Psychological Medicine, Page 1 of 12. © Cambridge University Press 2017
doi:10.1017/S0033291716003457

ORIGINAL ARTICLE

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

A. Masi[1], A. Lampit[2], M. M. DeMayo[1], N. Glozier[1], I. B. Hickie[1] and A. J. Guastella[1]*

[1] *Autism Clinic for Translational Research, Brain and Mind Centre, Central Clinical School, Sydney Medical School, University of Sydney, Camperdown, NSW, Australia*
[2] *Regenerative Neuroscience Group, Brain and Mind Centre, University of Sydney, Camperdown, NSW, Australia*

Background. Autism spectrum disorders (ASDs) are pervasive and multifactorial neurodevelopmental conditions, characterized by impairments in social communication and interaction, and restricted, repetitive patterns of behaviour, interests or activities. Treatment options to ameliorate symptoms of ASDs are limited. Heterogeneity complicates the quest for personalized medicine in this population. Our aim was to investigate if there are baseline characteristics of patients that moderate response or trial design features that impede the identification of efficacious interventions for ASDs.

Method. Literature searches of EMBASE, MEDLINE and PsycINFO identified 43 studies for qualitative assessment of baseline characterization of participants and 37 studies for quantitative analysis of moderators of treatment response. Criteria included blinded randomized controlled trials (RCTs) in paediatric ASD, with at least 10 participants per arm or 20 overall, of oral treatments, including pharmacological interventions and dietary supplements.

Results. Random-effects meta-analysis of 1997 participants (81% male) identified three moderators associated with an increase in treatment response: trials located in Europe and the Middle-East; outcome measures designated primary status; and the type of outcome measure. Inconsistent reporting of baseline symptom severity and intellectual functioning prevented analysis of these variables. Qualitative synthesis of baseline characteristics identified at least 31 variables, with only age and gender reported in all trials. Biological markers were included in six RCTs.

Conclusions. Few trials reported adequate baseline characteristics to permit detailed analysis of response to treatment. Consideration of geographical location, baseline severity and intellectual function is required to ensure generalizability of results. The use of biological markers and correlates in ASD trials remains in its infancy. There is great need to improve the application of baseline characterization and incorporation of biological markers and correlates to permit selection of participants into homogeneous subgroups and to inform response to treatment in ASD.

Received 2 October 2016; Revised 8 December 2016; Accepted 8 December 2016

Key words: Autism spectrum disorders, meta-analyses, paediatrics, treatment trials.

Introduction

Autism spectrum disorders (ASDs) are a diverse group of complex, persistent and pervasive neurodevelopmental conditions characterized by impairments in social communication and interaction and restrictive, repetitive patterns of behaviour, interests or activities (American Psychiatric Association, 2013). Clinical manifestation of impairment ranges from mild to profound where all aspects of daily functioning and quality of life are affected (Perry *et al.* 2009; de Vries & Geurts, 2015). Heterogeneity of clinical presentation is a hallmark of ASDs and considered a function of complex and heterogeneous genetic underpinnings (Huguet *et al.* 2013; Persico & Napolioni, 2013), that is further complicated by concomitant co-morbidities, including mental health (e.g. anxiety, depression and attention-deficit/hyperactivity disorder), intellectual disability (Simonoff *et al.* 2008; Matson & Shoemaker, 2009) and medical co-morbidities (e.g. seizure and sleep disorders, gastrointestinal, mitochondrial, and immune system dysregulation and oxidative stress markers) (Rossignol & Frye, 2012; Chaidez *et al.* 2014; Cohen *et al.* 2014; El Achkar & Spence, 2015).

* Address for correspondence: A. J. Guastella, Ph.D., Brain and Mind Centre, University of Sydney, 100 Mallett Street, Camperdown, NSW 2050, Australia.
(Email: adam.guastella@sydney.edu.au)

Behavioural interventions intensively implemented early in life are considered the 'gold standard' treatment for behavioural symptoms associated with ASDs (Eldevik *et al.* 2009). However, these types of interventions are focused on early childhood and are expensive and time consuming to implement, rendering them inaccessible to a substantial proportion of the ASD population. Oral treatments offer significant advantages over intensive therapies across all age groups and socio-economic groups. However, treatment options within this class of intervention continue to be limited. The US Food and Drug Administration has approved only two drugs, risperidone and aripiprazole, for treatment of symptoms associated with ASD, including severe tantrums, aggression and self-injurious behaviour. The efficacy, safety and tolerability of prescription medications, some prescribed off-label for symptoms associated with ASD, have been the subject of numerous reviews (Jesner *et al.* 2007; McPheeters *et al.* 2011; Carrasco *et al.* 2012; Ching & Pringsheim, 2012; Hurwitz *et al.* 2012; Williams *et al.* 2013; Guastella & Hickie, 2016). Furthermore, the efficacy and safety of treatments for specific symptoms, such as irritability (Fung *et al.* 2016), and the role of the response to placebo in treatment outcome (Masi *et al.* 2015) have been investigated. Despite this extensive body of literature, many trials of interventions initially identified as efficacious have not been replicated and there are no approved treatments for the core social and behavioural symptoms of ASD.

The timely identification of efficacious treatments is partially dependent on the recruitment of homogeneous subgroups that are likely to benefit from the treatment (Ousley & Cermak, 2014). Additionally, a thorough analysis of moderators of treatment response in randomized controlled trials (RCTs) informs inclusion and exclusion criteria in subsequent trials (Kraemer *et al.* 2002). Moderators of treatment response in a large RCT examining the safety and efficacy of risperidone in children with autism and serious behaviouural problems have been thoroughly investigated and reported (Arnold *et al.* 2010); however, this analysis occurred as exploratory investigations after the original study was reported (McCracken *et al.* 2002). RCTs of oral interventions in ASDs appear to focus primarily on efficacy and safety with the reporting of moderators of treatment response generally not considered.

The objective of this review is to use meta-analytic techniques to identify clinical characteristics and trial design features that are effect modifiers, or 'moderators', of treatment response in RCTs of oral treatments for children with ASDs. Additionally, a systematic review of trial design features and baseline clinical, functional and demographic characteristics of participants will be used to qualitatively assess participant characterization at baseline to inform trial design and the identification of baseline characteristics important in paediatric ASD RCTs.

Method

Data sources and study selection

Relevant articles were identified through an electronic literature search of MEDLINE, EMBASE and PsycINFO conducted in December 2015 using the following keywords: autism, Aspergers, pervasive developmental disorder and randomized controlled trial. An example of the search strategy is included in online Supplementary Fig. S1. The Cochrane Database of Systemic Reviews and the clinicaltrials.gov database were then searched for additional trials not identified in the initial search. Duplicates and articles published before 1990 were removed to ensure diagnoses were established based on criteria in the Diagnostic and Statistical Manual for Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association, 1994) or a more recent DSM version. A reviewer (A.M.) screened titles and abstracts of articles from the search strategy for eligible studies. The full text of remaining studies identified as meeting inclusion criteria described below were then assessed for eligibility independently by two researchers (A.M. and M.M.D.) and agreement reached on the inclusion of each study. The reference lists of eligible studies were then searched for studies meeting inclusion criteria. A flowchart detailing the stages of the assessment of studies was constructed according to preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (Fig. 1) (Moher *et al.* 2009).

