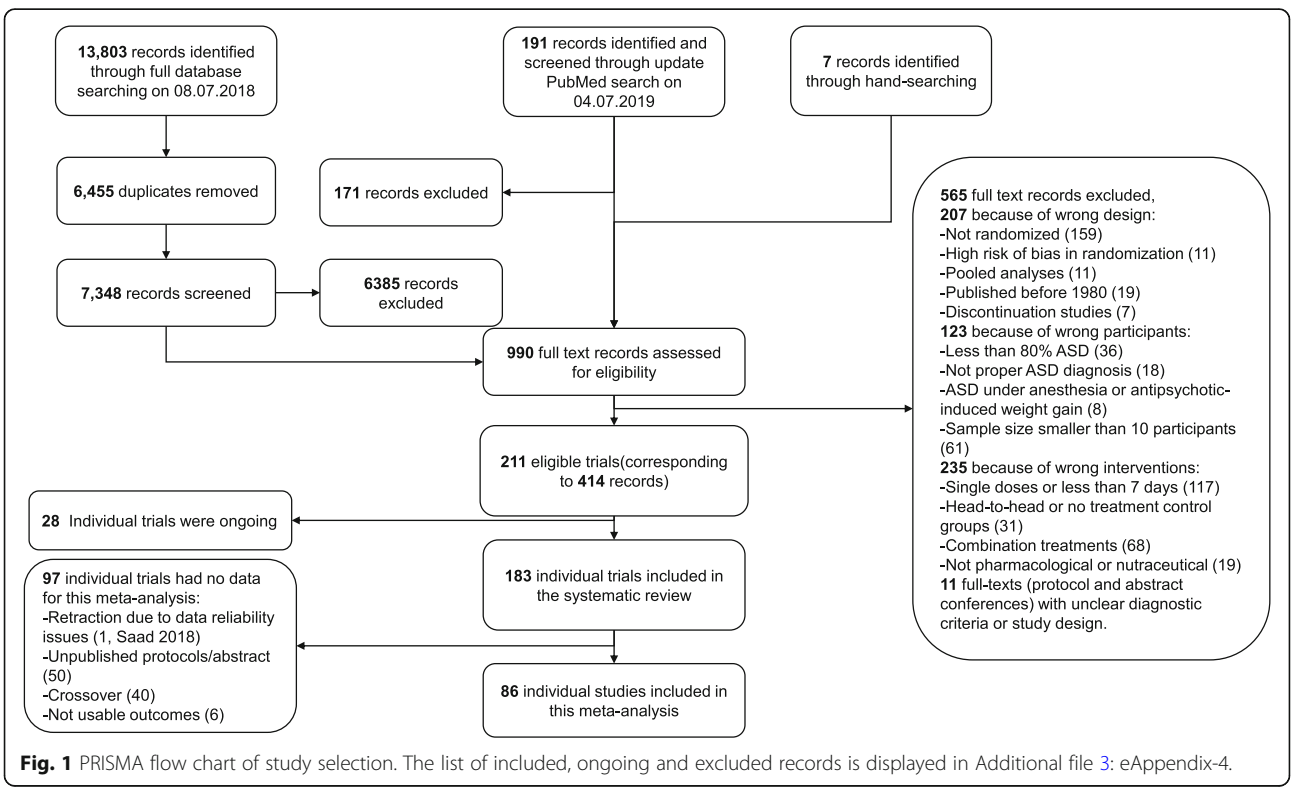


Fig. 1 PRISMA flow chart of study selection. The list of included, ongoing and excluded records is displayed in Additional file 3: eAppendix-4.

required for inclusion in 29 studies (irritability in 52% of them), and a genetic syndrome in one (neurofibromatosis-type-1) [60]. Core symptoms was the primary focus in 34 trials, while the focus was unclear in 23 studies. Nine studies included participants using a threshold of core symptom severity, using ABC-L/SW, CARS, RLRS, SRS, and YBOCS-versions (in five out of eight). Participants on placebo had a median age of 8.63 years ([6.55–10.16], k = 82), 17.16% were female ([0–20.9%], k = 80), and 54.2% had comorbid intellectual disability ([0–75%], k = 30). The median of baseline CGI-Severity was 4.73 ([4.39–5.00], k = 38) and ABC-Irritability was 17.18 ([13.71–22.70], k = 36).

Overall, 40% of the studies had an overall low risk of bias, 52% moderate and 8% high. Description of the methods was adequately reported in more than half of the studies for sequence generation (63%), allocation concealment (54%), and blinding (72%). Missing outcomes were adequately addressed in 60%, with a median overall dropout rate from placebo of 12.93% ([6.25–22.6%] and k = 70 trials out of 86 reported attrition rates). Of the studies, 23% had a high risk of selective reporting, and 13% high risk in other biases, mainly due to imbalances between groups (Additional file 3: eAppendix-5.2). Finally, 38% of the studies were industry-sponsored (including five in which investigators applied for a patent on the experimental intervention), and sponsorship was unclear in three studies.

Primary outcomes

Social-communication difficulties

Primary analysis In the primary analysis, 52 studies with 1497 participants on placebo were included. Most of the scales were filled by caregivers (77%). ABC-L/SW was the most used scale (56%) followed by VABS-Socialization (13%). Pooled placebo response was SMC = − 0.32 [95%CI − 0.39, − 0.25], with moderate levels of heterogeneity (I^2 = 31.88%, χ^2 = 74.87, p = 0.02) (Fig. 2).

Sensitivity analysis and publication bias The results did not change materially in sensitivity analyses (Additional file 3: eAppendix-6.1). There was no indication of small-study effects from a visual inspection of the funnel plot and Egger's test or publication bias (z = − 0.38, p = 0.70) (Additional file 3: eAppendix-6.2). Also, fixed and random effects summaries were identical, an indication that smaller and larger studies give similar results.

Meta-regressions The results of the univariable and multivariable meta-regression analyses are presented in Table 1, Figure-S, and eAppendix-6.3. Placebo response in social-communication difficulties decreased over years (by 0.016 [0.003, 0.030] SMC units per publication year). Larger placebo response was associated with caregiver ratings (−0.164 [−0.315, −0.012]), low risk of bias in other bias (− 0.160 [− 0.299, − 0.021]), and higher baseline ABC-Irritability (− 0.017 [− 0.028, − 0.006] per point). In the multivariable meta-regression, using the backward selection procedure, other bias, baseline ABC-Irritability, and the type of rater remained as covariates in the model, but the latter two were not significant due to their interaction with the other covariates. In a model without ABC-Irritability (available in 31 studies), publication year, other bias, and type of rater remained, the latter was not significant.

Repetitive behaviors

Primary analysis Fifty-two studies were included in the primary analysis with 1492 participants. Caregivers filled about half of the scales (56%). The most frequently used scales were ABC-S (44%) and YBOCS-versions (33%). Overall placebo response was SMC = − 0.23 [− 0.32, − 0.15], and there was some heterogeneity (I^2 = 55%, χ^2 = 113.32, p < 0.001) (Fig. 3).

Sensitivity analysis and publication bias Sensitivity analyses did not change the results materially, though there was a small difference between fixed and random-effects summary estimates, indicating possible small-study effects. Egger's test was not significant and yielded a marginal p value (z = 1.71, p = 0.09); it has been suggested that for this test, a threshold of 0.1 should be employed. By visual inspection of the funnel plot, we detected a possible asymmetry (Additional file 3: eAppendix-6.2) and the trim-and-fill adjusted placebo response was − 0.33 [− 0.41, − 0.25].

Meta-regressions Higher placebo response was associated with larger sample sizes (− 0.02 [− 0.03, − 0.01] SMC units per ten participants), flexible-dosing (− 0.195 [− 0.351, − 0.038]), and the use of a threshold of core symptoms at inclusion (− 0.346 [− 0.516, − 0.175]). These covariates remained in the multivariable model, but the use of a threshold of core symptoms at inclusion was not significant. Nevertheless, the findings might have been driven by three antidepressant trials in children/adolescents [65–67], with larger sample sizes (~ 150) and multiple sites (3, 6, and 18), as well as using flexible-dosing and a threshold of CYBOCS-PDD for inclusion (Table 1).

Overall core symptoms

Primary analysis Forty-five studies with 1063 participants were included in the primary analysis. Caregivers filled about half of the scales (51%). The most frequently

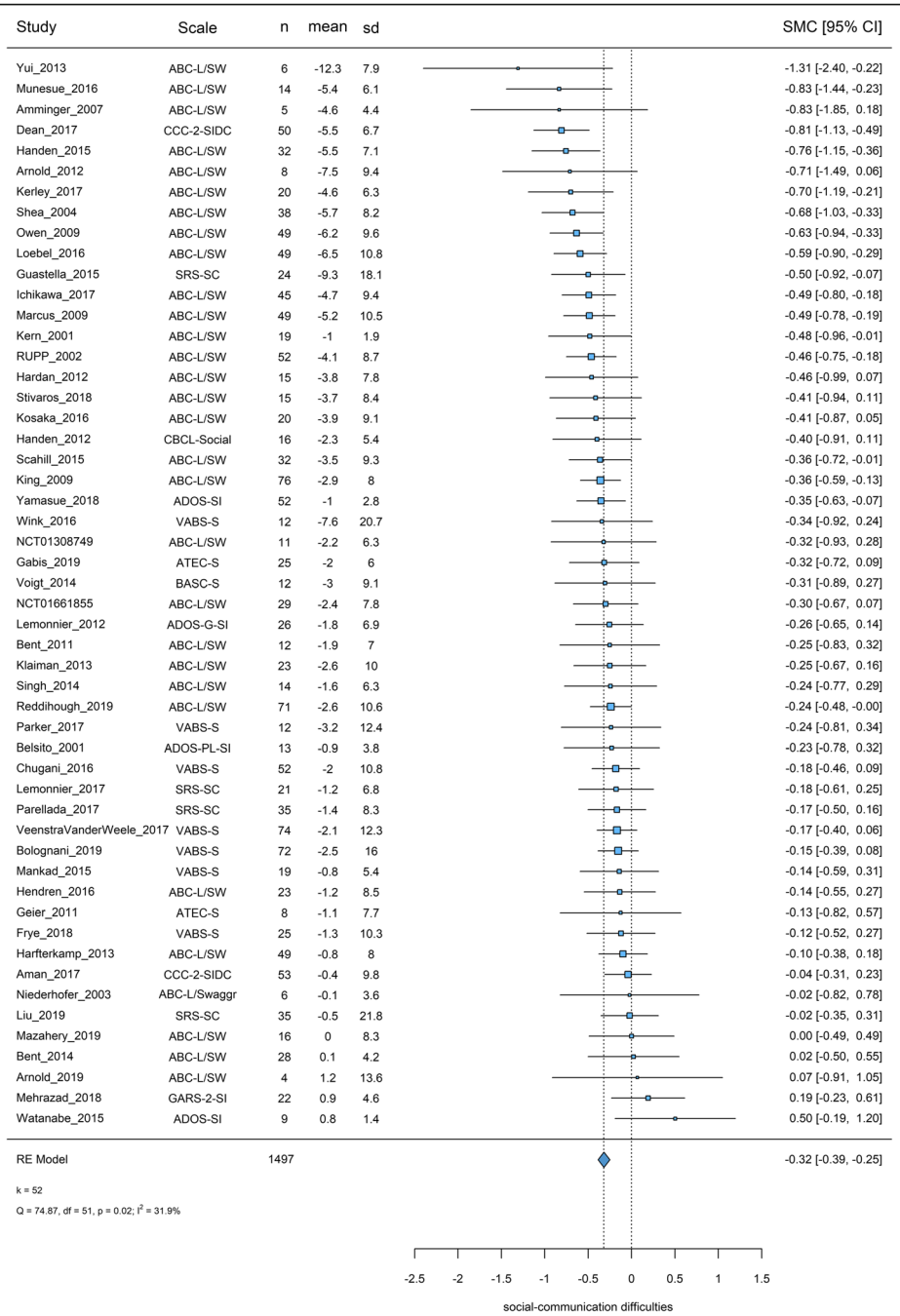


Fig. 2 Placebo response in scales measuring social-communication difficulties. Squares and bars represent standardized mean changes (SMC) and 95% confidence intervals for each study. The size of the square is proportional to the weight of the study in the meta-analysis. The diamond represent the pooled SMC. Heterogeneity is quantified with a χ^2 test (Q) and I^2. *In Chugani 2016, standard errors might have been reported as SDs. Therefore, we calculated SDs from the reported values (no reply from the corresponding author). It should be noted that in Niederhofer 2003, an aggregated score of ABC-L/SW rated by both caregivers and teachers were reported, in Amminger 2007, ABC-L/SW was rated by clinicians of the day care center. Scale: the scale used (clinician rated scales based on observation or interviews were preferred in the primary analysis); *n*: the number of participants on placebo; mean: mean change from baseline to endpoint (negative values for improvement); sd: the standard deviation used for the standardization (baseline standard deviations were preferred); SMC: standardized mean changes, 95% CI: 95% confidence intervals, *k* = total number of studies included in the analysis

Table 1 Univariable meta-regression analyses

Covariate		Social-communication difficulties				Repetitive behaviors and restricted interests				Overall core symptoms			
		k, *n*	β [95% CI]	*p*	R^2 (%)	*k*, *n*	β [95% CI]	*p*	R^2 (%)	*k*, *n*	β [95% CI]	*p*	R^2 (%)
Intervention	Route (ref. not oral)	52, 1497	0 [− 0.205, 0.206]	0.998	0	52, 1492	− 0.166 [− 0.391, 0.059]	0.148	8.13	45, 1063	− 0.042 [− 0.301, 0.217]	0.749	0
	Experimental intervention (ref. dietary supplement)	52, 1497	− 0.082 [− 0.229, 0.064]	0.268	2.61	52, 1492	− 0.081 − 0.262, 0.1]	0.382	0	45, 1063	− 0.023 [− 0.228, 0.181]	0.822	0
	Dose schedule (ref. fixed)**	51, 1491	− 0.095 [− 0.229, 0.038]	0.163	2.57	**51, 1486**	**− 0.195 [− 0.351, − 0.038]**	**0.015**	27.03	44, 1057	− 0.044 [− 0.276, 0.189]	0.713	0
Study design	Publication year	**52, 1497**	**0.016 [0.003, 0.03]**	**0.017**	**31.53**	52, 1492	− 0.006 [− 0.022, 0.009]	0.416	0	45, 1063	− 0.004 [− 0.021, 0.013]	0.607	0
	Country (ref. outside US/mixed)	52, 1497	− 0.002 [− 0.139, 0.135]	0.973	0	52, 1492	− 0.024 [− 0.194, 0.145]	0.778	0	45, 1063	− 0.054 [− 0.257, 0.148]	0.598	0.03
	Sponsorship (ref. no)	50, 1469	0.078 [− 0.054, 0.21]	0.246	2.89	51, 1470	0.002 [− 0.172, 0.175]	0.984	0	43, 996	0.057 [− 0.152, 0.266]	0.590	0
	No. sites*	48, 1395	− 0.003 [− 0.008, 0.003]	0.295	0	48, 1427	− 0.004 [− 0.011, 0.003]	0.287	0	**41, 957**	**− 0.025 [− 0.045, − 0.005]**	**0.015**	**31.62**
	% academic sites	46, 1340	− 0.196 [− 0.435, 0.044]	0.11	10.66	47, 1383	− 0.08 [− 0.397, 0.236]	0.619	0	41, 957	− 0.117 [− 0.461, 0.227]	0.505	0
	No. arms	52, 1497	− − 0.036 [− 0.208, 0.136]	0.684	0	52, 1492	0.035 [− 0.182, 0.252]	0.753	0	Insufficient data			
	Duration (weeks)	52, 1497	0.002 [− 0.007, 0.012]	0.669	0	52, 1492	− 0.009 [− 0.021, 0.003]	0.124	5.63	45, 1063	0.00 3[− 0.009, 0.015]	0.677	0
	Washout (ref. no)	46, 1404	− 0.064 [− 0.204, 0.076]	0.369	0	47, 1339	− 0.006 [− 0.165, 0.153]	0.943	0	39, 995	− 0.021 [− 0.248, 0.205]	0.855	0
	Sample size**	52, 1497	0 [− 0.001, 0.001]	0.967	0	**52, 1492**	**− 0.002 [− 0.003, − 0.001]**	**0.004**	**29.38**	45, 1063	− 0.001 [− 0.004, 0.001]	0.213	9.89
	% participants on placebo	52, 1497	0.172 [− 0.625, 0.968]	0.673	0	52, 1492	− 0.333 [− 1.31, 0.643]	0.503	0	45, 1063	0.401 [− 0.981, 1.783]	0.569	0
	Rater (ref. clinician)***	**51, 1491**	**− 0.164 [− 0.315, − 0.012]**	**0.034**	**21.48**	51, 1483	0.131 [− 0.033, 0.294]	0.117	12.20	43, 1009	− 0.148 [− 0.361, 0.065]	0.174	0
	Sequence generation (ref. unclear)	52, 1497	0.147 [− 0.047, 0.34]	0.138	6.46	52, 1492	− 0.043 [− 0.302, 0.216]	0.743	0	45, 1063	0.243[− 0.048, 0.533]	0.102	1.70
	Allocation concealment (ref. unclear)	52, 1497	− 0.045 [− 0.228, 0.138]	0.631	0	52, 1492	− 0.012 [− 0.233, 0.208]	0.912	0	**45, 1063**	**− 0.252 [− 0.485, − 0.019]**	**0.034**	**17.23**
	Blinding (ref. unclear/high)	52, 1497	0.083 [− 0.068, 0.235]	0.28	0	52, 1492	0.006 [− 0.191, 0.202]	0.953	0	45, 1063	− 0.019 [− 0.269, 0.231]	0.881	0
	Missing outcome (ref. unclear/high)	52, 1497	− 0.141 [− 0.293, 0.01]	0.067	12.03	52, 1492	− 0.052 [− 0.247, 0.143]	0.6	0	45, 1063	− 0.162 [− 0.362, 0.038]	0.111	12.72
	Selective reporting (ref. unclear/high)	52, 1497	− 0.014 [− 0.254, 0.227]	0.912	0	52, 1492	− 0.156 [− 0.471, 0.160]	0.333	0.33	Insufficient data			