Types of studies meeting eligibility criteria included published, peer-reviewed journal articles in the English language reporting results from double-blind RCTs comparing treatment response between active agent and placebo, using either parallel or crossover designs with at least 10 participants per treatment arm or 20 participants overall. Data were extracted from the first phase of crossover studies due to the risk of carry-over effect (Elbourne *et al.* 2002). Participants were aged up to 20 years old and diagnosed with ASD, Asperger's syndrome or pervasive developmental disorder according to the DSM-IV or DSM-5, or the International Classification of Diseases-10, or using standardized diagnostic instruments, such as the Autism Diagnostic Observation Schedule, the Autism Diagnostic Interview Revised, or the Childhood Autism Rating Scale (CARS). Treatments were orally ingested, compared against placebo and taken for at

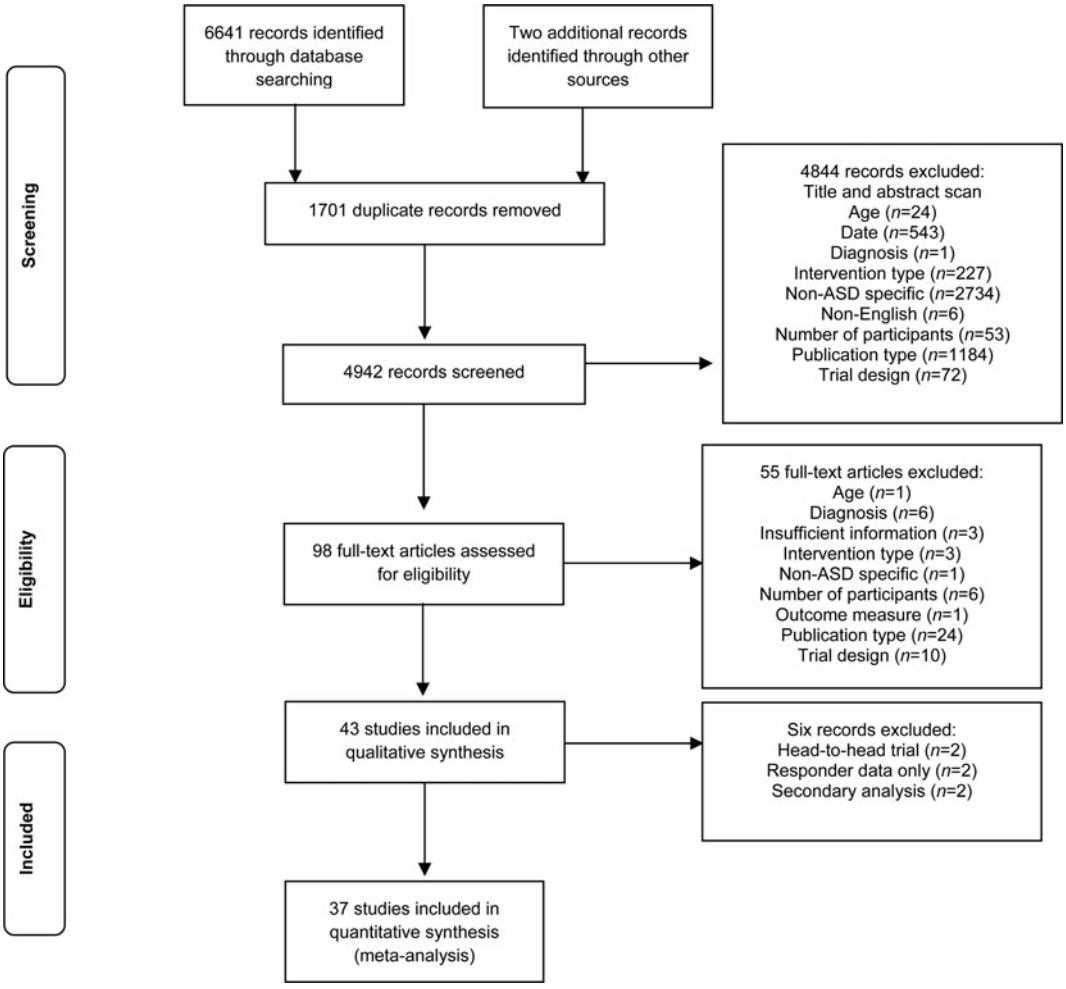

Fig. 1. Flow of information through the stages of the meta-analysis, including reasons for excluding full-text articles. ASD, Autism spectrum disorder.

least 14 days. Data reported using standardized instruments were included.

Data extraction

Continuous data reported for standardized outcome measures assessing treatment response in active intervention and placebo groups were extracted into an Excel spreadsheet. Data were independently extracted from each study by at least two researchers (A.M., M.M.D. and K. Szolusha) to ensure accuracy. If medians were reported, a mean was estimated using previously published methodology (Hozo *et al.* 2005). Authors of studies published in the previous 10 years were contacted to request raw data when data were not reported in the required format.

Data for potential moderators of response to treatment were extracted for the following participant characteristics and trial design features: number and age of participants, gender distribution, baseline severity, intellectual functioning, name and type of intervention, whether the treatment was a pharmacological agent and an adjunctive treatment (where adjunctive implies a combination of pharmacological agent approved for use in ASD with another oral treatment not approved for treatment in ASD), outcome measure, whether the outcome measure was designated as primary or secondary, the type of rater (independent evaluators or caregiver, including parents), geographical location of the trial, number of study sites, trial duration, frequency of contact for assessment purposes (telephone call or clinic visit), study quality, year of publication and trial sponsor details. All reported demographic and baseline clinical and functional characteristics not previously described were also extracted for each individual trial to facilitate a qualitative assessment of participant characterization at baseline.

A quality assessment to investigate the risk of bias of included RCTs was conducted by two independent researchers (A.M. and M.M.D.). The Cochrane Collaboration's tool

for the assessment of risk of bias (Higgins & Green, 2011) required an assessment of the risk of bias associated with specific features, including sequence generation, allocation concealment, blinding and outcome data reporting, as 'low risk', 'high risk' or 'unclear risk'. The same procedure was used to evaluate methodological quality of each RCT included in the meta-analysis. The Jadad Scale (Jadad *et al.* 1996) included three items directly related to the control of bias and eight items related to study design and features and had a maximum possible score of 13.

Data analysis

The quantitative analysis was based on the difference between active and placebo treatment groups using mean and standard deviation data at baseline and end point, or as treatment difference, for each included continuous outcome measure. Effect sizes were calculated using Comprehensive Meta-Analysis version 3 (Biostat, Inc., USA). The calculated effect size, Hedges' g, was the summary statistic and represented the difference between response to treatment in the active group and response to treatment in the placebo group, taking into account the post-score standard deviation. The inverse of study variance was used to weight each study. Hedges' g of 0.2, 0.5 and 0.8 were evaluated as small, moderate and large, respectively (Cohen, 2013). For the primary analysis, when a study reported baseline and end-point data for more than one outcome measure or for multiple subscales of an outcome measure, a single effect estimate based on mean effect size and variance across outcome measures or subscales was calculated (Borenstein *et al.* 2009). A random-effects model was used in all analyses as heterogeneity amongst the studies was expected. Subgroup analyses were conducted using mixed-effect models to investigate potential moderators of response to treatment. Studies within subgroups were combined using a random-effects model and differences between subgroups were assessed using a fixed-effects model (Borenstein & Higgins, 2013). Meta-regression was used to test whether there was a significant relationship between continuous variables and effect size.

A forest plot was prepared as a graphical illustration of effect size and associated 95% confidence interval (CI) for each included trial, with an overall effect size representing a summary measure of treatment response for all included trials. The I^2 statistic was used to assess between-study heterogeneity across studies (i.e. the proportion of variance across studies out of total observed variance), with values of 25, 50 and 75% implying small, moderate and high levels of heterogeneity, respectively (Higgins *et al.* 2003). Sensitivity analyses were conducted by removing outliers from analyses to determine if the subsequent effect size was significant. The Cochrane's Q statistic was used to test between-subgroup heterogeneity (Higgins & Green, 2011). Small study effect resulting from publication bias, insufficient reporting of outcomes, selective inclusion of study participants or other sources were assessed by visually inspecting funnel plots of Hedges' g against standard error (Sterne *et al.* 2011) and tested using Egger's test of the intercepts (Egger *et al.* 1997).

To facilitate a qualitative assessment of whether the reporting of demographic and baseline clinical and functional characteristics allowed for a thorough analysis of potential moderators, or effect modifiers, of treatment response, variables relating to demographics and baseline clinical and functional characteristics of randomized participants were collated in a graphical representation of the frequency of reporting of each variable across all qualifying studies.

Results

The screening and eligibility procedure resulted in 43 articles reporting on 41 RCTs in paediatric ASD being included in the qualitative synthesis and 37 studies included in the quantitative synthesis (Fig. 1). Two studies reported on secondary outcome measures not included in the original published article (McDougle *et al.* 2005; Harfterkamp *et al.* 2014). Two head-to-head trials were not included in the quantitative synthesis (Miral *et al.* 2008; Ghanizadeh *et al.* 2014) and outcome was reported as responder data in two studies (Handen *et al.* 2009; Voigt *et al.* 2014). There were 1997 participants (n active treatment = 1003, n placebo = 994). The proportion of total participants who were female was 19%, and the age range of the study population was 3–20 years. Other participant characteristics, trial design features and a measure of study quality are reported in online Supplementary Table S1. Using the Cochrane Risk of Bias Tool, 19 out of 37 studies were assessed as 'low risk' in all categories (Higgins & Green, 2011). All participants completed the trials in eight studies, and a further 17 studies used intention-to-treat analysis. A risk assessment for each study across every domain is presented in online Supplementary Table S2 and the proportion of studies within each risk level is shown in online Supplementary Fig. S2.

Baseline clinical, functional and demographic characteristics

Data relating to the baseline characteristics of study participants in the 41 included RCTs were itemized (online Supplementary Table S3). The numbers of studies including data relating to each of these variables are summarized in Fig. 2. All studies reported age and gender distributions. Ethnicity was reported in 41% of trials.

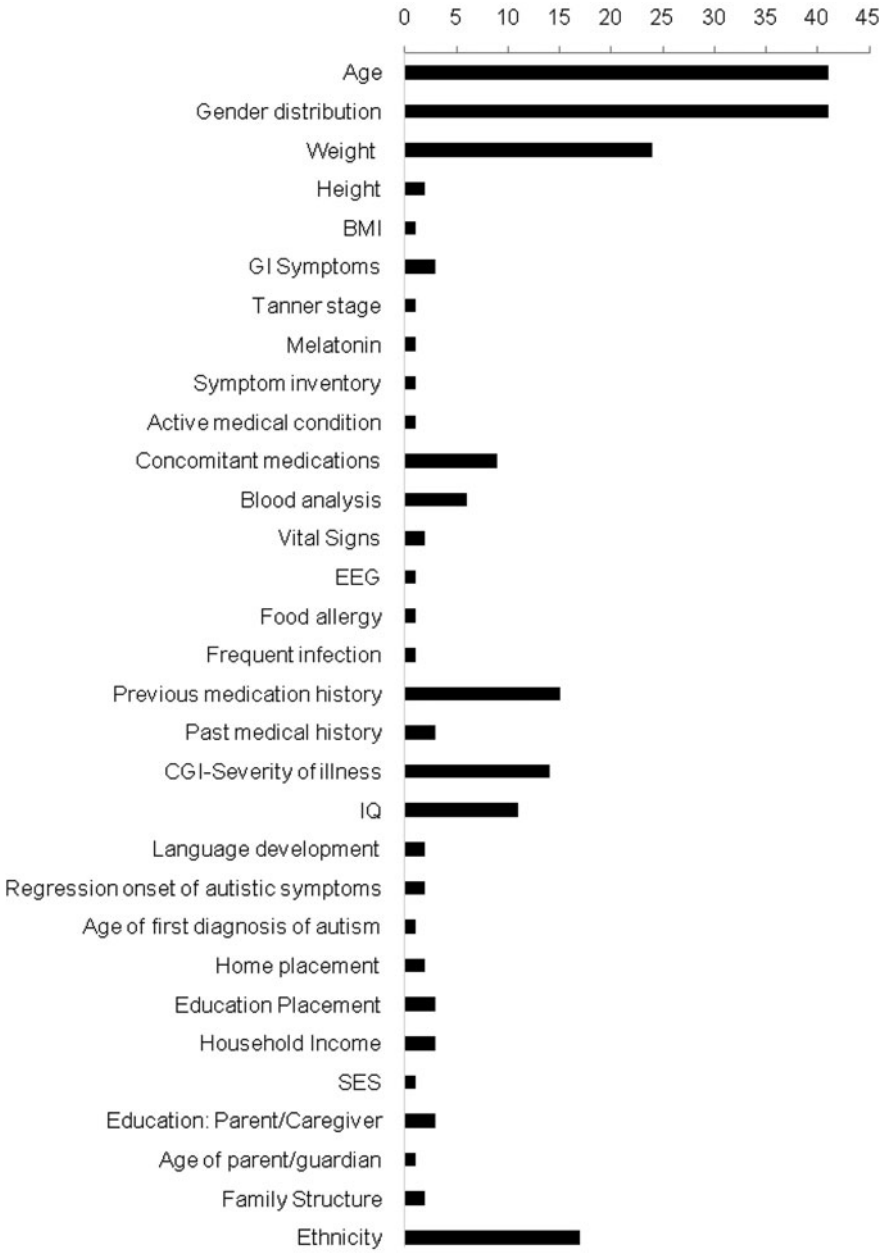

Fig. 2. Baseline characteristics for participants included in randomized controlled trials in paediatric autism. BMI, Body mass index; GI, gastrointestinal; EEG, electroencephalogram; CGI, Clinical Global Impression; IQ, intelligence quotient; SES, socio-economic status.