Table 1 Univariable meta-regression analyses *(Continued)*

Covariate		Social-communication difficulties				Repetitive behaviors and restricted interests				Overall core symptoms			
		k, *n*	β [95% CI]	*p*	R^2 (%)	*k*, *n*	β [95% CI]	*p*	R^2 (%)	*k*, *n*	β [95% CI]	*p*	R^2 (%)
	Other bias (ref. unclear/high)	**52, 1497**	**− 0.160 [− 0.299, − 0.021]**	**0.024**	**27.38**	52, 1492	− 0.091 [− 0.303, 0.120]	0.398	0	45, 1063	− 0.033 [− 0.254, 0.188]	0.768	0
Participant	Age group (ref. adults/mixed)	52, 1497	− 0.057 [− 0.27, 0.155]	0.597	0	52, 1492	− 0.107 [− 0.323, 0.109]	0.33	1.55	45, 1063	0.148 [− 0.124, 0.42]	0.287	9.35
	Mean age	51, 1478	0.002 [− 0.011, 0.014]	0.809	0	51, 1486	0.003 [− 0.008, 0.014]	0.547	0	45, 1063	− 0.007 [− 0.021, 0.007]	0.297	7.37
	% female	50, 1453	0.131 [− 0.65, 0.911]	0.743	0	50, 1482	− 0.282 − 1.114, 0.55]	0.507	0	45, 1063	0.277 [− 0.650, 1.203]	0.559	0
	% intellectual disability	17, 620	− 0.328 [− 0.672, 0.017]	0.063	42.21	21, 672	0.128 [− 0.335, 0.59]	0.588	0	14, 316	0.382 [− 0.105, 0.869]	0.124	20.87
	% Caucasian or Hispanic	29, 963	0.036 [− 0.351, 0.423]	0.855	0	29, 971	− 0.157 [− 0.715, 0.402]	0.583	0	21, 545	− 0.42 [− 1.204, 0.364]	0.294	0
	Associated conditions at baseline (ref. no)	52, 1497	− 0.068 [− 0.212, 0.075]	0.351	3.45	52, 1492	0.111 [− 0.065, 0.287]	0.216	5.83	45, 1063	− 0.012 [− 0.242, 0.218]	0.919	0
	Baseline mean BMI	12, 445	− 0.063 [− 0.156, 0.029]	0.179	19.01	12, 461	− 0.02 [− 0.132, 0.091]	0.720	0	8, 254	− 0.033 [− 0.122, 0.057]	0.476	0
	Baseline mean CGI− S	26, 917	− 0.042 [− 0.25, 0.167]	0.694	0	26, 925	0.033 [− 0.227, 0.292]	0.805	0	17, 513	− 0.077 [− 0.442, 0.288]	0.678	0
	Baseline mean ABC− Irritability	**31, 917**	**− 0.017 [-0.028, − 0.006]**	**0.002**	**100**	30, 884	− 0.005 [− 0.022, 0.013]	0.608	0	17, 340	− 0.006 [− 0.035, 0.023]	0.68	0
	Minimum threshold of core symptoms for inclusion (ref. no)**	52, 1497	0.085 [− 0.094, 0.264]	0.358	0.88	**52, 1492**	**− 0.346 [− 0.516, − 0.175]**	**< 0.001**	**53.85**	Insufficient data			

A negative coefficient represent an increase of placebo response. For dichotomous covariates, the reference level is mentioned. Meta-regression with dichotomous covariates were not performed when there were less than five data points for a level of the covariate (e.g., placebo lead-in or number of medications). For continuous covariates, the covariate refer to a change of 1 point of the variable, e.g., per year for publication year: decrease of placebo response in social-communication difficulties by 0.016 per year, % percentage of intellectual disability: increase of placebo response in social-communication difficulties by − 0.328 from 0 to 1 (100%) of the participants had intellectual disability. *The effect of number of sites on placebo response in overall core symptoms was not significant ($k = 40$, coefficient 0.0182 [− 0.268, 0.0631], $p = 0.4287$), when one outlier study with 26 sites was excluded (Bolognani 2019). **Placebo response in repetitive behaviors was not found to be predicted by sample size ($k = 49$, coefficient − 0.001 [− 0.002, 0.000], $p = 0.052$), flexible-dosing ($k = 48$, coefficient − 0.073 [− 0.189, 0.043], $p = 0.218$) and using a minimum threshold of core symptoms ($k = 49$, coefficient − 0.083 [− 0.24, 0.073], $p = 0.2950$), when three antidepressant trials (Herscu 2019, King 2009, and Reddihough 2019) were excluded. ***In the meta-regression of type of rater, Niederhofer 2003 was not included in social-communication difficulties (aggregated caregiver/teacher rating of ABC-L/SW), Saad 2015 was not included in overall core symptoms (it was not clear if CARS was rated only by parents or also filled by clinicians), and Anagnostou_2012 was not included in overall core symptoms and repetitive behaviors (SRS and RBS-R might have been rated as self-reports)

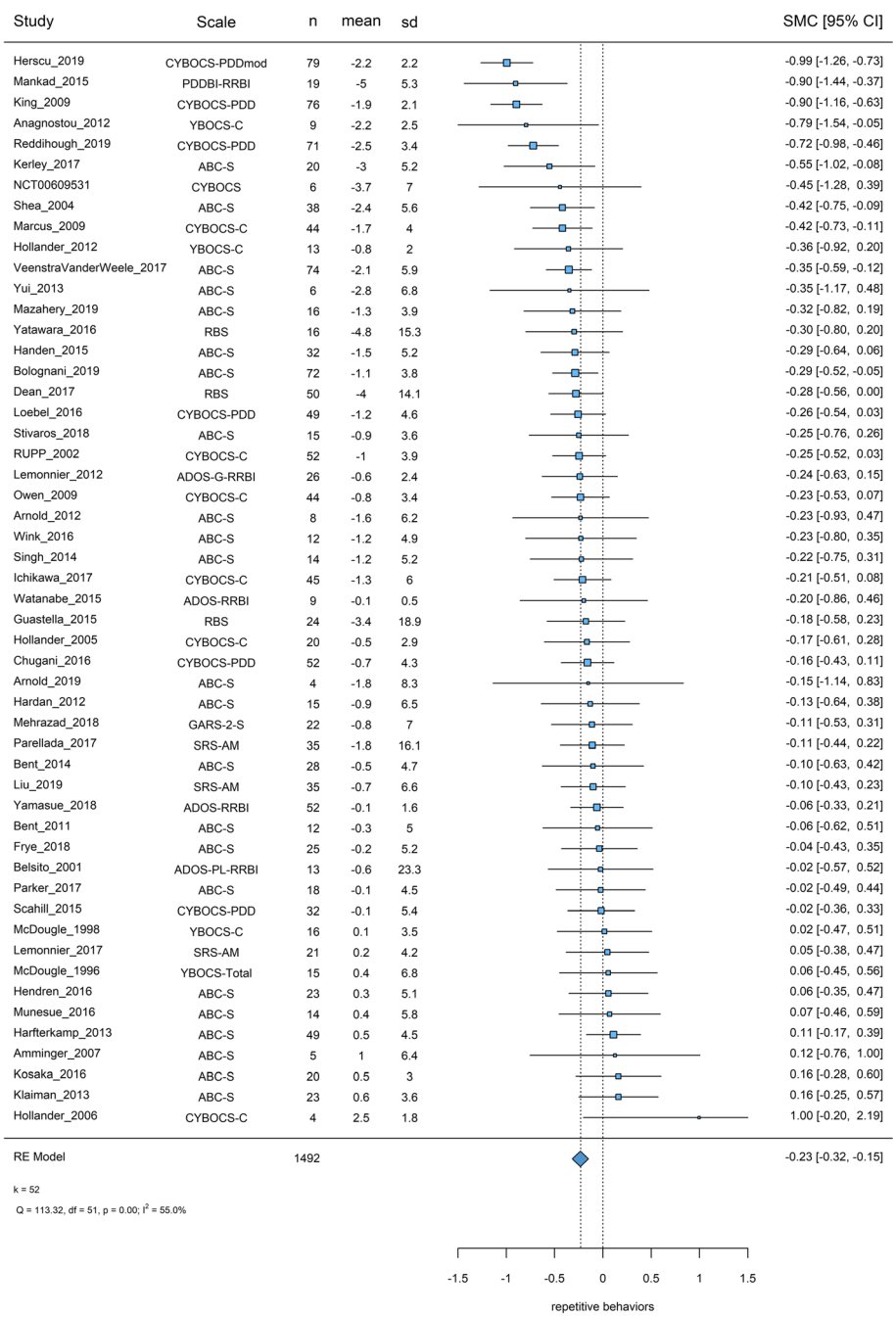


Fig. 3 Placebo response in scales measuring repetitive behaviors. Squares and bars represent standardized mean changes (SMC) and 95% confidence intervals for each study. The size of the square is proportional to the weight of the study in the meta-analysis. The diamond represent the pooled SMC. Heterogeneity is quantified with a χ^2 test (Q) and I^2. *In Chugani 2016, standard errors might have been reported as SDs. Therefore, we calculated SDs from the reported values (no reply from the corresponding author). In Amminger 2007, ABC-S was rated by clinicians of the day care center. Scale: the scale used (clinician rated scales based on observation or interviews were preferred in the primary analysis); n: the number of participants on placebo; mean: mean change from baseline to endpoint (negative values for improvement); sd: the standard deviation used for the standardization (baseline standard deviations were preferred); SMC: standardized mean changes, 95% CI: 95% confidence intervals, k = total number of studies included in the analysis

used scales were SRS (49%) and CARS (24%). Overall placebo response was SMC = − 0.36 [− 0.46, − 0.26] and heterogeneity was considerable (I^2 = 55.53%, χ^2 = 98.94, $p < 0.001$) (Fig. 4).

Sensitivity analysis and publication bias Sensitivity analyses did not change the results materially, no asymmetry was detected in the funnel plot, and Egger's test yielded a marginal p value (z = − 1.82, p = 0.07).

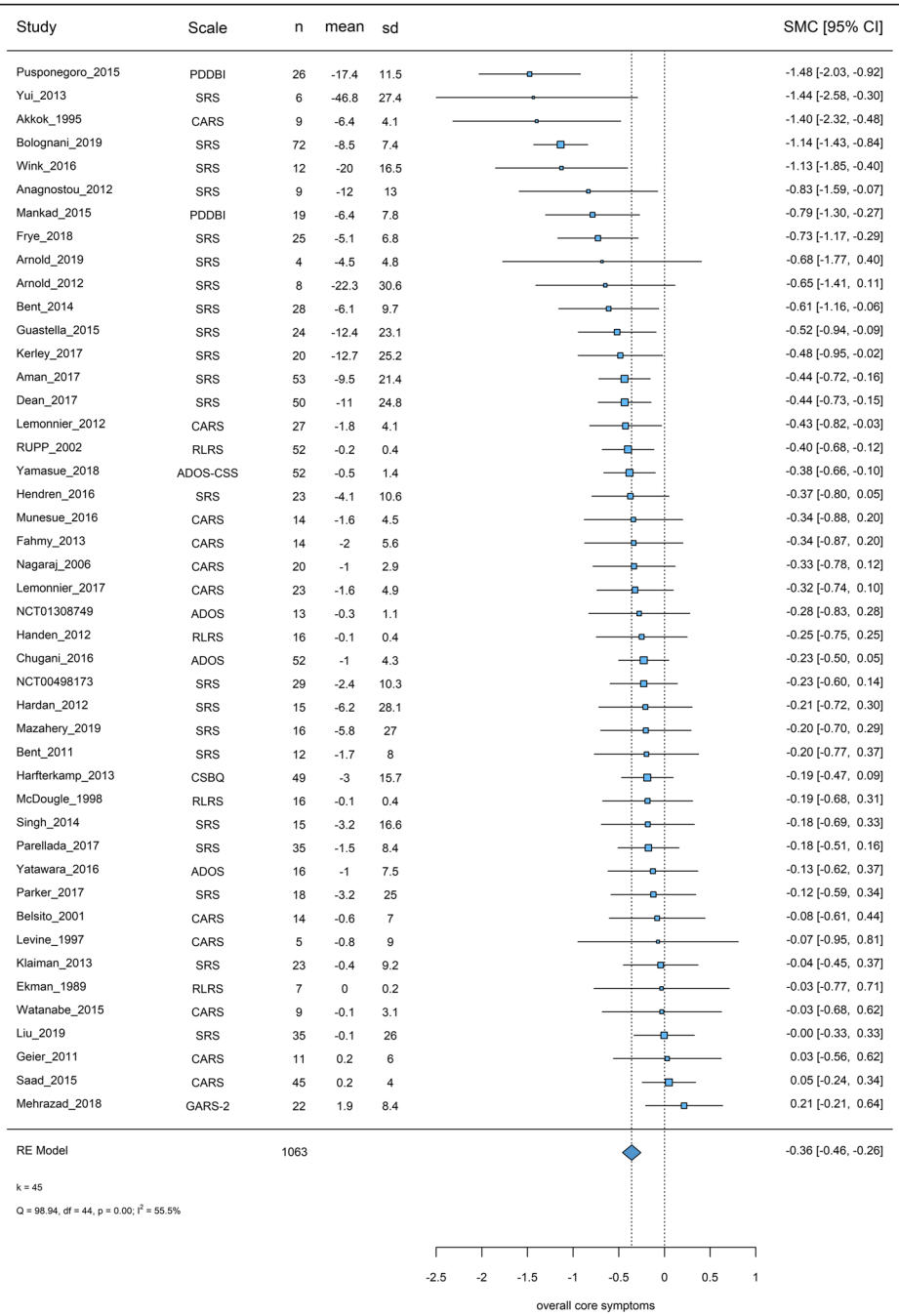


Fig. 4 Placebo response in scales measuring overall core symptoms. Squares and bars represent standardized mean changes (SMC) and 95% confidence intervals for each study. The size of the square is proportional to the weight of the study in the meta-analysis. The diamond represent the pooled SMC. Heterogeneity is quantified with a χ^2 test (*Q*) and I^2. *In Anagnostou 2012, we reversed baseline and endpoint values of SRS: in the manuscript, original baseline values were lower than endpoint in both placebo and oxytocin arms, meaning an increase of severity of symptoms during the study, which is not consistent with the reported positive effect size and the other outcomes (no reply from the corresponding author), in Saad 2015, CARS was rated by caregivers but it was unclear if also filled by clinicians (no reply from the corresponding author), as well as in RUPP 2002 and Handen 2012 the Ritvo-Freeman Life Rating Scale was rated by caregivers. Scale: the scale used (clinician rated scales based on observation or interviews were preferred in the primary analysis); *n*: the number of participants on placebo; mean: mean change from baseline to endpoint (negative values for improvement); sd: the standard deviation used for the standardization (baseline standard deviations were preferred); SMC: standardized mean changes, 95% CI: 95% confidence intervals, k= total number of studies included in the analysis

Meta-regressions Larger placebo response was associated with more trial sites (– 0.025 [– 0.045, – 0.005] SMC units per site, not significant when an outlier was removed) [68], and low risk of bias in allocation concealment (– 0.252 [– 0.485, – 0.019]). In the multivariable model, allocation concealment and number of sites were both significant.