Baseline severity, measured using the Clinical Global Impression-severity (CGI-S) scale of illness, was reported for participants in 14 trials and a measure of intellectual functioning was included in 11 trials. Previous medication use was reported for participants in 15 trials and concomitant medications were reported for participants in nine trials. Past medical history and gastrointestinal symptoms were reported in two trials. A symptom inventory was reported in one trial. Regression onset of autistic symptoms was reported in two trials and age of first diagnosis in one trial. Diagnosis was confirmed for participants in 34 RCTs; however, specific diagnostic categories were noted in only 12 studies. Demographic variables such as caregiver education and socio-economic status were not routinely reported. Reporting of other functional variables, including level of language development, was generally limited.

Treatment response

Effect size was calculated for the overall response to treatments for all included studies ($k = 37$, $g = 0.63$,

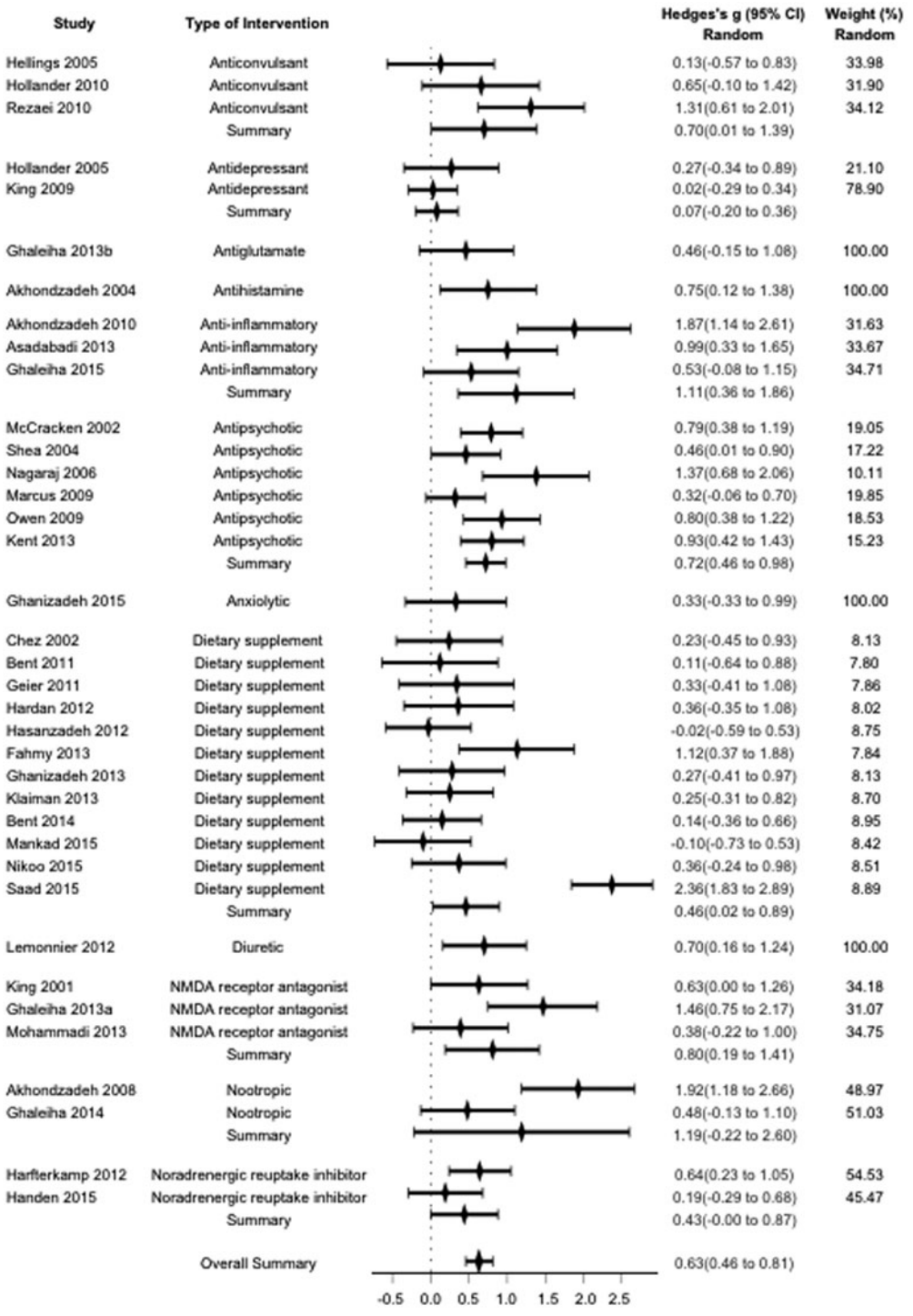

Fig. 3. Forest plot for treatment response in randomized controlled trials in paediatric autism grouped by class of intervention. CI, Confidence interval; NMDA, *N*-methyl-D-aspartate.

95% CI 0.46–0.81, $p < 0.001$, Fig. 3. See the online Supplementary material for study references). The level of between-study heterogeneity was moderate ($I^2 = 71.95\%$), implying about 28% of the heterogeneity was due to random error (i.e. variance within studies). A funnel plot did not show significant asymmetry (Egger's intercept = 2.00, $p = 0.13$). In a sensitivity analysis, one notable outlier study was removed (Saad *et al.* 2015), resulting in a reduction of heterogeneity ($I^2 = 58.22\%$). Effect size remained moderate ($k = 36$, $g = 0.57$, 95% CI 0.42–0.72, $p < 0.001$). Of the trials, 10 showed large effect sizes, eight trials showed moderate