Number of medications and the use of placebo lead-in had not sufficient data for all outcomes, while number of arms, selective reporting, and the use of the threshold of core symptoms did not have sufficient data for meta-regressions in overall core symptoms.

Secondary outcomes

Placebo response by type of rater

Results based on scales filled by different type of raters (Additional file 3: eAppendix-6.4) were similar to those of meta-regressions by type of rater (one effect size per study, clinician ratings were preferred whenever available).

CGI-I positive response rates

The overall positive response rate as defined by at least much improvement in the CGI-I was 19% [16–22%] (k = 57, n = 1686, I^2 = 53%) (Fig. 5). The anchoring system of CGI was unclear in 35 studies, while seven considered both core and associated symptoms (three used OACIS [69]), three reported separate evaluations for global autism symptoms and for the trial target symptom, and three considered mainly core symptoms and nine associated symptoms (two reported the RUPP-framework [70]) (Table-S).

Post-hoc correlation analyses

Between covariates Exploratory correlations between covariates are presented in Additional file 3: eAppendix-5.4. Significant correlations with a Spearman's $|\rho| > 0.5$ were found between sample size and number of sites (ρ = 0.77), percentage of academic sites and sponsorship (ρ = – 0.52), as well as number of sites (ρ = – 0.51), risk of bias domain of sequence generation and the domain of selective reporting (ρ = 0.57), number of arms and percentage of participants on placebo (ρ = – 0.54), and between other covariates with a large proportion of missing data, e.g., baseline ABC-Irritability, BMI, CGI-S, and percentage of participants with intellectual disability (Additional file 3: eAppendix-5.3).

Between placebo and drug response SMCs of placebo and experimental intervention were correlated in social-communication difficulties (Spearman's ρ = 0.525, p < 0.001) and overall core symptoms (ρ = 0.539, p < 0.001), but no correlation was found in repetitive behaviors (ρ = 0.233, p = 0.096) (Additional file 3: eAppendix-6.5).

Discussion

In pharmacological and dietary supplement ASD trials, placebo response was substantial and comparable among core symptoms; about 20% of the participants were at least much improved with placebo. We found potential predictors of larger placebo response in at least one symptom domain, i.e., baseline irritability, the use of a threshold of core symptoms at inclusion, caregiver ratings, larger sample size and number of sites, lower risk of bias, flexible-dosing, and less recent publication year.

Predictors of placebo response

Participant-related factors

It has been argued that placebo response might be larger in children/adolescents than adults [71]. We did not find a difference between age groups or an effect of mean age. Nonetheless, extrapolations between age groups should be interpreted with caution because the majority of studies were in pediatric populations (87%). Other participant characteristics did not predict placebo response (e.g., sex, ethnicity, BMI, intellectual disability).

Low baseline severity has been found to predict placebo response in most psychiatric conditions [50]. We did not find an effect of baseline global severity (CGI-S), yet available data were sparse (baseline CGI-S was reported in less than half studies, k = 38, 44%) and narrowly ranged between 3.88 and 6 (Additional file 3: eAppendix-5.1); also because most of the studies required participants to be at least moderately ill (i.e., CGI-S ≥ 4). Baseline severity in core symptoms could not be analyzed as a potential predictor due to the large diversity of scales. On the other hand, we found that trials using a cut-off of core symptoms for inclusion might have a larger placebo response in repetitive behaviors, yet this association was not significant in a multivariable meta-regression and it might have been driven by three antidepressant trials that used a cut-off of the clinician-administered scale CYBOCS-PDD [65–67]. Trials that utilize a baseline score cut-off could be prone to regression to the mean effects as well as baseline score inflation, especially for clinician-administered scales and under participant recruitment pressure [72]. These effects could be partially avoided by using different scales at assessing participants for inclusion and as primary outcomes [73], yet this might be challenging given the lack of optimal scales in ASD. Centralized raters blind to inclusion criteria might also reduce baseline inflation and increase inter-rater reliability, yet the execution of the trial could become complicated [72]. Since inflated scores are usually very close to the inclusion cut-off, a potential solution could be that the primary analysis is conducted by including participants with a higher cut-off (that is blinded to the investigators) than the inclusion cut-off [74].

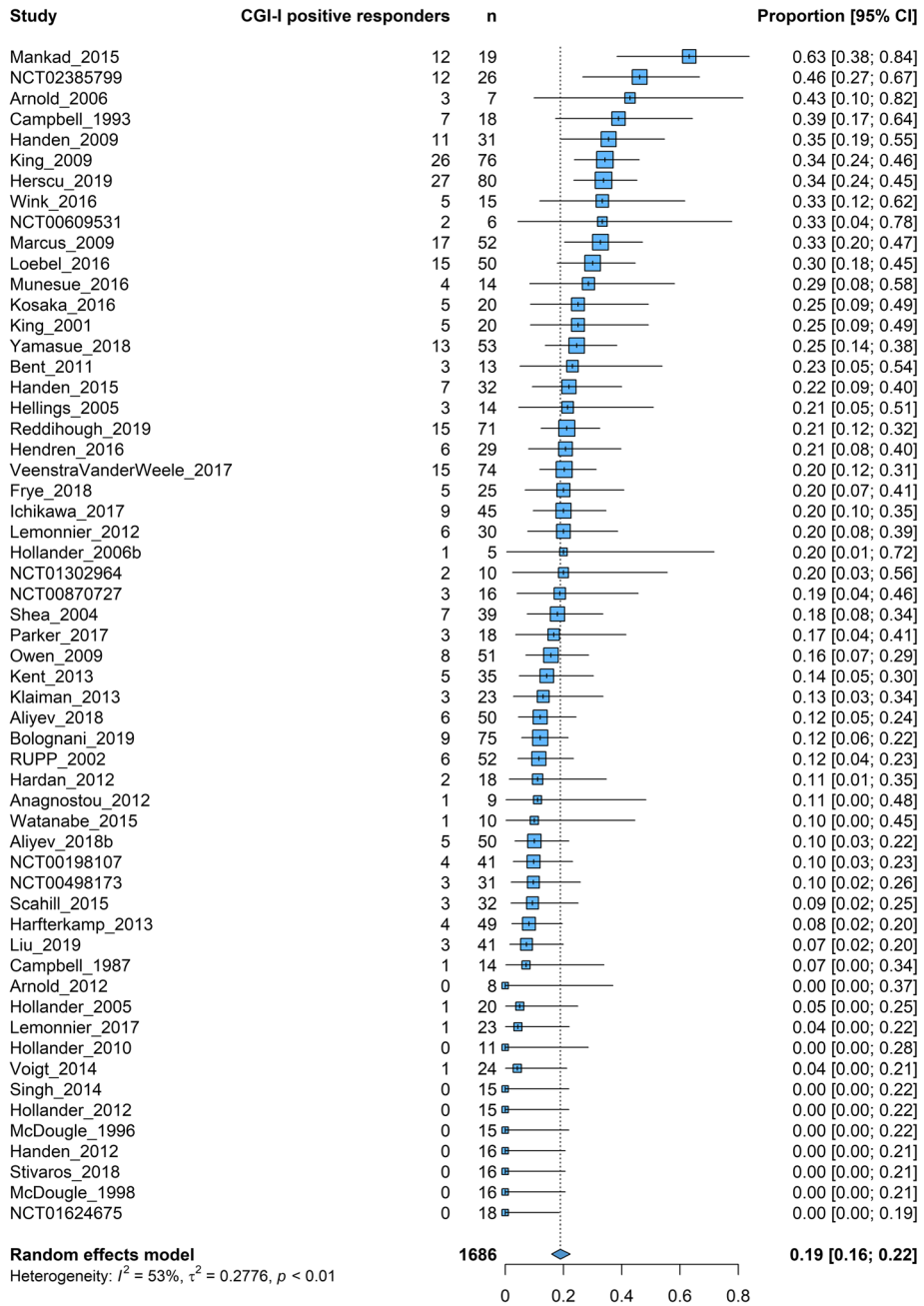


Fig. 5 CGI-I positive placebo response. Square and bars represent the point estimate of the proportion of responders and its 95% confidence interval for each study. The size of the squares is proportional to the weight of the study. The diamonds represent the pooled proportion and its 95% confidence intervals for each subgroup and overall. Heterogeneity is quantified with a χ^2 test (Q) and I^2. CGI-I positive responders: number of participants with a positive response defined as at least much improvement in the CGI-I (if not reported, it was imputed using a validated method); Total: total number of participants on placebo

The presence of an associated condition was required as inclusion criteria in about one-third of the trials (29 out of 86), and it was not found to predict placebo response. Nevertheless and since co-occurring symptoms and diagnoses are highly prevalent in participants with ASD [5], it can be expected that participants in other studies had also associated symptoms of varying levels. Accordingly, the median of baseline ABC-Irritability was 17.18 IQR [13.71–22.70], while normative data suggested a mean of 12.8 [75]. Thus, our sample in general could be consisted of participants with somewhat higher levels of irritability. Indeed, the most frequently investigated

associated condition in our sample was irritability (k = 15) and the presence of an associated condition was correlated to baseline ABC-Irritability (ρ = 0.49, p < 0.001, Additional file 3: eAppendix-5.4). We found that baseline ABC-Irritability was associated with a larger placebo response in social-communication difficulties, yet this association was not significant in a multivariable meta-regression. The contrary was found in a quite large trial (n = 149) investigating citalopram for repetitive behaviors, yet participants had lower levels of irritability (mean ABC-Irritability = 11.2) [9]. Additionally, a small 8-week observational study investigating the effects of participation in a study protocol suggested that placebo-effects may be mainly observed in children with higher levels of irritability [76]. Such participation effects could be decreased by a screening phase with adequate duration, which could also investigate the stability of symptoms and incorporate a potential washout of psychotropic drugs. However, no effect was found for the use of a washout phase and there were not enough data to investigate the use of a placebo lead-in phase, which is in general not recommended [72].

Design- and intervention-related factors

Caregiver ratings seemed to be more prone to placebo response in social-communication difficulties, but the effect was not consistent in multivariable meta-regressions. It has been argued that placebo-by-proxy effects are important components of placebo response in child/adolescent psychiatry, since they can alter caregiver perception of symptoms (thus improving directly scores in caregiver scales), and/or modify caregiver behaviors toward children and subsequently improving symptoms (thus improving scores also in non-caregiver scales) [71, 77]. In addition, many of the existing scales were not designed to measure change but rather as screening (e.g., SRS [31]) or diagnostic tools (CARS [32] and ADOS [78]), and efforts have been made for their improvement and adaptation, such as the ADOS calibrated severity score [79]. Given the lack of optimal scales, CGI has been extensively used and it is recommended for all trials irrespective of their target in order to investigate global autism symptoms and incorporate both core and associated symptoms [80, 81]. However, the anchoring system of CGI should be clearly reported, since it could vary materially among trials with different target symptoms (Table-S).

Therefore, there is a critical need to develop standardized and sensitive measures of core symptoms, which do not solely depend on caregivers [82, 83]. The semi-structured interview of VABS might be a promising measure of change in social-communication difficulties [33], with potential sensitivity to detect efficacy [68, 84] and empirically derived cut-offs of minimal-clinical-important differences [85]. Recent instruments have also been developed, among others the Brief Observation of Social Communication Change (BOSCC) [86, 87], the Autism Behavior Inventory [88], and the Autism Impact Measure (AIM) [89], but their utilization is yet to be determined. Patient- (or parent-) reported outcomes have also gained recently greater attention [90], yet they should not be considered immune to placebo-effects [91]. The utilization of scales that require more extensive training and experience (e.g., ADOS, BOSCC, and VABS) might be challenging in larger scale trials, and thus a low inter-rater reliability could increase the variance of measurements and subsequently decrease drug-placebo differences. A notable example is the multi-center arbaclofen trial [84], in which VABS should have been completed by the same clinician and caregiver for each participant. However, there was quite low adherence to the protocol (rater change in about 25% of the participants), potentially because VABS-Socialization was a secondary outcome, not expected to be sensitive in the context of the trial. A post-hoc per-protocol analysis of no rater change found a significant improvement of arbaclofen in comparison to placebo, in contrast to the non-significant difference of the primary analysis [84]. Therefore, proper training of the raters and inter-rater reliability of the measurements as well as guidance and adherence to the protocol should be ensured, especially in multi-site trials.

Sample size and number of sites have been suggested as predictors of placebo response [50, 51, 92]. We also found that a larger sample size was associated with a larger placebo response in repetitive behaviors, yet the results might be driven by three antidepressant trials [65–67]. This association could also be explained by a potential publication bias and the small-study effects found in the funnel plot (see Additional file 3: eAppendix-6.2), since the results of less precise trials with larger placebo response in repetitive behaviors might have been not published. Additionally, sample size was closely related to the number of sites (Spearman's ρ = 0.77, p < 0.001, see Additional file 3: eAppendix-6), which predicted placebo response in overall core symptoms, yet the latter was driven by another outlier study with 26 sites [68]. Trials with more sites were more frequently industry-sponsored (ρ = 0.27, p = 0.04) and consisted of less academic sites (ρ = − 0.51, p < 0.001). It should be noted though that the majority of included studies were single-center (median number of sites 1 IQR [1–4]), had academic sites (about 83% consisted only of academic sites), and small sample sizes (median 45 IQR [30–91]); therefore, the results could not be extrapolated to a wider range of potential values. Nevertheless, more sites and the recruitment of non-academic professional sites, which could have less experience and enroll

competitively, might increase variability, be prone to less rigorous participant selection and baseline score inflation [73, 74, 92]. Therefore, trials should be well powered, yet extremely large sample sizes could be avoided, as well as sites should be carefully selected and their number should be kept at the minimum feasible.

Studies with low risk of bias in other biases (mainly baseline imbalance) and allocation concealment were associated with larger placebo response in social-communication and overall core symptoms, respectively. It is intriguing that studies with a better quality in terms of risk of bias might have a larger placebo response. However, the above risk of bias domains evaluate the randomization process, and in inadequately randomized trials, control groups might have a poorer prognosis [93].

The association between dosing schedule and placebo response can be puzzling, e.g., both flexible- [94] and fixed-dosing schedules [95] have been associated with larger placebo responses in depression. We found an association between flexible-dosing and larger placebo response in repetitive behaviors, yet it was driven by three antidepressant trials [65–67]. Flexible-dosing could allow dose optimization guided by clinical response and/or the occurrence of side effects. The dose titration schedule and criteria as well as the starting dose and dose ranges should be carefully selected in the context of large placebo responses. For example, in one the aforementioned antidepressant trials, large placebo responses (> 25% reduction from baseline in CYBOCS-PDD) might have impeded dose escalation from a low starting dose (2 mg of fluoxetine) to a stable appropriate dose (> 10 mg) for sufficient duration of treatment (> 4 weeks) [67]. On the other hand, dose-response studies are a special type of fixed-dosing studies that might be prone to larger placebo responses. They are multi-arm and participants have an increased chance to receive active medication, as well as larger sample sizes and multiple sites are usually required. These factors have been associated with a larger placebo response in psychiatry [50], yet not all of them were replicated in our analysis, probably due to the limited number of studies with those characteristics. A notable example is the dose-response study of aripiprazole [96], which had a placebo positive response rate of 33% in comparison to 16% in the similarly designed but flexible-dosing study [97] (Fig. 3). However, this has not always been observed, such as in risperidone trials, i.e., 14% in the dose-response study [98] in comparison to 12% [53] and 18% [99] in the flexible-dose studies.

Country of origin and type of experimental intervention (pharmacological or dietary supplement) was not found to predict placebo response, in contrast to a previous meta-analysis [11], which included also many Iranian trials with risperidone-combined treatments that were excluded from our review (combination treatments such as risperidone + placebo were excluded, see Additional file 3: eAppendix-4). Therefore, the findings in the previous meta-analysis could have been confounded by larger responses in combined placebo groups, i.e., response of risperidone + placebo.

There is no clear consensus about the adequate trial duration and half of the included studies had a duration of at least 12 weeks, yet the duration of the trial should be based on the mechanism of action of the experimental intervention and a longer duration could be required in order to observe sustained changes in core symptoms [100]. We did not find an effect of trial duration, yet shorter-term trials have been associated with larger placebo response in psychiatry [50]. However, in longer-term trials including young children, anticipated developmental trajectories could also explain placebo effects and subsequently mask drug-placebo differences [101]. Therefore, developmentally based scales might be necessary to overcome this challenge [82] as well as trial designs could include additional follow-up assessments in order to confirm stability of improvement [101].

In most psychiatric disorders, placebo response has increased over a period of 60 years [49, 50, 102], but this trend was not replicated in ASD trials, which were more recent, mainly published between 2009 and 2017. Even, placebo response in social-communication difficulties might have decreased over years. However, this effect was not found when ABC-Irritability was included in multivariable meta-regression. Temporal changes in the definition of ASD and research practices might play an important role per se, as differences between ASD and neurotypical populations might have been decreased over the years [103].

Limitations

Our analysis has limitations. First, our analysis focused on placebo response in core symptoms of pharmacological and dietary supplement interventions. Therefore, we did not investigate placebo response in associated symptoms or of psychological/behavioral or multimodal interventions, which could also be of interest. However, core symptoms was the apparent focus in about 40% of the included trials, while many trials focused on associated symptoms, mainly irritability or ADHD symptoms. Second, there was a large diversity of scales used as well as a wide variability of their use, e.g., different CGI-I anchoring systems. Third, moderators of drug-placebo differences were not investigated and efforts to minimize placebo response could also affect drug response, since they were correlated in social-communication difficulties and overall core symptoms, but not in repetitive

behaviors (Additional file 3: eAppendix-6.5). In addition, some predictors might have a different impact on placebo and drug response [51]. Nevertheless, a more fine-grained analysis was impeded by the use of diverse experimental interventions with essentially different mechanisms of action (Additional file 3: eAppendix-5.1), e.g., contrary to schizophrenia [49, 51, 102], for which antipsychotics are the cornerstone of treatment [104].

Fourth, a common estimated pre-post correlation was used, but effect sizes were not materially changed in sensitivity analyses (Additional file 3: eAppendix-6.1). Fifth, despite the large number of eligible studies, about half did not provide data in spite of our efforts (authors of 85% of included studies published after 1990 could be contacted, with a reminder e-mail in case of no response, and 17% of them provided additional data/clarifications, Additional file 3: eAppendix-4), and a priori we did not use data from the whole crossover period (in forty trials), in order to avoid carry-over effects [14]. Sixth, due to the fact that information for many predictors, especially for participant-related factors (Additional file 3: eAppendix-5), was missing in many studies, we could not employ a full multivariable meta-regression and we focused on a series of univariable meta-regressions. Therefore, we cannot exclude the possibility of omitted variable bias in the results, i.e., the fact that the effect of the omitted variables may be added to the predictor considered in the univariable meta-regression. It should be noted that meta-regressions of aggregate data have an observational nature and they are prone to ecological fallacy, thus our findings should be considered exploratory and hypothesis-generating, considering also that there was no adjustment to multiple testing. Accordingly, individual-participant-data meta-analysis could allow for a more fine-grained analysis and further elucidate the impact of participant-level factors, such as age, sex, as well as baseline severity of core/associated symptoms.

Conclusions

In order to increase the detection of efficacy of experimental interventions for ASD, high-quality and adequately powered trials are required, and predictors of placebo response should be considered. Extremely large sample sizes could be avoided and when multiple sites are needed, they should be carefully selected, trained, and monitored as well as their number should be kept at the minimum feasible. This would also facilitate a more rigorous selection of participants and a higher inter-rater reliability of measurements. Furthermore, scales that do not solely depend on caregiver reports could be selected as primary outcomes, since placebo-by-proxy effects are expected. Nevertheless, our findings highlight the urgent need for optimal and developmentally-based measures of change in core symptoms [82, 83]. The mechanism of action of the experimental intervention could guide the selection of an appropriate, yet sufficiently long, trial duration as well as of the dose schedule and dose ranges. Participant-related factors, such as age, sex, and baseline severity of core/associated symptoms as well as factors that could differentially moderate drug response warrant further investigation. Last, in order to facilitate comparability between studies and synthesis of evidence, trials should better characterize their participants and improve their reporting, including the CGI anchoring system.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13229-020-00372-z.

Additional file 1.
Additional file 2.
Additional file 3.

Abbreviations

ABC-L/SW: Aberrant Behavior Checklist-Lethargy/Social Withdrawal; ABC-S: Aberrant Behavior Checklist-Stereotypic Behavior; ADHD: Attention-deficit/hyperactivity disorder; ADI-R: Autism Diagnostic Interview-Revised; ADOS: Autism Diagnostic Observation Scale; ASD: Autism spectrum disorders; BOSCC: Brief Observation of Social Communication Change; BMI: Body mass index; CARS: Childhood Autism Rating Scale; CGI: Clinical global impression; (C)YBOCS-PDD: (Children) Yale Obsessive Compulsive Scale-Pervasive Developmental Disorders; DSM: Diagnostic and Statistical Manual of Mental Disorders; ICD: International Classification of Diseases; RLRS: Ritvo-Freeman Real Life Rating Scale; SD: Standard deviation; SMC: Standardized mean change; SRS: Social Responsiveness Scale; VABS: Vineland Adaptive Behavior Scale

Acknowledgements

We would like to thank Farhad Sokraneh, information specialist of the Cochrane Schizophrenia Group, who conducted the first search in electronic databases, Dr. Yikang Zhu for the translation of a Chinese study and Prof. Toshi Furukawa for the translation of a Japanese study. We would like to thank the following authors that the kindly contribution to the review by provided additional data and/or clarifications about their studies: Adi Aran, Kaat Alaerts, Nadir Aliyev, Eugene Arnold, Haim Belmaker, Yéhézkel Ben-Ari, Leventhal Bennet, Stephen Bent, Helena Brentani, Jan Buitelaar, Ana Maria Castejon, Michael Chez, Torsten Danfors, Paulo Fontoura, Robert Grimaldi, Paul Gringras, Alexander Häge, Antonio Hardan, Robert Hendren, Janet Kern, Bruno Leheup, Wenn Liu, Raquel Martinez, James McCracken, Tali Nir, Deborah Pearson, Jeanette Ramer, Dan Rossignol, Kevin Sanders, Elisa Santocchi, Renato Scifo, Sarah Shea, Lawrence Scahill, Jeremy Veenstra-Vanderweele Paul Wang, David Wilensky, Hidenori Yamasue, and Lingli Zhang.

Authors' contributions

SS (study design, study selection, data extraction, contacting authors for additional data, statistical analysis, interpretation of the data, drafting the first version of the manuscript, study supervision), OC (study selection, data extraction, contacting authors for additional data), JST(data extraction, technical support with Access database), IB (study selection), MK (study selection), AR (data extraction), AC (study selection), GD (study selection), MH (technical support with Access database), DF (study design, interpretation of the data), DM (statistical analysis, interpretation of the data), TC (interpretation of the data), DGM (interpretation of the data), MP (study design, interpretation of the data), CA (study design, interpretation of the data), SL (study design, data extraction, statistical analysis, interpretation of the data, drafting the first version of the manuscript, study supervision). All authors critically reviewed the manuscript for important intellectual content. The author(s) read and approved the final manuscript.

Funding
This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 777394 for the project AIMS-2-TRIALS. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and AUTISM SPEAKS, Autistica, SFARI. CA, MP and DF were supported by the Spanish Ministry of Science, Innovation and Universities. Instituto de Salud Carlos III, co-financed by ERDF Funds from the European Commission, "A way of making Europe", CIBERSAM, Madrid Regional Government (B2017/BMD-3740 AGES-CM-2), European Union Structural Funds and European Union Seventh Framework Program and H2020 Program; Fundación Familia Alonso, Fundación Alicia Koplowitz and Fundación Mutua Madrileña. The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Open access funding provided by Projekt DEAL.

Availability of data and materials
All data generated during this study are included in this published article (and its supplementary information files). The datasets analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
In the last 3 years, Stefan Leucht has received honoraria as a consultant/advisor and/or for lectures from LB Pharma, Otsuka, Lundbeck, Boehringer Ingelheim, LTS Lohmann, Janssen, Johnson&Johnson, TEVA, MSD, Sandoz, SanofiAventis, Angelini, Recordati, Sunovion, and Geodon Richter. David Fraguas has been a consultant and/or has received fees from Angelini, Eisai, IE4Lab, Janssen, Lundbeck, and Otsuka. He has also received grant support from Instituto de Salud Carlos III (Spanish Ministry of Science, Innovation and Universities) and from Fundación Alicia Koplowitz. Mara Parellada has received educational honoraria from Otsuka, research grants from FAK and Fundación Mutua Madrileña (FMM), Instituto de Salud Carlos III (Spanish Ministry of Science, Innovation and Universities) and European ERANET and H2020 calls, travel grants from Otsuka and Janssen. Consultant for Exeltis and Servier. Celso Arango has been a consultant to or has received honoraria or grants from Acadia, Angelini, Gedeon Richter, Janssen Cilag, Lundbeck, Otsuka, Roche, Sage, Sanofi, Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion and Takeda. In the last 3 years, Maximilian Huhn has received speakers honoraria from Janssen. Declan Murphy has received consulting fees from Roche. The other authors have nothing to disclose.

Author details
[1]Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, Ismaningerstr. 22, 81675 Munich, Germany. [2]Department of Child and Adolescent Psychiatry, School of Medicine, Balçova Dokuz Eylul University, İzmir, Turkey. [3]Department of Experimental and Clinical Medicine, Psychiatric Clinic University Hospital 'Gaspare Rodolico', University of Catania, Catania, Italy. [4]Department of Psychiatry, Spedali Civili Hospital, Brescia, Italy. [5]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy Social Foundation Bamberg, Teaching Hospital of the University of Erlangen, Erlangen, Germany. [6]Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, Institute of Psychiatry and Mental Health, IiSGM, CIBERSAM, Madrid, Spain. [7]School of Medicine, Universidad Complutense, Madrid, Spain. [8]Department of Primary Education, University of Ioannina, Ioannina, Greece. [9]Faculté de Médecine, Université Paris Descartes, Paris, France. [10]Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. [11]Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.

Received: 1 April 2020 Accepted: 14 August 2020
Published online: 26 August 2020

References

1. American Psychiatric A. Diagnostic and statistical manual of mental disorders (DSM-5®): American Psychiatric Pub; 2013.
2. Baio J, Wiggins L, Christensen DL, Maenner MJ, Daniels J, Warren Z, Kurzius-Spencer M, Zahorodny W, Robinson Rosenberg C, White T, et al. Prevalence of Autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 Sites, United States, 2014. MMWR Surveill Summ. 2018;67(6):1–23.
3. Baxter AJ, Brugha TS, Erskine HE, Scheurer RW, Vos T, Scott JG. The epidemiology and global burden of autism spectrum disorders. Psychol Med. 2015;45(3):601–13.
4. Buescher AVS, Cidav Z, Knapp M, Mandell DS. Costs of Autism Spectrum disorders in the United Kingdom and the United States costs of ASD in the UK and US costs of ASD in the UK and US. JAMA Pediatrics. 2014;168(8):721–8.
5. Howes OD, Rogdaki M, Findon JL, Wichers RH, Charman T, King BH, Loth E, McAlonan GM, McCracken JT, Parr JR, et al. Autism spectrum disorder: consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. J Psychopharmacol. 2018;32(1):3–29.
6. Jobski K, Hofer J, Hoffmann F, Bachmann C. Use of psychotropic drugs in patients with autism spectrum disorders: a systematic review. Acta Psychiatr Scand. 2017;135(1):8–28.
7. Berry-Kravis EM, Lindemann L, Jonch AE, Apostol G, Bear MF, Carpenter RL, Crawley JN, Curie A, Des Portes V, Hossain F, et al. Drug development for neurodevelopmental disorders: lessons learned from fragile X syndrome. Nat Rev Drug Discov. 2018;17(4):280–99.
8. Gribkoff VK, Kaczmarek LK. The need for new approaches in CNS drug discovery: Why drugs have failed, and what can be done to improve outcomes. Neuropharmacology. 2017;120:11–9.
9. King BH, Dukes K, Donnelly CL, Sikich L, McCracken JT, Scahill L, Holl e E, Bregman JD, Anagnostou E, et al. Baseline factors predicting placebo response to treatment in children and adolescents with autism spectrum disorders: a multisite randomized clinical trial. JAMA Pediatrics. 2013;167(11):1045–52.
10. Arnold LE, Farmer C, Kraemer HC, Davies M, Witwer A, Chuang S, Disilvestro R, McDougle CJ, McCracken J, Vitiello B *et al*: Moderators, mediators, and other predictors of risperidone response in children with autistic disorder and irritability. J Child Adolesc Psychopharmacol. 2010;20(2):83–93..
11. Masi A, Lampit A, Glozier N, Hickie IB, Guastella AJ. Predictors of placebo response in pharmacological and dietary supplement treatment trials in pediatric autism spectrum disorder: a meta-analysis. Transl Psychiatry. 2015;5:e640.
12. Masi A, Lampit A, DeMayo MM, Glozier N, Hickie IB, Guastella AJ. A comprehensive systematic review and meta-analysis of pharmacological and dietary supplement interventions in paediatric autism: moderators of treatment response and recommendations for future research. Psychol Med. 2017;47(7):1323–34.
13. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
14. Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. Int J Epidemiol. 2002;31(1):140–9.
15. Whiting-O'Keefe QE, Henke C, Simborg DW. Choosing the correct unit of analysis in Medical Care experiments. Med Care. 1984;22(12):1101–14.
16. Reichow B, Volkmar FR, Cicchetti DV. Development of the evaluative method for evaluating and determining evidence-based practices in autism. J Autism Dev Disord. 2008;38(7):1311–9.
17. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
18. Furukawa TA, Salanti G, Atkinson LZ, Leucht S, Ruhe HG, Turner EH, Chaimani A, Ogawa Y, Takeshima N, Hayasaka Y, et al. Comparative efficacy and acceptability of first-generation and second-generation antidepressants in the acute treatment of major depression: protocol for a network meta-analysis. BMJ Open. 2016;6(7):e010919.