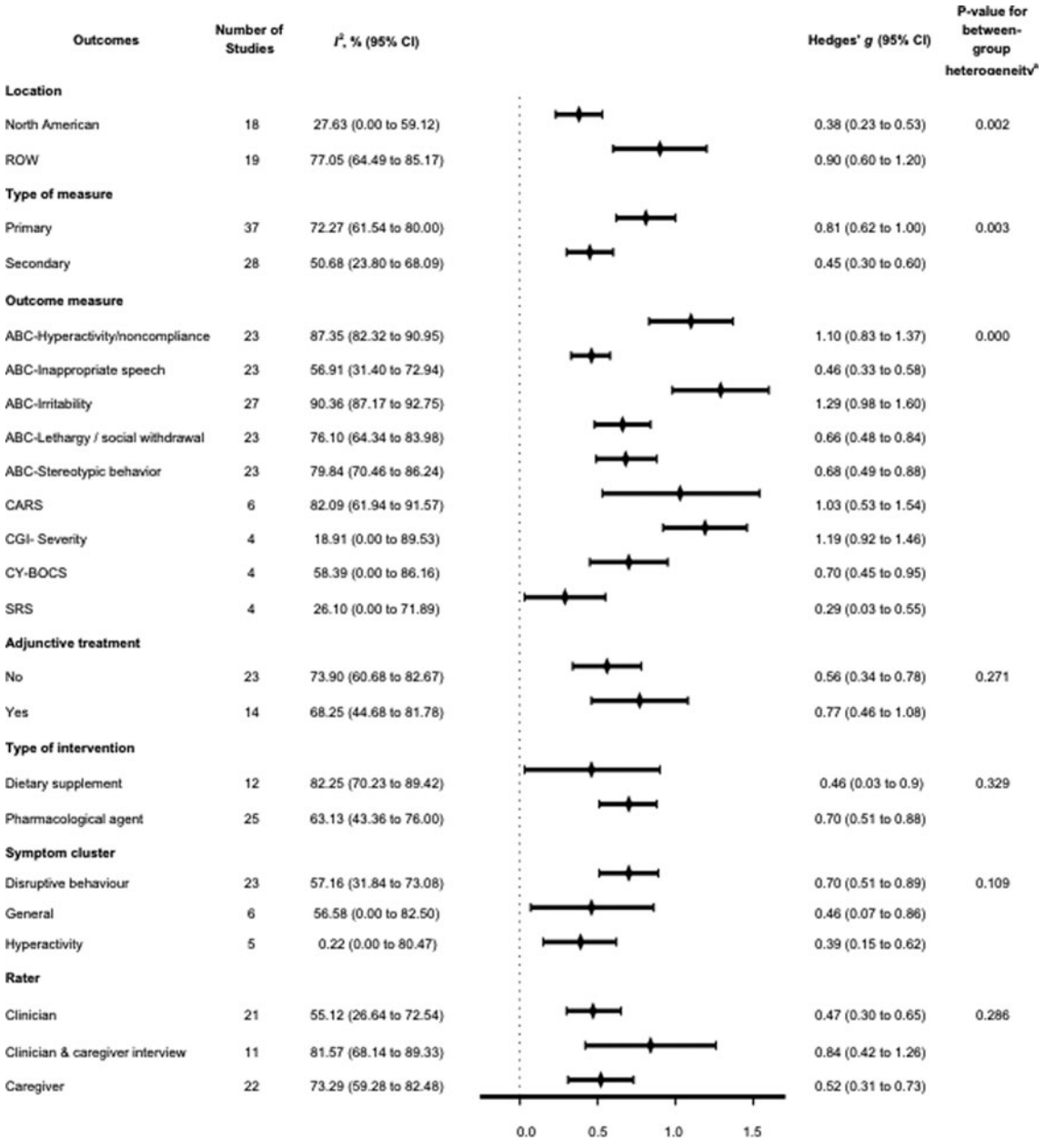

Fig. 4. Subgroup analyses of moderators of treatment response in randomized controlled trials in paediatric autism spectrum disorder. CI, Confidence interval; ROW, rest of world; ABC, Aberrant Behavior Checklist; CARS, Childhood Autism Rating Scale; CGI, Clinical Global Impression; CY-BOCS, Children's Yale–Brown Obsessive–Compulsive Scale; SRS, Social Responsiveness Scale. [a] *Q* test for between-group heterogeneity, mixed-effects model.

to large effect sizes, 13 trials showed small to moderate effect size and five trials resulted in a small effect size. Two trials showed no improvement in the target of the intervention. The results of 17 trials were significant (Fig. 3).

Moderators of response to treatment

Subgroup analyses were used to investigate potential categorical moderators of treatment response (Fig. 4). There was a significant difference for outcome measure [$Q = 60.90$, degrees of freedom (df) = 8, $p < 0.001$; Fig. 4], with the largest effect size for outcome measured using the Aberrant Behavior Checklist-irritability (ABC-I) subscale ($g = 1.29$). The effect size for primary outcome measures was significantly greater than secondary outcome measures ($Q = 8.63$, df = 1, $p = 0.003$; Fig. 4). A significant effect for geographical location of the RCT was found ($Q = 16.65$, df = 6, $p = 0.01$). Given the location of trials (online Supplementary Table S1), further

analysis was undertaken comparing the effect for trials located in North America and trials that were located in Europe and the Middle East. A significant effect for geographical location remained ($Q = 9.37$, df = 1, $p = 0.002$; Fig. 4), with trials from Europe and the Middle East reporting larger effect sizes than trials located in North America. The rater of the outcome measure, the targeted symptom cluster, adjunctive treatments, and whether the intervention was a drug or a dietary supplement were not effect modifiers of response to treatment (Fig. 4). The type of intervention was further investigated with subgroup analysis of all classes of pharmacological agents (i.e. excluding dietary supplements) not identifying a significant effect modifier of treatment response ($Q = 17.40$, df = 10, $p = 0.07$, Fig. 3).

A random-effects meta-regression analysis of continuous variables was completed using the Knapp–Hartung method (Knapp & Hartung, 2003). Year of publication, number of study sites and contact visits, trial duration, the number and age of trial participants, gender distribution and study quality (as measured by the Jadad Score) were not significant moderators of response to treatment (online Supplementary Table S4).

Discussion

This systematic review and meta-analysis is the most comprehensive synthesis of baseline variables and data assessing moderators of treatment response in RCTs in paediatric ASD. The systematic review component combined baseline demographics and clinical and functional data reported in paediatric RCTs in ASDs. The meta-analysis identified three effect modifiers: the measure used to assess outcome, the primary or secondary outcome status of the efficacy measure and an effect for geographical region. The main aim of this review was to evaluate the state of the evidence for a variety of the potential interventions in ASD. Thus, the meta-analysis was designed to investigate different sources of heterogeneity across studies, with a treatment class being a critical moderator.

Large effect sizes for outcome measures, including the CGI-S scale, CARS and the hyperactivity and irritability subscales of the ABC, may reflect that these measures are capturing a broader interpretation of response to treatment than other measures that are more specific to certain features, such as communication and social impairments and repetitive behaviours. The effect for primary outcome may indicate that the measure is detecting improvements more relevant to the clinical targets of the interventions than the secondary outcome assessments.