19. Carrasco M, Volkmar FR, Bloch MH. Pharmacologic treatment of repetitive behaviors in autism spectrum disorders: evidence of publication bias. Pediatrics. 2012;129(5):e1301–10.
20. Sturman N, Deckx L, van Driel ML. Methylphenidate for children and adolescents with autism spectrum disorder. Cochrane Database Syst Rev. 2017;11:Cd011144.
21. Fung LK, Mahajan R, Nozzolillo A, Bernal P, Krasner A, Jo B, Coury D, Whitaker A, Veenstra-Vanderweele J, Hardan AY. Pharmacologic treatment of severe irritability and problem behaviors in autism: a systematic review and meta-analysis. Pediatrics. 2016;137(Suppl 2):S124–35.
22. Hirsch LE, Pringsheim T. Aripiprazole for autism spectrum disorders (ASD). Cochrane Database Syst Rev. 2016;2016(6):Cd009043.
23. Williamson E, Sathe NA, Andrews JC, Krishnaswami S, McPheeters ML, Fonnesbeck C, Sanders K, Weitlauf A, Warren Z. AHRQ comparative effectiveness reviews. In: Medical Therapies for Children With Autism Spectrum Disorder-An Update. Rockville (MD): Agency for Healthcare Research and Quality (US); 2017.
24. Williams K, Brignell A, Randall M, Silove N, Hazell P. Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). Cochrane Database Syst Rev. 2013;(8):Cd004677. https://community.cochrane.org/style-manual/references/reference-types/cochrane-publications.
25. Hurwitz R, Blackmore R, Hazell P, Williams K, Woolfenden S. Tricyclic antidepressants for autism spectrum disorders (ASD) in children and adolescents. Cochrane Database Syst Rev. 2012;(3):Cd008372. https://community.cochrane.org/style-manual/references/reference-types/cochrane-publications.
26. Nye C, Brice A. Combined vitamin B6-magnesium treatment in autism spectrum disorder. Cochrane Database Syst Rev. 2005;(4):Cd003497. https://community.cochrane.org/style-manual/references/reference-types/cochrane-publications.
27. Fraguas D, Díaz-Caneja CM, Pina-Camacho L, Moreno C, Durán-Cutilla M, Ayora M, González-Vioque E, de Matteis M, Hendren RL, Arango C. Dietary interventions for autism spectrum disorder: a meta-analysis. Pediatrics. 2019;144(5):e20183218.
28. Aman MG, Singh NN, Stewart AW, Field CJ. Psychometric characteristics of the aberrant behavior checklist. Am J Ment Defic. 1985;89(5):492–502.
29. Sparrow SS. Vineland Adaptive Behavior Scales. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology. New York, NY: Springer New York; 2011. p. 2618–21.
30. Scahill L, McDougle CJ, Williams SK, Dimitropoulos A, Aman MG, McCracken JT, Tierney E, Arnold LE, Cronin P, Grados M, et al. Children's Yale-Brown obsessive compulsive scale modified for pervasive developmental disorders. J Am Acad Child Adolesc Psychiatry. 2006;45(9):1114–23.
31. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, Metzger LM, Shoushtari CS, Splinter R, Reich W. Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. J Autism Dev Disord. 2003;33(4):427–33.
32. Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). J Autism Dev Disord. 1980;10(1):91–103.
33. Anagnostou E, Jones N, Huerta M, Halladay AK, Wang P, Scahill L, Horrigan JP, Kasari C, Lord C, Choi D, et al. Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. Autism. 2014;19(5):622–36.
34. Scahill L, Aman MG, Lecavalier L, Halladay AK, Bishop SL, Bodfish JW, Grondhuis S, Jones N, Horrigan JP, Cook EH, et al. Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. Autism. 2013;19(1):38–52.
35. European Medicines Agency: Guideline on the clinical development of medicinal products for the treatment of Autism Spectrum Disorder (ASD). In. Edited by CHMP; 2018: EMA/CHMP/598082/592013.
36. Bolte EE, Diehl JJ: Measurement tools and target symptoms/skills used to assess treatment response for individuals with autism spectrum disorder. 2013, 43(11):2491-2501.
37. Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. Int Clin Psychopharmacol. 2005;20(1):49–52.
38. Samara MT, Spineli LM, Furukawa TA, Engel RR, Davis JM, Salanti G, Leucht S. Imputation of response rates from means and standard deviations in schizophrenia. Schizophr Res. 2013;151(1-3):209–14.
39. Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Available from https://www.cochrane-handbook.org. (see https://community.cochrane.org/style-manual/references/reference-types/cochrane-publications.
40. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5(1):13.
41. Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ: Empirical assessment of within-arm correlation imputation in trials of continuous outcomes. In. Rockville: Agency for Healthcare Research and Quality (US); 2012: EHC141-EF.
42. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. J Clin Epidemiol. 2006;59(1):7–10.
43. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–88.
44. Becker BJ. Synthesizing standardized mean-change measures. Br J Math Stat Psychol. 1988;41(2):257–78.
45. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36(3):1–48.
46. Schwarzer G, Chemaitelly H, Abu-Raddad LJ, Rücker G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. Res Synth Methods. 2019;10(3):476–83.
47. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
48. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000;56(2):455–63.
49. Leucht S, Chaimani A, Leucht C, Huhn M, Mavridis D, Helfer B, Samara M, Cipriani A, Geddes JR, Salanti G, et al. 60years of placebo-controlled antipsychotic drug trials in acute schizophrenia: Meta-regression of predictors of placebo response. Schizophr Res. 2018;201:315–23.
50. Weimer K, Colloca L, Enck P. Placebo eff ects in psychiatry: mediators and moderators. Lancet Psychiatry. 2015;2(3):246–57.
51. Leucht S, Chaimani A, Mavridis D, Leucht C, Huhn M, Helfer B, Samara M, Cipriani A, Geddes JR, Davis JM. Disconnection of drug-response and placebo-response in acute-phase antipsychotic drug trials on schizophrenia? Meta-regression analysis. Neuropsychopharmacology. 2019;44(11):1955–66.
52. Sandler A. Placebo effects in developmental disabilities: implications for research and practice. Ment Retard Dev Disabil Res Rev. 2005;11(2):164–70.
53. RUPP. Risperidone in Children with Autism and Serious Behavioral Problems. N Engl J Med. 2002;347(5):314–21.
54. Agid O, Kapur S, Arenovich T, Zipursky RB. Delayed-onset hypothesis of antipsychotic action: a hypothesis tested and rejected. Arch Gen Psychiatry. 2003;60(12):1228–35.
55. Schwarzer G. meta: An R package for meta-analysis. R news. 2007;7(3):40–5.
56. R Core Team. R: A language and Environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019. https://www.R-project.org/.
57. Akkok FG. Bahar; Oktem, Ferhunde; Reid, Larry D.; Sucuoglu, Bulbin: Otizm'de naltrekson sagaltiminin davranissal ve biyokimyasal boyutlari. Turk Psikiyatri Dergisi. 1995;6(4):251–62.
58. Pusponegoro HD, Ismael S, Firmansyah A, Sastroasmoro S, Vandenplas Y. Gluten and casein supplementation does not increase symptoms in children with autism spectrum disorder. Acta Paediatr. 2015;104(11):e500–5.
59. Campbell M, Anderson LT, Small AM, Adams P, Gonzalez NM, Ernst M. Naltrexone in autistic children: behavioral symptoms and attentional learning. J Am Acad Child Adolesc Psychiatry. 1993;32(6):1283–91.
60. Stivaros S, Garg S, Tziraki M, Cai Y, Thomas O, Mellor J, Morris AA, Jim C, Szumanska-Ryt K, Parkes LM, et al. Randomised controlled trial of simvastatin treatment for autism in young children with neurofibromatosis type 1 (SANTA). Mol Autism. 2018;9:12.
61. Bent S, Hendren RL, Zandi T, Law K, Choi JE, Widjaja F, Kalb L, Nestle J, Law P. Internet-based, randomized, controlled trial of omega-3 fatty acids for hyperactivity in autism. J Am Acad Child Adolesc Psychiatry. 2014;53(6):658–66.
62. Fahmy SF, El-Hamamsy M, Zaki O, Badary OA. Effect of L-carnitine on behavioral disorder in autistic children. Value Health. 2013;16(3):A15.
63. Geier DA, Kern JK, Davis G, King PG, Adams JB, Young JL, Geier MR. A prospective double-blind, randomized clinical trial of levocarnitine to treat autism spectrum disorders. Med Sci Monit. 2011;17(6):Pi15–23.

64. Belsito KM, Law PA, Kirk KS, Landa RJ, Zimmerman AW. Lamotrigine therapy for autistic disorder: a randomized, double-blind, placebo-controlled trial. J Autism Dev Disord. 2001;31(2):175–81.
65. King BH, Hollander E, Sikich L, McCracken JT, Scahill L, Bregman JD, Donnelly CL, Anagnostou E, Dukes K, Sullivan L, et al. Lack of efficacy of citalopram in children with autism spectrum disorders and high levels of repetitive behavior: citalopram ineffective in children with autism. Arch Gen Psychiatry. 2009;66(6):583–90.
66. Reddihough DS, Marraffa C, Mouti A, O'Sullivan M, Lee KJ, Orsini F, Hazell P, Granich J, Whitehouse AJO, Wray J, et al. Effect of fluoxetine on obsessive-compulsive behaviors in children and adolescents with autism spectrum disorders: a randomized clinical trial. JAMA. 2019;322(16):1561–9.
67. Herscu P, Handen BL, Arnold LE, Snape MF, Bregman JD, Ginsberg L, Hendren R, Kolevzon A, Melmed R, Mintz M, et al. The SOFIA study: negative multi-center study of low dose fluoxetine on repetitive behaviors in children and adolescents with autistic disorder. J Autism Dev Disord. 2019;50(9):3233–44.
68. Bolognani F, Del Valle RM, Squassante L, Wandel C, Derks M, Murtagh L, Sevigny J, Khwaja O, Umbricht D, Fontoura P. A phase 2 clinical trial of a vasopressin V1a receptor antagonist shows improved adaptive behaviors in men with autism spectrum disorder. Sci Transl Med. 2019;11(491):eaat7838.
69. Butter E, Mulick J. The Ohio autism clinical impressions scale (OACIS). Columbus, OH: Children's Research Institute; 2006.
70. Arnold LE, Aman MG, Martin A, Collier-Crespin A, Vitiello B, Tierney E, Asarnow R, Bell-Bradshaw F, Freeman BJ, Gates-Ulanet P, et al. Assessment in multisite randomized clinical trials of patients with autistic disorder: the Autism RUPP Network. Research Units on Pediatric Psychopharmacology. J Autism Dev Disord. 2000;30(2):99–111.
71. Weimer K, Gulewitsch MD, Schlarb AA, Schwille-Kiuntke J, Klosterhalfen S, Enck P. Placebo effects in children: a review. Pediatr Res. 2013;74(1):96–102.
72. Rutherford BR, Roose SP. A model of placebo response in antidepressant clinical trials. Am J Psychiatry. 2013;170(7):723–33.
73. Parellada M, Moreno C, Moreno M, Espliego A, de Portugal E, Arango C. Placebo effect in child and adolescent psychiatric trials. Eur Neuropsychopharmacol. 2012;22(11):787–99.
74. Mancini M, Wade AG, Perugi G, Lenox-Smith A, Schacht A. Impact of patient selection and study characteristics on signal detection in placebo-controlled trials with antidepressants. J Psychiatr Res. 2014;51:21–9.
75. Kaat AJ, Lecavalier L, Aman MG. Validity of the aberrant behavior checklist in children with autism spectrum disorder. J Autism Dev Disord. 2014;44(5):1103–16.
76. Jones RM, Carberry C, Hamo A, Lord C. Placebo-like response in absence of treatment in children with Autism. Autism Res. 2017;10(9):1567–72.
77. Whalley B, Hyland ME. Placebo by proxy: the effect of parents' beliefs on therapy for children's temper tantrums. J Behav Med. 2013;36(4):341–6.
78. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, Pickles A, Rutter M. The autism diagnostic observation schedule—Generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord. 2000;30(3):205–23.
79. Shumway S, Farmer C, Thurm A, Joseph L, Black D, Golden C. The ADOS calibrated severity score: relationship to phenotypic variables and stability over time. Autism Res. 2012;5(4):267–76.
80. Provenzani U, Fusar-Poli L, Brondino N, Damiani S, Vercesi M, Meyer N, Rocchetti M, Politi P. What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. Autism. 2020;24(2):274–84.
81. Aman MG, Novotny S, Samango-Sprouse C, Lecavalier L, Leonard E, Gadow KD, King BH, Pearson DA, Gernsbacher MA, Chez M. Outcome measures for clinical drug trials in autism. CNS spectrums. 2004;9(1):36–47.
82. Bishop S, Farmer C, Kaat A, Georgiades S, Kanne S, Thurm A. The need for a Developmentally Based measure of social communication skills. J Am Acad Child Adolesc Psychiatry. 2019;58(6):555–60.
83. Grzadzinski R, Janvier D, Kim SH. Recent developments in treatment outcome measures for young children with autism spectrum disorder (ASD). Semin Pediatr Neurol. 2020;34:100806.
84. Veenstra-VanderWeele J, Cook EH, King BH, Zarevics P, Cherubini M, Walton-Bowen K, Bear MF, Wang PP, Carpenter RL. Arbaclofen in children and adolescents with autism spectrum disorder: a randomized, controlled, phase 2 trial. Neuropsychopharmacology. 2017;42(7):1390–8.
85. Chatham CH, Taylor KI, Charman T, Liogier D'ardhuy X, Eule E, Fedele A, Hardan AY, Loth E, Murtagh L, Del Valle RM, et al. Adaptive behavior in autism: Minimal clinically important differences on the Vineland-II. Autism Res. 2018;11(2):270–83.
86. Grzadzinski R, Carr T, Colombi C, McGuire K, Dufek S, Pickles A, Lord C. Measuring changes in Social communication behaviors: preliminary development of the brief observation of social communication change (BOSCC). *J Autism Dev Disord*. 2016;46:2464–79.
87. Gengoux GW, Abrams DA, Schuck R, Millan ME, Libove R, Ardel CM, Phillips JM, Fox M, Frazier TW, Hardan AY. A Pivotal response treatment package for children with autism spectrum disorder: An RCT. Pediatrics. 2019;144(3):e20190178.
88. Bangerter A, Ness S, Aman MG, Esbensen AJ, Goodwin MS, Dawson G, Hendren R, Leventhal B, Khan A, Opler M, et al. Autism behavior inventory: a novel tool for assessing core and associated symptoms of autism spectrum disorder. J Child Adolesc Psychopharmacol. 2017;27(9):814–22.
89. Kanne SM, Mazurek MO, Sikora D, Bellando J, Branum-Martin L, Handen B, Katz T, Freedman B, Powell MP, Warren Z. The autism impact measure (AIM): initial development of a new tool for treatment outcome measurement. J Autism Dev Disord. 2014;44(1):168–79.
90. Graham Holmes L, Zampella CJ, Clements C, JP MC, Maddox BB, Parish-Morris J, Udhnani MD, Schultz RT, Miller JS. A Lifespan approach to patient-reported outcomes and quality of life for people on the autism spectrum. Autism Res. 2020;13(6):970–87.
91. Estevinho MM, Afonso J, Rosa I, Lago P, Trindade E, Correia L, Dias CC, Magro F. Gedii: Placebo effect on the health-related quality of life of inflammatory bowel disease patients: a systematic review with meta-analysis. J Crohns Colitis. 2018;12(10):1232–44.
92. Bridge JA, Birmaher B, Iyengar S, Barbe RP, Brent DA. Placebo response in randomized controlled trials of antidepressants for pediatric major depressive disorder. Am J Psychiatry. 2009;166(1):42–9.
93. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ. 1998;317(7167):1185–90.
94. Furukawa TA, Cipriani A, Atkinson LZ, Leucht S, Ogawa Y, Takeshima N, Hayasaka Y, Chaimani A, Salanti G. Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. Lancet Psychiatry. 2016;3(11):1059–66.
95. Khan A, Khan SR, Walens G, Kolts R, Giller EL. Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the food and drug administraton summary basis of approval reports. Neuropsychopharmacology. 2003;28(3):552–7.
96. Marcus RN, Owen R, Kamen L, Manos G, McQuade RD, Carson WH, Aman MG. A placebo-controlled, fixed-dose study of aripiprazole in children and adolescents with irritability associated with autistic disorder. J Am Acad Child Adolesc Psychiatry. 2009;48(11):1110–9.
97. Owen R, Sikich L, Marcus RN, Corey-Lisle P, Manos G, McQuade RD, Carson WH, Findling RL. Aripiprazole in the treatment of irritability in children and adolescents with autistic disorder. Pediatrics. 2009;124(6):1533–40.
98. Kent JM, Kushner S, Ning X, Karcher K, Ness S, Aman M, Singh J, Hough D. Risperidone dosing in children and adolescents with autistic disorder: a double-blind, placebo-controlled study. J Autism Dev Disord. 2013;43(8):1773–83.
99. Shea S, Turgay A, Carroll A, Schulz M, Orlik H, Smith I, Dunbar F. Risperidone in the treatment of disruptive behavioral symptoms in children with autistic and other pervasive developmental disorders. Pediatrics. 2004;114(5):e634–41.
100. European Medicines Agency (EMA). Guideline on the clinical development of medicinal products for the treatment of Autism Spectrum Disorder (ASD). Comittee: Committee for Medicinal Products for Human Use (CHMP), Reference number: EMA/CHMP/598082/2013. Published on: 21/11/2017. Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-development-medicinal-products-treatment-autism-spectrum-disorder-asd_en.pdf.
101. Potter LA, Scholze DA, Biag HMB, Schneider A, Chen Y, Nguyen DV, Rajaratnam A, Rivera SM, Dwyer PS, Tassone F, et al. A Randomized controlled trial of sertraline in young Children with autism spectrum disorder. Front Psychiatry. 2019;10:810.
102. Leucht S, Leucht C, Huhn M, Chaimani A, Mavridis D, Helfer B, Samara M, Rabaioli M, Bacher S, Cipriani A, et al. Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, bayesian meta-analysis, and meta-regression of efficacy predictors. Am J Psychiatry. 2017;174(10):927–42.