Significant differences in treatment effect size were also found based on the location of the RCTs, with trials located in North America reporting a small to moderate effect size overall and trials conducted in Europe and the Middle East reporting a large effect size overall. This variation of effect modification across geographical locations raises the question of whether the results from RCTs investigating treatment efficacy in ASD can be extrapolated across geographical regions. Trial location has previously been identified as a predictor of placebo response in RCTs in paediatric ASD (Masi *et al.* 2015) and an individual patient data meta-analysis from trials of antipsychotics in the treatment of schizophrenia reported smaller effect sizes in North American patients compared with other regions (Mattila *et al.* 2014). Furthermore, pharmacological treatments in bipolar disorder also highlighted the possibility of differential efficacy across geographical regions, with studies from the USA reporting lower effect sizes, an effect which could not be explained by differences in baseline characteristics (Welten *et al.* 2015). Participants in the RCTs conducted in North America were on average older than the participants in the European and Middle Eastern trials. However, given that the population is paediatric, it is unlikely that the difference in age renders North American participants less amenable to therapeutic change. The participants in North American trials may have been more ethnically diverse; however, limited data prevented an analysis of whether outcome was modified by ethnicity. Variation in treatment compliance may contribute to the difference in treatment response between the regions, an explanation previously suggested following the identification of a similar effect in major depressive disorder (Thase *et al.* 2016). Recruitment settings for a number of the trials conducted in the Middle East were out-patient clinics of psychiatric hospitals, whereas recruitment for trials conducted in North America were primarily community settings, possibly reflecting differing healthcare and clinical practices between the regions. However, it is possible that more severely affected children are recruited in a hospital setting. Contextual analysis of the different health systems, e.g. whether those in North America were receiving more behavioural therapy thus enhancing the change in the placebo group and masking differences, could provide some clarity around the factors resulting in this geographical variability in efficacy. This is beyond the scope of this review.

At least 31 different baseline characteristics were reported across included studies, with substantial variability in the number of characteristics and type of data included at baseline in individual studies and the incorporation of biological markers or correlates in only six studies (Fig. 2, online Supplementary Table S3). The limited reporting of most of these variables prevented an

investigation of whether the variable moderated response to treatment. There was a notable lack of consistency in the reporting of baseline symptom severity, which has previously been identified as a moderator of treatment response in an RCT investigating the efficacy of risperidone in paediatric ASD (Arnold *et al.* 2010). Baseline measures of severity were noted as completed or referred to as inclusion criteria but not reported as a baseline characteristic. The ABC-I subscale was the most frequently reported outcome measure at baseline and end point (27 trials). However, the ABC-I was noted as the baseline measure of symptom severity in only 13 trials. CGI-S, recommended as a measure that should be included in all ASD clinical trials (Aman *et al.* 2004), was reported in only 13 trials. Reporting of a specific measure of baseline symptom severity should be included in all ASD RCTs, using a level of detail that facilitates analysis of the relationship between baseline severity and treatment efficacy similar to that undertaken in trials of antidepressants (Kirsch *et al.* 2008). This will establish the efficacy of intervention across the range of initial symptom severity. Intellectual functioning was also inconsistently reported and often not as a full-scale measure. Given that the rates of co-occurrence of intellectual disability and ASD have been reported at between 50 and 75% (Charman *et al.* 2011), and that severe ASD increases the risk of co-occurring intellectual disability (Vivanti *et al.* 2013), it is imperative that characterization of participants in RCTs includes an assessment of intellectual functioning. In addition to facilitating an assessment of whether intellectual functioning is an effect modifier of treatment response, it would also enable an assessment of whether intellectual functioning represents a barrier to inclusion in the RCT; this is a major concern, as previous findings have highlighted the exclusion of persons with intellectual disability in medical research trials (Feldman *et al.* 2014). Ethnicity was reported in approximately one-third of the included trials, which raises concern about generalization of results across ethnicities and may mask racial inconsistencies in the recognition and diagnosis of ASD in children of some ethnicities (Mandell *et al.* 2009). There was a notable lack of reporting of co-occurring symptoms and co-morbidities at baseline, with only one study reporting a symptom inventory at baseline (Nagaraj *et al.* 2006). Without consideration of these characteristics, it is not possible to investigate whether co-morbidity moderated treatment outcome. Additionally, the role of concurrent behaviour therapy, such as an intensive behavioural intervention (IBI), e.g. applied behavioural analysis or parent training, could not be investigated, as only one trial identified whether participants were undertaking IBI (Mankad *et al.* 2015) and it was generally not noted as an exclusion criterion. Given that behavioural interventions have long been considered the main evidence-based treatment for the amelioration of behavioural symptoms associated with ASD (Eldevik *et al.* 2009), and previous trials of combination therapies have shown greater reduction in behaviour problems (Aman *et al.* 2009; Frazier *et al.* 2010), the role of combined behavioural interventions in treatment outcome should be investigated.

Clinical heterogeneity poses significant challenges to the identification of efficacious treatments for ASDs. An important focus for ASD research has been to minimize heterogeneity of treatment subgroups as a means of improving characterization of ASD and the subsequent identification of treatment responders (Ousley & Cermak, 2014). Reducing the heterogeneity of treatment subgroups in clinical trials can only be possible if participants are thoroughly characterized. Based on the results of this systematic review and meta-analysis and previous guidelines for participant characterization in clinical trials (Scahill & Lord, 2004), we highlight the importance of including a baseline measure of symptom severity and intellectual functioning, and demographic details such as ethnicity. Additionally, given the prevalence in ASDs, participant characterization should include mental health and medical co-morbidities. We note that despite National Institute of Mental Health calls for novel markers of treatment response in ASDs, few clinical trials have yet incorporated biological markers, such as basic laboratory sampling to facilitate genetic investigations and metabolic screening, and biological correlates, such as imaging, to facilitate neurobiological investigations and the identification of involved circuitry and associated mechanisms of action. The designs of two ongoing large multi-site randomized trials for antidepressants highlight the comprehensive nature of assessments required to establish biological signatures and personalization of treatment, and represent models for future research investigating treatments for symptoms of ASDs (Williams *et al.* 2011; Trivedi *et al.* 2016).

Conclusion

This systematic review and meta-analysis highlights issues with respect to the conduct of RCTs in paediatric ASD. Geographical variations in trial outcome are interesting and require further investigation and contextual analysis. Our qualitative assessment of baseline characterization of participants in RCTs in paediatric autism indicates that narrow reporting of baseline characteristics is preventing a thorough assessment of moderators of treatment response and limiting the benefits of trial results guiding future trials on the selection of informed inclusion and exclusion criteria. In a population that is highly heterogeneous, the identification of efficacious treatments and the generalization of trial

results are highly dependent on thorough baseline characterizations. An assessment of intellectual functioning and reporting of a specific measure of baseline severity should be included in all trials to facilitate an assessment of effect modification. Furthermore, investigations of biological markers and correlates should also be included in all trials to facilitate the identification of biomarkers that could be used to stratify and select participants with ASD who may benefit from a specific intervention or as a treatment-sensitive objective response measure.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0033291716003457

Acknowledgements

The authors thank Professor Salvador-Carulla for his comments on the manuscript, Harry Hallock for assisting with analysis and presentation of data and Kerribeth Szolusha for assisting with data extraction and the evaluation of study quality. The authors acknowledge a National Health and Medical Research Council (NHMRC) Dementia Research Development Fellowship (ID 1108520) to A.L. and a NHMRC career development fellowship (APP1061922) and project grant (1043664) to A.J.G.

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Declaration of Interest

I.B.H. is a Commissioner in Australia's new National Mental Health Commission from 2012. He was a director of headspace: the national youth mental health foundation until January 2012. He was previously the chief executive officer (till 2003) and clinical adviser (till 2006) of beyondblue, an Australian National Depression Initiative.

He is the Co-Director, Health and Policy at the Brain and Mind Centre which operates two early-intervention youth services under contract to headspace. He has led a range of community-based and pharmaceutical industry-supported depression awareness and education and training programmes. He has led projects for health professionals and the community supported by governmental, community agency and pharmaceutical industry partners (Wyeth, Eli Lily, Servier, Pfizer, AstraZeneca) for the identification and management of depression and anxiety. He has received honoraria for presentations of his own work at educational seminars supported by a number of non-government organizations and the pharmaceutical industry (including Servier, Pfizer, AstraZeneca, and Eli Lilly).

He is a member of the Medical Advisory Panel for Medibank Private and also a Board Member of Psychosis Australia Trust. He leads an investigator-initiated study of the effects of agomelatine on circadian parameters (supported in part by Servier) and has participated in a multicentre clinical trial of the effects of agomelatine on sleep architecture in depression and a Servier-supported study of major depression and sleep disturbance in primary care settings.

References

Aman MG, McDougle CJ, Scahill L, Handen B, Arnold L, Johnson C, Stigler KA, Bearss K, Butter E, Swiezy NB, Sukhodolsky DD, Ramadan Y, Pozdol S, Nikolov R, Lecavalier L, Kohn AE, Koenig K, Hollway JA, Korzekwa P, Gavaletz A, Mulick JA, Hall KL, Dziura J, Ritz L, Trollinger S, Yu S, Vitiello B, Wagner A (2009). Medication and parent training in children with pervasive developmental disorders and serious behavior problems: results from a randomized clinical trial. *Journal of the American Academy of Child and Adolescent Psychiatry* **48**, 1143–1154.

Aman MG, Novotny S, Samango-Sprouse C, Lecavalier L, Leonard E, Gadow KD, King BH, Pearson DA, Gernsbacher MA, Chez M (2004). Outcome measures for clinical drug trials in autism. *CNS Spectrums* **9**, 36–47.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn. American Psychiatric Association: Washington, DC.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Association: Washington, DC.

Arnold L, Farmer C, Kraemer HC, Davies M, Witwer A, Chuang S, DiSilvestro R, McDougle CJ, McCracken J, Vitiello B, Aman MG, Scahill L, Posey DJ, Swiezy NB (2010). Moderators, mediators, and other predictors of risperidone response in children with autistic disorder and irritability. *Journal of Child and Adolescent Psychopharmacology* **20**, 83–93.

Borenstein M, Hedges LV, Higgins JP, Rothstein HR (2009). *Introduction to Meta-Analysis*. John Wiley & Sons, Ltd: Hoboken, NJ.

Borenstein M, Higgins JPT (2013). Meta-analysis and subgroups. *Prevention Science* **14**, 134–143.

Carrasco M, Volkmar FR, Bloch MH (2012). Pharmacologic treatment of repetitive behaviors in autism spectrum disorders: evidence of publication bias. *Pediatrics* **129**, e1301–e1310.

Chaidez V, Hansen RL, Hertz-Picciotto I (2014). Gastrointestinal problems in children with autism, developmental delays or typical development. *Journal of Autism and Developmental Disorders* **44**, 1117–1127.

Charman T, Pickles A, Simonoff E, Chandler S, Loucas T, Baird G (2011). IQ in children with autism spectrum

disorders: data from the Special Needs and Autism Project (SNAP). *Psychological Medicine* **41**, 619–627.

Ching H, Pringsheim T (2012). Aripiprazole for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, issue **5**, CD009043.

Cohen J (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge Academic: London.

Cohen S, Conduit R, Lockley SW, Rajaratnam SMW, Cornish KM (2014). The relationship between sleep and behavior in autism spectrum disorder (ASD): a review. *Journal of Neurodevelopmental Disorders* **6**, 44.

De Vries M, Geurts H (2015). Influence of autism traits and executive functioning on quality of life in children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders* **45**, 2734–2743.

Egger M, Davey Smith G, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634.

El Achkar CM, Spence SJ (2015). Clinical characteristics of children and young adults with co-occurring autism spectrum disorder and epilepsy. *Epilepsy and Behavior* **47**, 183–190.

Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A (2002). Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* **31**, 140–149.

Eldevik S, Hastings RP, Hughes JC, Jahr E, Eikeseth S, Cross S (2009). Meta-analysis of early intensive behavioral intervention for children with autism. *Journal of Clinical Child and Adolescent Psychology* **38**, 439–450.

Feldman MA, Bosett J, Collet C, Burnham-Riosa P (2014). Where are persons with intellectual disabilities in medical research? A survey of published clinical trials. *Journal of Intellectual Disability Research* **58**, 800–809.

Frazier TW, Youngstrom EA, Haycook T, Sinoff A, Dimitriou F, Knapp J, Sinclair L (2010). Effectiveness of medication combined with intensive behavioral intervention for reducing aggression in youth with autism spectrum disorder. *Journal of Child and Adolescent Psychopharmacology* **20**, 167–177.

Fung LK, Mahajan R, Nozzolillo A, Bernal P, Krasner A, Jo B, Coury D, Whitaker A, Veenstra-Vanderweele J, Hardan AY (2016). Pharmacologic treatment of severe irritability and problem behaviors in autism: a systematic review and meta-analysis. *Pediatrics* **137**, S124–S135.

Ghanizadeh A, Sahraeizadeh A, Berk M (2014). A head-to-head comparison of aripiprazole and risperidone for safety and treating autistic disorders, a randomized double blind clinical trial. *Child Psychiatry and Human Development* **45**, 185–192.

Guastella AJ, Hickie IB (2016). Oxytocin treatment, circuitry, and autism: a critical review of the literature placing oxytocin into the autism context. *Biological Psychiatry* **79**, 234–242.

Handen BL, Melmed RD, Hansen RL, Aman MG, Burnham DL, Bruss JB, McDougle CJ (2009). A double-blind, placebo-controlled trial of oral human immunoglobulin for gastrointestinal dysfunction in children with autistic disorder. *Journal of Autism and Developmental Disorders* **39**, 796–805.

Harfterkamp M, Buitelaar JK, Minderaa RB, Van De Loo-Neus G, Van Der Gaag R-J, Hoekstra PJ (2014). Atomoxetine in autism spectrum disorder: no effects on social functioning; some beneficial effects on stereotyped behaviors, inappropriate speech, and fear of change. *Journal of Child and Adolescent Psychopharmacology* **24**, 481–485.

Higgins JPT, Green S (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011 (http://www.cochrane-handbook.org/).

Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560.

Hozo SP, Djulbegovic B, Hozo I (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology* **5**, 13.

Huguet G, Ey E, Bourgeron T (2013). The genetic landscapes of autism spectrum disorders. *Annual Review of Genomics and Human Genetics* **14**, 191–213.

Hurwitz R, Blackmore R, Hazell P, Williams K, Woolfenden S (2012). Tricyclic antidepressants for autism spectrum disorders (ASD) in children and adolescents. *Cochrane Database of Systematic Reviews*, issue **3**, CD008372.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, McQuay HJ (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials* **17**, 1–12.