103. Rødgaard E-M, Jensen K, Vergnes J-N, Soulières I, Mottron L. Temporal changes in effect sizes of studies comparing individuals with and without autism: a meta-analysis. JAMA Psychiatry. 2019;76(11):1124–32.
104. Huhn M, Nikolakopoulou A, Schneider-Thoma J, Krause M, Samara M, Peter N, Arndt T, Bäckers L, Rothe P, Cipriani A, et al. Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. Lancet. 2019;394(10202):939–51.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC

Publication III

Validation of an imputation method to estimate the number of responders using the mean and standard deviation of the Clinical Global Impression Improvement scale in autism spectrum disorder

Citation: Siafis, S., Rodolico, A., Çıray, O., Murphy, D.G., Parellada, M., Arango, C., and Leucht, S. 2021. 'Imputing the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder', *Brain Sci*, 11. doi: 10.3390/brainsci11070908.

This article is licensed under a Creative Commons Attribution 4.0 International License: https://creativecommons.org/licenses/by/4.0/

brain sciences

MDPI

Brief Report

Imputing the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder

Spyridon Siafis [1,*], Alessandro Rodolico [2], Oğulcan Çıray [3], Declan G. Murphy [4], Mara Parellada [5,6], Celso Arango [5,6] and Stefan Leucht [1]

1 Department of Psychiatry and Psychotherapy, School of Medicine, Technical University of Munich, 81675 Munich, Germany; stefan.leucht@tum.de
2 Department of Experimental and Clinical Medicine, Psychiatric Clinic University Hospital 'Gaspare Rodolico', University of Catania, 95125 Catania, Italy; alessandro.rodolico@me.com
3 Department of Child and Adolescent Psychiatry, Mardin State Hospital, 47100 Artuklu, Mardin, Turkey; remziogulcanciray@gmail.com
4 Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London WC2R 2LS, UK; declan.murphy@kcl.ac.uk
5 Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, 28003 Madrid, Spain; parellada@hggm.es (M.P.); carango@hggm.es (C.A.)
6 School of Medicine, Universidad Complutense, 28040 Madrid, Spain
* Correspondence: spyridon.siafis@tum.de; Tel.: +49-894-140-6415; Fax: +49-894-140-4888

check for updates

Citation: Siafis, S.; Rodolico, A.; Çıray, O.; Murphy, D.G.; Parellada, M.; Arango, C.; Leucht, S. Imputing the Number of Responders from the Mean and Standard Deviation of CGI-Improvement in Clinical Trials Investigating Medications for Autism Spectrum Disorder. *Brain Sci.* **2021**, *11*, 908. https://doi.org/10.3390/brainsci11070908

Academic Editor: Shinichiro Nakajima

Received: 8 June 2021
Accepted: 7 July 2021
Published: 9 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Abstract: Introduction: Response to treatment, according to Clinical Global Impression-Improvement (CGI-I) scale, is an easily interpretable outcome in clinical trials of autism spectrum disorder (ASD). Yet, the CGI-I rating is sometimes reported as a continuous outcome, and converting it to dichotomous would allow meta-analysis to incorporate more evidence. Methods: Clinical trials investigating medications for ASD and presenting both dichotomous and continuous CGI-I data were included. The number of patients with at least much improvement (CGI-I $\leq$ 2) were imputed from the CGI-I scale, assuming an underlying normal distribution of a latent continuous score using a primary threshold $\theta = 2.5$ instead of $\theta = 2$, which is the original cut-off in the CGI-I scale. The original and imputed values were used to calculate responder rates and odds ratios. The performance of the imputation method was investigated with a concordance correlation coefficient (CCC), linear regression, Bland–Altman plots, and subgroup differences of summary estimates obtained from random-effects meta-analysis. Results: Data from 27 studies, 58 arms, and 1428 participants were used. The imputation method using the primary threshold ($\theta = 2.5$) had good performance for the responder rates (CCC = 0.93 95% confidence intervals [0.86, 0.96]; β of linear regression = 1.04 [0.95, 1.13]; bias and limits of agreements = 4.32% [−8.1%, 16.74%]; no subgroup differences $\chi^2 = 1.24$, p-value = 0.266) and odds ratios (CCC = 0.91 [0.86, 0.96]; $\beta = 0.96$ [0.78, 1.14]; bias = 0.09 [−0.87, 1.04]; $\chi^2 = 0.02$, p-value = 0.894). The imputation method had poorer performance when the secondary threshold ($\theta = 2$) was used. Discussion: Assuming a normal distribution of the CGI-I scale, the number of responders could be imputed from the mean and standard deviation and used in meta-analysis. Due to the wide limits of agreement of the imputation method, sensitivity analysis excluding studies with imputed values should be performed.

Keywords: response; meta-analysis; continuous outcomes; dichotomous outcomes

1. Introduction

There is still no approved medication for the core symptoms of autism spectrum disorder (ASD) (i.e., social communication difficulties and repetitive restricted behaviors [1]), yet a large number of medications are being investigated in an increasing number of randomized controlled trials (RCTs), with this number increasing sharply after 2008 [2]. Many

of these trials are pilot trials with small sample sizes and cannot provide definite answers, and given their increasing number, there is an ongoing need to comprehensively synthesize their evidence [2].

However, the lack of agreement on the selection of outcome measures for the core symptoms in clinical trials precludes the synthesis of evidence [3–5]. The available scales are, at best, "appropriate with conditions" [3,4], and given the lack of a "gold standard", the Clinical Global Impression scales (CGI-Severity and CGI-Improvement) [6,7] have been widely used in clinical trials of ASD [8,9] not only as important secondary outcomes, but also as the primary outcome [10]. CGI-Severity (CGI-S) is a seven-point scale used by clinicians to assess the current severity of illness, ranging from one ("normal, not at all ill") to seven ("among the most extremely ill patients") and usually measured at the trial's baseline and endpoint. CGI-Improvement (CGI-I) is a seven-point scale used by clinicians to measure global response compared to the baseline, ranging from one ("very much improved") to seven ("very much worse"). A clinically important response is frequently defined as at least much improvement (i.e., a number of participants with a CGI-I score of one or two) [11].

In addition, a comprehensive synthesis of evidence would require the combination of all available studies; however, some of them may present the CGI-I as a continuous outcome (i.e., with a mean and standard deviation). The conversion of continuous outcomes to dichotomous ones would allow the combination all available data across studies. Imputation methods of the number of responders from the means and standard deviations have been validated with depression [12] and schizophrenia scales [13]. The appropriateness of these methods might be questioned with the CGI-I, given the limited number of points of the CGI, as well as in ASD, given its heterogeneity and the small sample sizes of clinical trials (only 8.7% of RCTs included more than 100 participants [2]). Therefore, our aim was to validate the imputation of the responder rates from the means and standard deviations of the CGI-I in ASD trials. We compared the responder rates and odds ratios calculated from the original and imputed numbers of participants with a clinically important response to treatment.

2. Methods

2.1. Dataset

This is a secondary analysis which uses part of the dataset from a systematic review and meta-analysis on pharmacological and dietary supplement interventions for ASD (PROSPERO ID: CRD42019125317) [14,15]. A comprehensive literature search, study selection, and data extraction by at least two independent reviewers were conducted (last update search on 31 August 2020). Response to treatment was investigated as a secondary outcome in the reviews, and the CGI-I was extracted as continuous and dichotomous outcomes. In this analysis, we used 27 studies with 58 arms and 1428 participants that provided data on (1) the means and standard deviations (SDs) of the CGI-I and (2) the number of responders defined at least as much improved in the CGI-I (CGI-I $\leq$ 2). Data from the endpoint of the studies were used (the minimum duration of treatment was set at seven days). The intention-to-treat (ITT) data were preferred, and when only completer data was available, we assumed that participants lost to the follow-up did not respond.

The cut-off of the least much improvement (CGI-I 1 or 2) was investigated, which represents a clinically important response [11] and is frequently reported in clinical trials [10]. The responder rates using the original or imputed number of responders were calculated in each arm. The odds ratios (ORs) were also calculated for each non-reference arm in a study, using as a reference the placebo arm of the study or another active treatment (in the case of non-placebo-controlled trials).

2.2. Imputation Method

We used an imputation method validated with depression [12] and schizophrenia scales [13] which assumed a normal distribution of the scale (CGI-I in this analysis) given a

mean (μ) and standard deviation (σ). The number of responders of a threshold (θ) in the CGI-I (i.e., participants with a CGI-I score $\leq$ θ) could be calculated using the total number of participants assessed (n) and the probability of the lower tail of the distribution (*p*) for Z-score = $(\theta - \mu)/\sigma$ (Figure 1). Then, the number of responders was n * *p*.

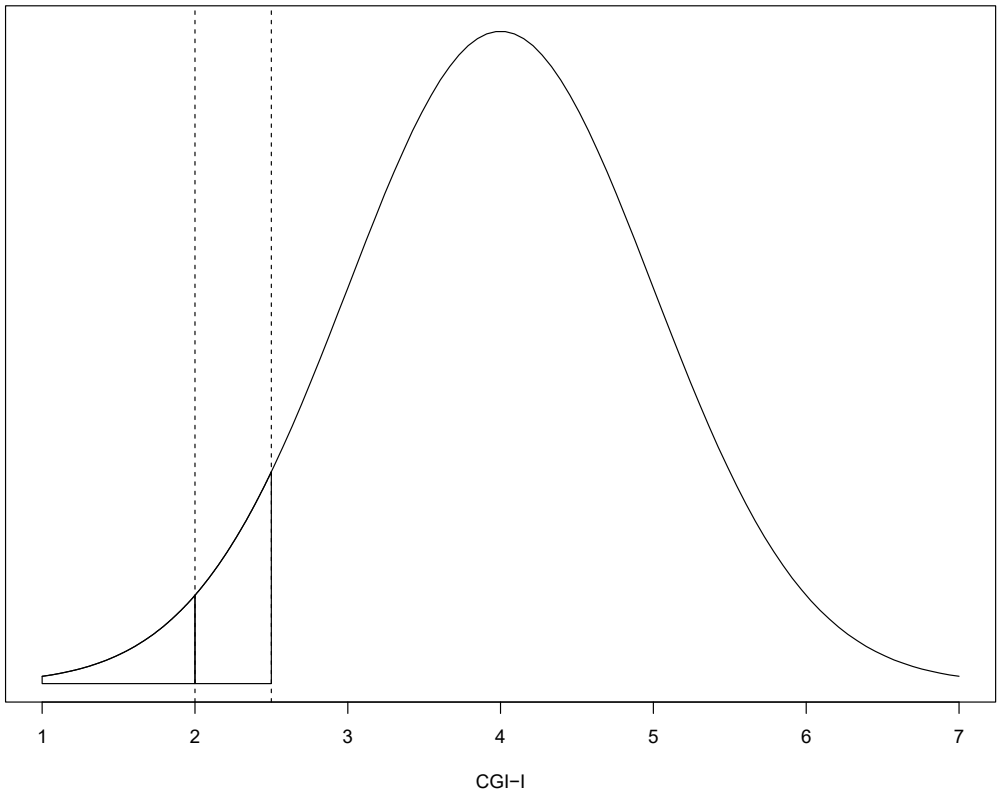


Figure 1. Underlying distribution of a latent CGI-I score, using an assumed normal distribution of the CGI-I, such as with μ = 4 and σ = 1. Under the assumption of a normal distribution, the probability (p) of at least much improvement (CGI-I = 2) could be calculated with Z-score = $(\theta - \mu)/\sigma$, where θ is a threshold of the response. As a primary threshold, we used θ = 2.5 for at least much improvement (CGI-I of 1 or 2, the blue and red shaded parts of the distribution), since it could be assumed that a patient with a score between 2 and 2.5 in the underlying latent continuous variable would have been classified as at least much improved. As a secondary threshold, we used θ = 2 (red shaded part of the distribution).

According to the work of Furukawa et al. in 2005 [12], when the CGI-I was used, responders were imputed using the threshold of θ = 2 (at least "much improved"). However, the CGI-I is a seven-point Likert-type scale, and an underlying latent continuous variable could be assumed which could have had different thresholds of mapping the discrete responses [16]. Both the ordinal scale scores and the scores of the latent continuous variable would have the same μ and σ, but the threshold θ for the discrete responses (e.g., of at least "much improved") would differ [16]. Therefore, we used a threshold of θ = 2.5 as the primary threshold to impute the number of responders (Figure 1), since a participant with a latent CGI-I continuous score ranging from 2 to 2.5 would have also been considered as at least "much improved". In a secondary analysis, we used a secondary threshold of θ = 2 to impute responders from the assumed normal distribution of the ordinal scale.

We calculated the responder rates from the original and imputed numbers of responders using the randomized number of participants as the denominator. We also calculated the odds ratios (OR) between the experimental and control investigations (placebo or another active treatment). The natural logarithm of the ORs (lnOR) was used in the analysis.

2.3. Assessment of Performance of the Imputation Method

2.3.1. Concordance Correlation Coefficient (CCC)

The agreement between the original and imputed responder rates and the lnORs were investigated with the concordance correlation coefficient (CCC) [17] and its 95% confidence intervals. The CCC ranged between −1 and 1 (perfect agreement).

2.3.2. Predictive Accuracy and Linear Regression Model

Linear regression models were used to determine the predictive accuracy of the imputation method, and a good imputation method should have a slope (β) and R^2 close to one and a low mean squared error (MSE).

2.3.3. Limits of Agreement and Bland–Altman Analysis

The Bland–Altman method was used to investigate the limits of agreement of the bias (i.e., the difference between the original and imputed values) [18,19]. In the Bland–Altman plot, the difference of the original and imputed values is presented in the *y*-axis, and their average is in the *x*-axis. The distribution of the difference was inspected for normality, and a Shapiro–Wilk test was conducted. The limits of agreement were represented with 95% confidence intervals, considering acceptable the ones found in the validation of the method in schizophrenia scales [13], i.e., −0.7% 95% CI (−9.8%, 8.4%) for the difference of the original and imputed responder rates and 0.06 95% CI (−0.24, 0.35) for the difference of the original and imputed lnORs. To investigate if the bias was proportional to the mean, a linear regression model of the differences on their mean (using the natural logarithms for both the responder rates and odds ratios) was conducted [18].

2.4. Meta-Analysis

We compared the pooled estimates from the meta-analysis using the original and imputed values. The responder rates (logit transformed and back-transformed for presentation) [20] and odds ratios (natural logarithm and back-transformed for presentation) were pooled in a random-effects meta-analysis [21]. Subgroup analysis was conducted to investigate the differences of the pooled estimates from the meta-analysis using the original and the imputed values (primary and secondary thresholds).

Analysis was conducted in R v4.0.3 [22]. The CCC, linear regression, and Bland–Altman limits were calculated with base R and epiR v2.0.17 [23]. The effect sizes and meta-analysis were calculated with metafor v2.4−0 [24] and meta v4.15−1 [25]. The data cleaning and graphs were completed using packages of tidyverse v13.0 [26]. The statistical threshold was set at two-sided alpha 5%.

3. Results

The results of the CCC, linear regression, and Bland–Altman analysis are presented in Table 1 and Figure 2 (responder rates) and Figure 3 (odds ratios).

3.1. Responder Rates

The responder rates derived from the imputed values using the primary threshold (θ = 2.5) were in good agreement with the original values (CCC 0.93, 95% confidence interval [0.89, 0.96]), and the imputation method had good predictive accuracy (β = 1.04 [0.95, 1.13], R^2 = 90.86%, MSE = 0.063) (Figure 2A, blue). The difference between the original and imputed values (normally distributed, Figure S1) was, on average, 4.32% with 95% confidence intervals [−8.1%, 16.74%] (Figure 2B, blue), and it was not proportional to the mean when natural logarithms were used (β = −0.034 [−0.135, 0.068]) (Figure 2C, blue).

Table 1. CCC and regression of response rates.

		Agreement	Predictive Accuracy			Bias	
	Number of Observations (k)	**CCC (95% CI)**	**β (95% CI) of Original (Y) and Imputed (X)**	**R^2 (%)**	**MSE**	**Bias and 95% Limits of Agreement**	**β (95% CI) of Difference (Y) and Mean (X)**
Responder Rates (Original 58 Observations)							
Primary Threshold	58	0.93 (0.89–0.96)	1.04 (0.95, 1.13)	90.86	0.063	4.32% (−8.1%, 16.74%)	−0.034 (−0.135, 0.068) *
Secondary Threshold	58	0.59 (0.48−0.69)	1.41 (1.26, 1.57)	85.01	0.0813	16.15% (−3.18%, 35.47%)	−0.028 (−0.177, 0.121) *
Log OR (Original 30 Observations)							
Primary Threshold	28	0.91 (0.81, 0.95)	0.96 (0.78, 1.14)	82.03%	0.495	0.09 (−0.87, 1.04)	0.06 (−0.120, 0.231)
Secondary Threshold	27	0.81 (0.63, 0.91)	0.90 (0.65, 1.15)	67.85%	0.664	0.24 (−1.05, 1.53)	0.086 (−0.164, 0.334)

* Natural logarithmic transformation of the responder rates. The dependent and independent variables of linear regressions are indicated with (Y) and (X), respectively.

On the other hand, the imputation method had poorer performance when the secondary threshold was used (θ = 2), with poor agreement (CCC = 0.59 [0.48, 0.69]) and predictive accuracy (β = 1.41 [1.26, 1.57], R^2 = 85.01%, MSE = 0.0813). This would mean that the original responder rates of 20% would correspond, on average, to imputed responder rates of 14.2% and from 50% to 35.46% (1.41 times higher) (Figure 2A, red). The difference between the original and imputed values (normally distributed, Figure S2) was larger on average (16.15% [−3.18%, 35.47%]) (Figure 2B, red) and not proportional to the mean when natural logarithms were used (β = −0.034 [−0.135, 0.068]) (Figure 2C, red). In comparison with the schizophrenia scales (bias −0.7% [−9.8%, 8.4%]) [13], the bias was larger and the limits of agreements were wider.

The summary estimates obtained from the meta-analysis of the imputed values using the secondary threshold (12.1% [8.8%, 16.4%]) were smaller than those obtained from the imputed values using the primary threshold (24.3% [19%, 30.4%]) or the original values (29.1% [23.2%, 35.8%]) (χ^2 = 22.22, *p*-value < 0.001) (Figure 2D). This was reflected in the post hoc two-by-two comparisons that found the summary estimates obtained from the imputed values using the secondary threshold were smaller than those using the primary threshold (χ^2 = 12.29, *p*-value < 0.001) or original values (χ^2 = 21, *p*-value < 0.001), while there was no difference between the latter two (χ^2 = 1.24, *p*-value = 0.266).

3.2. Odds Ratios

When the primary threshold was used (θ = 2.5), the imputed natural logarithm of the odds ratios was in good agreement with the original values (CCC 0.91, 95% confidence interval [0.81, 0.95]), and the imputation method had good predictive accuracy (β = 0.96 [0.78, 1.14], R^2 = 82.03%, MSE = 0.495) (Figure 3A, blue). The difference between the original and imputed values (normally distributed, Figure S3) was, on average, 0.09 with 95% confidence intervals [−0.87, 1.04] (Figure 3B, blue). This would mean that the original odds ratios were, on average, 1.1 (=$e^{0.09}$) times larger than the imputed values (95% CI [0.42, 2.83]). The differences were not proportional to the mean (β = 0.06 [−0.120, 0.231]) (Figure 3C, blue).

The imputation method using the secondary threshold (θ = 2) had poorer performance, with a CCC of 0.81 [0.63, 0.91]) and predictive accuracy of β = 0.90 [0.65, 1.15], R^2 = 67.85%, MSE = 0.664 (Figure 3A, red). The difference between the original and imputed values (normally distributed, Figure S4) was, on average, 0.24 [−1.05, 1.53] (Figure 2B, red), meaning that the original odds ratios were, on average, 1.27 (=$e^{0.24}$) times larger than the imputed values (95% [0.35, 4.62]). The differences were not proportional to the mean (β = 0.086 [−0.164, 0.334]) (Figure 3C, red). For both thresholds, the average bias was

similar, yet the limits of agreement were considerably wider than those found in the schizophrenia scales (0.06 [−0.24, 0.35]) [13].

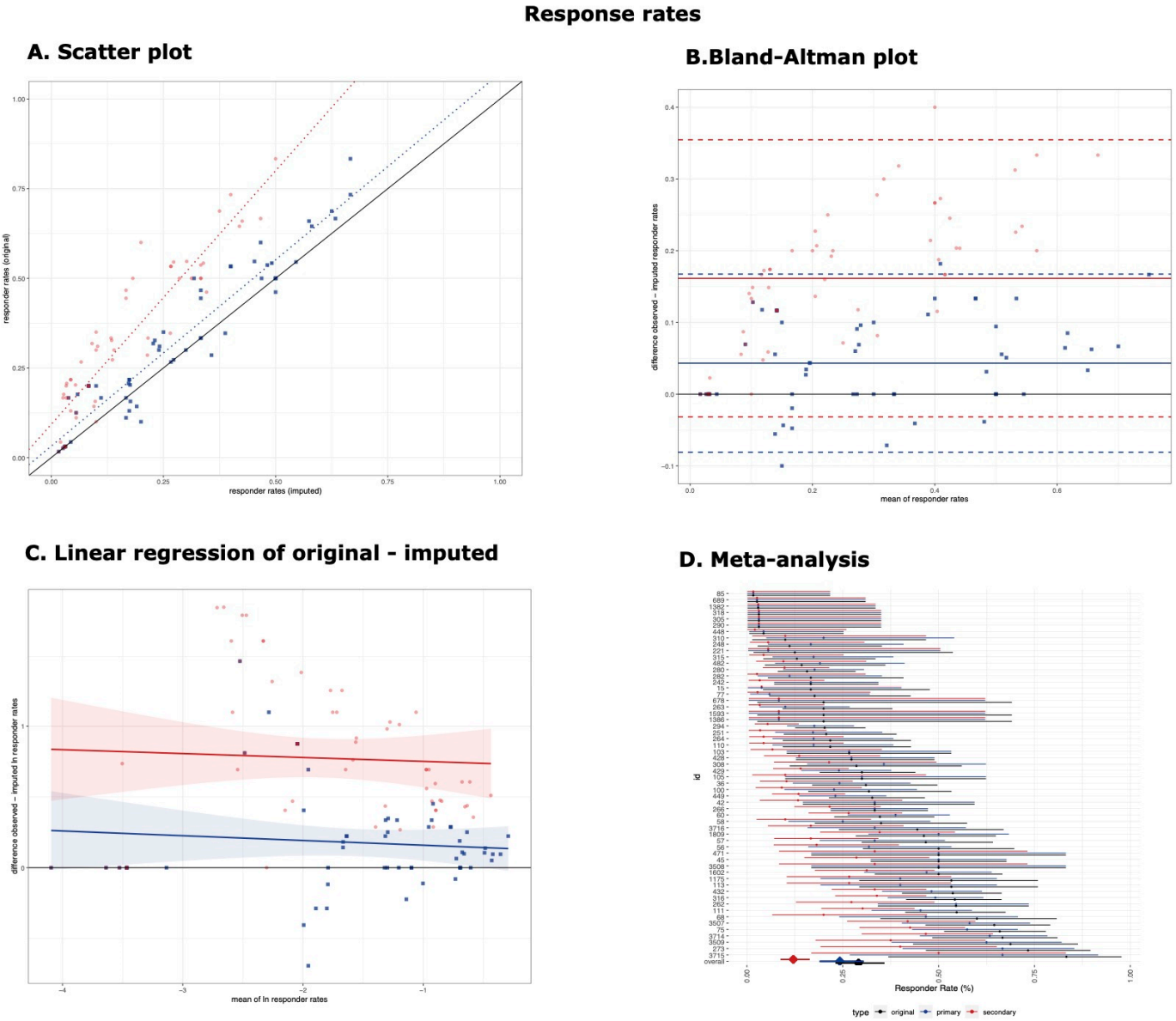


Figure 2. Response rates. (**A**) Scatter plot of response rates. Scatter plot of the comparison between original and imputed response rates (blue for the primary threshold and red for the secondary threshold). The black solid line represents the line of perfect correspondence. Blue and red dotted lines represent the linear regression model for the primary and secondary threshold. (**B**) Bland-Altman plot of response rates. The black solid line represents the optimal difference between original and imputed responder rates. The solid blue and red lines represent the median difference of the primary and secondary threshold, and the dashed blue and red dotted lines represent their 95% confidence intervals, corresponding to the limits of agreement. (**C**) Linear regression of original minus imputed ln responder rates. Linear regression of the difference between original and imputed natural logarithms of responder rates to their mean. Regression lines and its 95% confidence intervals are presented for the primary threshold (blue) and the secondary threshold (red). (**D**) Meta-analysis. Meta-analysis of responder rates using original values (black), imputed using the primary threshold (blue) and secondary threshold (red). Effect sizes with their 95% confidence intervals are presented with circles and error bars for individual arms and with diamonds and error bars for the pooled estimates.

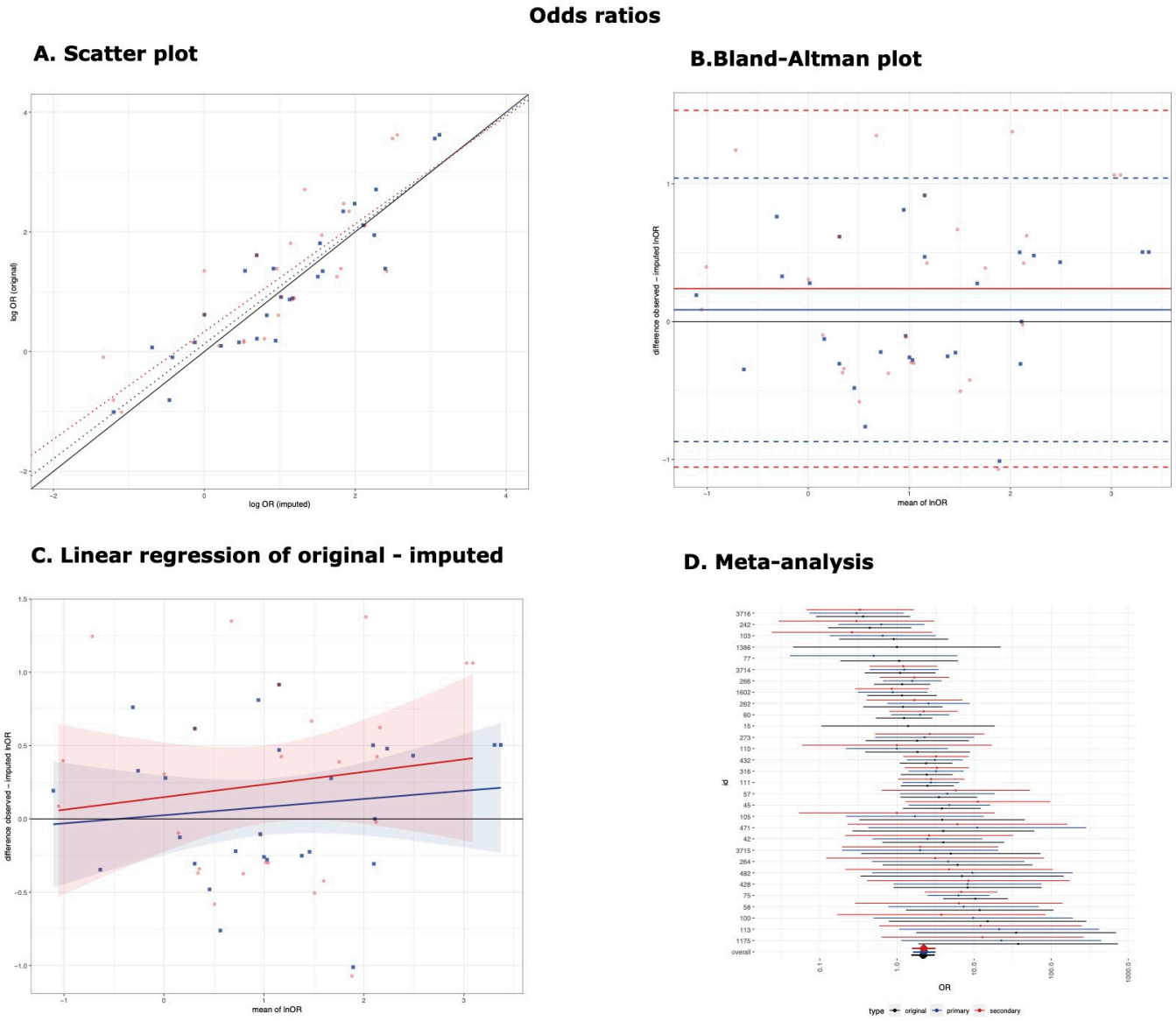


Figure 3. Odds ratios. (**A**) Scatter plot of lnORs. Scatter plot of the comparison between original and imputed lnORs (blue for the primary threshold and red for the secondary threshold). The black solid line represents the line of perfect correspondence. Blue and red dotted lines represent the linear regression model for the primary and secondary threshold. (**B**) Bland-Altman plot of lnORs. The black solid line represents the optimal difference between original and imputed lnORs. The solid blue and red lines represent the mean difference of the primary and secondary threshold, and the dashed blue and red dotted lines represent their 95% confidence interval of the difference, corresponding to the limits of agreement. (**C**) Linear regression of original minus imputed lnOR. Linear regression of the difference between original and imputed natural logarithms of odds ratios to their mean. Regression lines and its 95% confidence intervals are presented for the primary threshold (blue) and the secondary threshold (red). (**D**) Meta-analysis of odds ratios. Meta-analysis of odds ratios using original values (black), imputed using the primary threshold (blue) and secondary threshold (red). Effect sizes with their 95% confidence intervals are presented with circles and error bars for individual arms and with diamonds and error bars for the pooled estimates.

Nevertheless, no subgroup differences were found in the pooled estimates obtained from the meta-analysis, regardless of whether the original values (number of observations k = 30, 2.20 [1.56, 3.09]) or the imputed values using the primary (k = 28, 2.27 [1.64, 3.14]) or secondary threshold (k = 27, 2.23 [1.60, 3.11]) were used (χ^2 = 0.02, *p*-value = 0.991) (Figure 3D). No subgroup differences were found in the post hoc two-by-two comparisons (i.e., original versus imputed using the primary threshold (χ^2 = 0.02, *p*-value = 0.894),

original versus secondary threshold (χ^2 < 0.00, *p*-value = 0.949), and primary versus secondary threshold (χ^2 < 0.00, *p*-value = 0.945)). It should be noted that the odds ratios were not calculated in the case of double zeros (i.e., no responder in the experimental or control interventions). Therefore, some original observations were not paired with the imputed observations in these meta-analyses (2 out of 30 for the primary threshold and 3 out of 30 for the secondary threshold).

4. Discussion

In this analysis, we applied an imputation method previously validated mainly with depression [12] and schizophrenia scales [13] to estimate the number of responders from the means and standard deviations of the CGI-I in ASD. We further replicated the quite satisfactory performance of the imputation method, suggesting that the number of responders could be imputed from the CGI-I, and they could be used in the meta-analysis of the responder rates and odds ratios. Our findings also suggest that, since the imputation method assumed a normal distribution of the seven-point Likert-type CGI-I scale, an underlying latent continuous variable could be considered, and a higher threshold than the original could be used in the imputation method for better performance, such as with participants that were at least much improved (CGI-I $\leq$ 2), which would have had a score in the latent continuous variable $\leq$2.5. In a previous study validating the method in depression [12], the number of responders was imputed in a subset of studies from the CGI-I using the original threshold of "at least much improvement" (θ = 2), yet the specific performance on the CGI-I was not evaluated. Nevertheless, differences between the primary and secondary thresholds were less striking when the odds ratios were used in comparison with the response rates, since relative indices like odds ratios seem to remain constant across different thresholds and control event rates [27].

Our analysis would facilitate synthesis of evidence in ASD by allowing the conversion of the means and standard deviations of the CGI-I to number of responders and subsequent meta-analysis to incorporate all available data. There is still no consensus on the selection of the outcome measures of symptom change in ASD, so diverse scales that assess different symptom domains (e.g., social communication difficulties, repetitive behaviors, and problem behaviors) have been used across trials. The majority of them are not specifically designed to measure treatment response, and only a few have been used in more than 5% of clinical trials [9]. On the other hand, the CGI-I is recommended for use in clinical trials irrespective of their objective and clinical context in order to measure treatment response while incorporating all behavior symptom domains [8,9]. Therefore, pooled estimates derived from the number of responders according to the CGI-I might be more clinically interpretable than those from the standardized mean differences (SMDs) of diverse scales [28].

This analysis has certain limitations. First, there were considerable data for the responder rates (27 studies and 58 arms), yet the data points on the odds ratios were about half the amount (because a reference should be used in each study), also resulting in wider limits of agreements. Second, we focused on the clinically important response using the cut-off of "at least much improvement", or CGI-I $\leq$ 2. Therefore, the imputation method was not directly validated for the other cut-offs, such as "at least minimal improvement", or CGI-I $\leq$ 3. Third, our data were derived from clinical trials investigating pharmacological and dietary supplement interventions for ASD. Therefore, generalizability to psychosocial interventions or other fields of medicine should be further examined. Fourth, the imputation method assumes a normal distribution, yet scores from a Likert-type scale like the CGI-I might be frequently skewed. Indeed, potential skewness was suggested in 45% of the arms (when mean $-$ 1 < 2 * SD), and there was strong evidence of skewness in 5% of the arms (when mean $-$ 1 < SD) (Figure S5) [29]. Nevertheless, the performance of the imputation method was surprisingly satisfactory. Fifth, other methods to convert continuous to dichotomous effect sizes (e.g., from SMD to OR) have been proposed [30] and were not evaluated here, yet the method in this manuscript allows for the estimation

of the number of responders that could be used in meta-analysis of both the proportions (such as single-group meta-analysis of responder rates) and relative effects (such as odds ratios or relative risks).

In conclusion, the number of responders could be imputed when given a mean and standard deviation of CGI-I. The imputation method had better performance when an underlying latent continuous variable was considered and an appropriate threshold was used ($\theta = 2.5$ and not 2 for "at least much improvement"). The imputed number of responders could be used in meta-analysis of the responder rates and odds ratios. Given the wide limits of agreement between the original and imputed values, the robustness of the results of the main analysis should be investigated in a sensitivity analysis by excluding effect sizes derived from the imputed number of responders, as has been suggested previously [13].

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/brainsci11070908/s1, Figure S1: Histogram and QQ plot of original-imputed responder rates (primary threshold), Figure S2: Histogram and QQ plot of original-imputed responder rates (secondary threshold), Figure S3: Histogram and QQ plot of original-imputed lnOR (primary threshold), Figure S4: Histogram and QQ plot of original-imputed lnOR (secondary threshold), Figure S5: Investigation of skewness of CGI-I scores.

Author Contributions: S.S. (study design, study selection, data extraction, data analysis, interpretation of the data, and drafting the first version of the manuscript); A.R. (study selection, data extraction, and interpretation of the data); O.Ç. (study selection, data extraction, and interpretation of the data); D.G.M. (interpretation of the data); M.P. (interpretation of the data); C.A. (interpretation of the data); and S.L. (interpretation of the data and study supervision). All authors critically reviewed the manuscript for important intellectual content. The authors read and approved the final version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This project received funding from the Innovative Medicines Initiative 2 joint undertaking under grant agreement No. 777394 for the project AIMS−2-TRIALS. This joint undertaking received support from the European Union's Horizon 2020 research and innovation program, EFPIA, AUTISM SPEAKS, Autistica, and SFARI.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We would like to thank the authors that participated in the completion of the main systematic review and meta-analysis: Hui Wu, Marc Krause, Anna Ceraso, Giacomo Deste, Maximilian Huhn, David Fraguas, Antonia San José Cáceres, Dimitris Mavridis, and Tony Charman. We would like to thank the other contributors to the systematic review: Farhad Sokraneh, information specialist of the Cochrane Schizophrenia Group who conducted the first search in electronic databases, Yikang Zhu for the translation of a Chinese study, and Toshi Furukawa for the translation of a Japanese study. We would like to thank the following authors that kindly contributed to the review by providing additional data or clarifications about their studies: Adi Aran, Kaat Alaerts, Nadir Aliyev, Eugene Arnold, Haim Belmaker, Yéhézkel Ben-Ari, Leventhal Bennet, Stephen Bent, Helena Brentani, Jan Buitelaar, Ana Maria Castejon, Michael Chez, Torsten Danfors, Paulo Fontoura, Robert Grimaldi, Paul Gringras, Alexander Häge, Randi J. Hagerman, Antonio Hardan, Robert Hendren, Janet Kern, Bruno Leheup, Wenn Liu, Raquel Martinez, James McCracken, Tali Nir, Deborah Pearson, Laura Politte, Jeanette Ramer, Dan Rossignol, Kevin Sanders, Elisa Santocchi, Renato Scifo, Sarah Shea, Lawrence Scahill, Jeremy Veenstra-Vanderweele Paul Wang, David Wilensky, Hidenori Yamasue, and Lingli Zhang.

Conflicts of Interest: In the last 3 years, Stefan Leucht has received honoraria as a consultant and advisor and for lectures from LB Pharma, Otsuka, Lundbeck, Boehringer Ingelheim, LTS 760 Lohmann, Janssen, Johnson & Johnson, TEVA, MSD, Sandoz, SanofiAventis, Angelini, Recordati, Sunovion, and Geodon Richter. Celso Arango has been a consultant to or has received honoraria or grants from Acadia, Angelini, Gedeon Richter, Janssen Cilag, Lundbeck, Otsuka, Roche, Sage, Sanofi, Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion, and Takeda. The

other authors have nothing to disclose. Mara Parellada has received educational honoraria from Otsuka, research grants from FAK and the Fundación Mutua Madrileña (FMM), Instituto de Salud Carlos III (Spanish Ministry of Science, Innovation and Universities), and European ERANET and H2020 calls, and travel grants from Otsuka and Janssen. As a consultant for Exeltis and Servier. Declan G. Murphy has received consulting fees from Roche.

References

1. Díaz-Caneja, C.M.; State, M.W.; Hagerman, R.J.; Jacquemont, S.; Marín, O.; Bagni, C.; Umbricht, D.; Simonoff, E.; De Andrés-Trelles, F.; Kaale, A.; et al. A white paper on a neurodevelopmental framework for drug discovery in autism and other neurodevelopmental disorders. *Eur. Neuropsychopharmacol.* **2021**. [CrossRef]
2. Tromans, S.; Adams, C. Brief Report: Autism Spectrum Disorder: A Comprehensive Survey of Randomized Controlled Trials. *J. Autism. Dev. Disord.* **2018**, *48*, 3228–3232. [CrossRef] [PubMed]
3. Scahill, L.; Aman, M.G.; Lecavalier, L.; Halladay, A.K.; Bishop, S.L.; Bodfish, J.W.; Grondhuis, S.; Jones, N.; Horrigan, J.P.; Cook, E.H.; et al. Measuring repetitive behaviors as a treatment endpoint in youth with autism spectrum disorder. *Autism* **2013**, *19*, 38–52. [CrossRef] [PubMed]
4. Anagnostou, E.; Jones, N.; Huerta, M.; Halladay, A.K.; Wang, P.; Scahill, L.; Horrigan, J.P.; Kasari, C.; Lord, C.; Choi, D.; et al. Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. *Autism* **2014**, *19*, 622–636. [CrossRef] [PubMed]
5. McConachie, H.; Parr, J.R.; Glod, M.; Hanratty, J.; Livingstone, N.; Oono, I.P.; Robalino, S.; Baird, G.; Beresford, B.; Charman, T.; et al. Systematic review of tools to measure outcomes for young children with autism spectrum disorder. *Health Technol. Assess.* **2015**, *19*, 1–506. [CrossRef]
6. Guy, W. *ECDEU Assessment Manual for Psychopharmacology*; US Department of Health and Welfare: Washington, DC, USA, 1976; pp. 534–537.
7. Choque Olsson, N.; Bolte, S. Brief report: "Quick and (not so) dirty" assessment of change in autism: Cross-cultural reliability of the Developmental Disabilities CGAS and the OSU autism CGI. *J. Autism. Dev. Disord.* **2014**, *44*, 1773–1778. [CrossRef] [PubMed]
8. Aman, M.G.; Novotny, S.; Samango-Sprouse, C.; Lecavalier, L.; Leonard, E.; Gadow, K.D.; King, B.H.; Pearson, D.A.; Gernsbacher, M.A.; Chez, M. Outcome measures for clinical drug trials in autism. *CNS Spectr.* **2004**, *9*, 36–47. [CrossRef] [PubMed]
9. Provenzani, U.; Fusar-Poli, L.; Brondino, N.; Damiani, S.; Vercesi, M.; Meyer, N.; Rocchetti, M.; Politi, P. What are we targeting when we treat autism spectrum disorder? A systematic review of 406 clinical trials. *Autism* **2020**, *24*, 274–284. [CrossRef]
10. King, B.H.; Hollander, E.; Sikich, L.; McCracken, J.T.; Scahill, L.; Bregman, J.D.; Donnelly, C.L.; Anagnostou, E.; Dukes, K.; Sullivan, L.; et al. Lack of efficacy of citalopram in children with autism spectrum disorders and high levels of repetitive behavior: Citalopram ineffective in children with autism. *Arch. Gen. Psychiatry* **2009**, *66*, 583–590. [CrossRef]
11. Busner, J.; Targum, S.D. The clinical global impressions scale: Applying a research tool in clinical practice. *Psychiatry* **2007**, *4*, 28–37.
12. Furukawa, T.A.; Cipriani, A.; Barbui, C.; Brambilla, P.; Watanabe, N. Imputing response rates from means and standard deviations in meta-analyses. *Int. Clin. Psychopharmacol.* **2005**, *20*, 49–52. [CrossRef] [PubMed]
13. Samara, M.T.; Spineli, L.M.; Furukawa, T.A.; Engel, R.R.; Davis, J.M.; Salanti, G.; Leucht, S. Imputation of response rates from means and standard deviations in schizophrenia. *Schizophr. Res.* **2013**, *151*, 209–214. [CrossRef]
14. Siafis, S.; Çıray, O.; Schneider-Thoma, J.; Bighelli, I.; Krause, M.; Rodolico, A.; Ceraso, A.; Deste, G.; Huhn, M.; Fraguas, D.; et al. Placebo response in pharmacological and dietary supplement trials of autism spectrum disorder (ASD): Systematic review and meta-regression analysis. *Mol. Autism* **2020**, *11*, 66. [CrossRef] [PubMed]
15. Siafis, S.; Çıray, O.; Wu, H.; Schneider-Thoma, J.; Bighelli, I.; Krause, M.; Rodolico, A.; Ceraso, A.; Deste, G.; Huhn, M.; et al. Pharmacological and dietary supplement treatments for autism spectrum disorder: A systematic review and network meta-analysis. 2021; In preparation.
16. Liddell, T.M.; Kruschke, J.K. Analyzing ordinal data with metric models: What could possibly go wrong? *J. Exp. Soc. Psychol.* **2018**, *79*, 328–348. [CrossRef]
17. Lin, L.I. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* **1989**, *45*, 255–268. [CrossRef] [PubMed]
18. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **1999**, *8*, 135–160. [CrossRef]
19. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *1*, 307–310. [CrossRef]
20. Schwarzer, G.; Chemaitelly, H.; Abu-Raddad, L.J.; Rücker, G. Seriously misleading results using inverse of Freeman-Tukey double arcsine transformation in meta-analysis of single proportions. *Res. Synth. Methods* **2019**, *10*, 476–483. [CrossRef]
21. DerSimonian, R.; Laird, N. Meta-analysis in clinical trials revisited. *Contemp. Clin. Trials* **2015**, *45*, 139–145. [CrossRef]
22. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018; Available online: https://www.r-project.org/ (accessed on 22 March 2021).
23. Stevenson, M.; Nunes, T.; Heuer, C.; Marshall, J.; Sanchez, J.; Thornton, R.; Reiczigel, J.; Robison-Cox, J.; Sebastiani, P.; Solymos, P. epiR: Tools for the Analysis of Epidemiological Data. R Package Version 2.0.17. 2020. Available online: https://cran.r-project.org/package=epiR (accessed on 22 March 2021).
24. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **2010**, *36*, 1–48. [CrossRef]

25. Balduzzi, S.; Rücker, G.; Schwarzer, G. How to perform a meta-analysis with R: A practical tutorial. *Evid. Based Ment. Health* **2019**, *22*, 153–160. [CrossRef] [PubMed]
26. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.D.A.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J. Welcome to the Tidyverse. *J. Open Source Softw.* **2019**, *4*, 1686. [CrossRef]
27. Furukawa, T.A.; Akechi, T.; Wagenpfeil, S.; Leucht, S. Relative indices of treatment effect may be constant across different definitions of response in schizophrenia trials. *Schizophr. Res.* **2011**, *126*, 212–219. [CrossRef] [PubMed]
28. Johnston, B.C.; Alonso-Coello, P.; Friedrich, J.O.; Mustafa, R.A.; Tikkinen, K.A.O.; Neumann, I.; Vandvik, P.O.; Akl, E.A.; Da Costa, B.R.; Adhikari, N.K.; et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *Cmaj* **2016**, *188*, 25–32. [CrossRef]
29. Higgins, J.P.T.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. *Cochrane Handbook for Systematic Reviews of Interventions*; John Wiley & Sons: Hoboken, NJ, USA, 2019.
30. Meister, R.; Von Wolff, A.; Kriston, L. Odds ratios of treatment response were well approximated from continuous rating scale scores for meta-analysis. *J. Clin. Epidemiol.* **2015**, *68*, 740–751. [CrossRef] [PubMed]