Jesner OS, Aref-Adib M, Coren E (2007). Risperidone for autism spectrum disorder. *Cochrane Database of Systematic Reviews*, issue 1, CD005040.

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* **5**, e45.

Knapp G, Hartung J (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine* **22**, 2693–2710.

Kraemer HC, Wilson GT, Fairburn CG, Agras WS (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* **59**, 877–883.

Mandell DS, Wiggins LD, Carpenter LA, Daniels J, Diguiseppi C, Durkin MS, Giarelli E, Morrier MJ, Nicholas JS, Pinto-Martin JA, Shattuck PT, Thomas KC, Yeargin-Allsopp M, Kirby RS (2009). Racial/ethnic disparities in the identification of children with autism spectrum disorders. *American Journal of Public Health* **99**, 493–498.

Mankad D, Dupuis A, Smile S, Roberts W, Brian J, Lui T, Genore L, Zaghloul D, Iaboni A, Marcon PM, Anagnostou E (2015). A randomized, placebo controlled trial of omega-3 fatty acids in the treatment of young children with autism. *Molecular Autism* **6**, 18.

Masi A, Lampit A, Glozier N, Hickie IB, Guastella AJ (2015). Predictors of placebo response in pharmacological and dietary supplement treatment trials in pediatric autism spectrum disorder: a meta-analysis. *Translational Psychiatry* **5**, e640.

Matson JL, Shoemaker M (2009). Intellectual disability and its relationship to autism spectrum disorders. *Research in Developmental Disabilities* **30**, 1107–1114.

Mattila T, Wohlfarth T, Koeter M, Storosum J, Van Den Brink W, De Haan L, Leufkens H, Denys D (2014). Geographic variation in efficacy of atypical antipsychotics for the acute treatment of schizophrenia – an individual patient data meta-analysis. *European Neuropsychopharmacology* **24**, 1067–1077.

McCracken JT, McGough J, Shah B, Cronin P, Hong D, Aman MG, Arnold E, Lindsay R, Nash P, Hollway J, McDougle CJ, Posey D, Swiezy N, Kohn A, Scahill L, Martin A, Koenig K, Volkmar F, Carroll D, Lancor A, Tierney E, Ghuman J, Gonzalez NM, Grados M, Vitiello B, Ritz L, Davies M, Robinson J, McMahon D; Research Units on Pediatric Psychopharmacology Autism Network (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine* **347**, 314–321.

McDougle CJ, Scahill L, Aman MG, McCracken JT, Tierney E, Davies M, Arnold LE, Posey DJ, Martin A, Ghuman JK, Shah B, Chuang SZ, Swiezy NB, Gonzalez NM, Hollway J, Koenig K, McGough JJ, Ritz L, 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. *American Journal of Psychiatry* **162**, 1142–1148.

McPheeters ML, Warren Z, Sathe N, Bruzek JL, Krishnaswami S, Jerome RN, Veenstra-Vanderweele J (2011). A systematic review of medical treatments for children with autism spectrum disorders. *Pediatrics* **127**, e1312–e1321.

Miral S, Gencer O, Inal-Emiroglu FN, Baykara B, Baykara A, Dirik E (2008). Risperidone *versus* haloperidol in children and adolescents with AD: a randomized, controlled, double-blind trial. *European Child and Adolescent Psychiatry* **17**, 1–8.

Moher D, Liberati A, Tetzlaff J, Altman DG (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine* **151**, 264–269, w64.

Nagaraj R, Singhi P, Malhi P (2006). Risperidone in children with autism: randomized, placebo-controlled, double-blind study. *Journal of Child Neurology* **21**, 450–455.

Ousley O, Cermak T (2014). Autism spectrum disorder: defining dimensions and subgroups. *Current Developmental Disorders Reports* **1**, 20–28.

Perry A, Flanagan HE, Dunn Geier J, Freeman NL (2009). Brief report: the Vineland Adaptive Behavior Scales in young children with autism spectrum disorders at different cognitive levels. *Journal of Autism and Developmental Disorders* **39**, 1066–1078.

Persico AM, Napolioni V (2013). Autism genetics. *Behavioral Brain Research* **251**, 95–112.

Rossignol DA, Frye RE (2012). Mitochondrial dysfunction in autism spectrum disorders: a systematic review and meta-analysis. *Molecular Psychiatry* **17**, 290–314.

Saad K, Eltayeb AA, Mohamad IL, Al-Atram AA, Elserogy Y, Bjorklund G, El-Houfey AA, Nicholson B (2015). A randomized, placebo-controlled trial of digestive enzymes in children with autism spectrum disorders. *Clinical Psychopharmacology and Neuroscience* **13**, 188–193.

Scahill L, Lord C (2004). Subject selection and characterization in clinical trials in children with autism. *CNS Spectrums* **9**, 22–32.

Simonoff E, Pickles A, Charman T, Chandler S, Loucas T, Baird G (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of the American Academy of Child and Adolescent Psychiatry* **47**, 921–929.

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, Carpenter J, Rucker G, Harbord RM, Schmid CH, Tetzlaff J, Deeks JJ, Peters J, Macaskill P, Schwarzer G, Duval S, Altman DG, Moher D, Higgins JP (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* **343**, d4002.

Thase ME, Mahableshwarkar AR, Dragheim M, Loft H, Vieta E (2016). A meta-analysis of randomized, placebo-controlled trials of vortioxetine for the treatment of major depressive disorder in adults. *European Neuropsychopharmacology* **26**, 979–993.

Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, Oquendo MA, Bruder G, Pizzagalli D, Toups M, Cooper C, Adams P, Weyandt S, Morris DW, Grannemann BD, Ogden RT, Buckner R, McInnis M, Kraemer HC, Petkova E, Carmody TJ, Weissman MM (2016). Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care (EMBARC): rationale and design. *Journal of Psychiatric Research* **78**, 11–23.

Vivanti G, Barbaro J, Hudry K, Dissanayake C, Prior M (2013). Intellectual development in autism spectrum disorders: new insights from longitudinal studies. *Frontiers in Human Neuroscience* **7**, **354**.

Voigt RG, Mellon MW, Katusic SK, Weaver AL, Matern D, Mellon B, Jensen CL, Barbaresi WJ (2014). Dietary docosahexaenoic acid supplementation in children with autism. *Journal of Pediatric Gastroenterology and Nutrition* **58**, 715–722.

Welten CC, Koeter M, Wohlfarth T, Storosum J, Van Den Brink W, Gispen-De Wied C, Leufkens H, Denys D (2015). Efficacy of drug treatment for acute mania differs across geographic regions: an individual patient data meta-analysis of placebo-controlled studies. *Journal of Psychopharmacology* **29**, 923–932.

Williams K, Brignell A, Randall M, Silove N, Hazell P (2013). Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews*, issue **8**, CD004677.

Williams LM, Rush AJ, Koslow SH, Wisniewski SR, Cooper NJ, Nemeroff CB, Schatzberg AF, Gordon E (2011). International Study to Predict Optimized Treatment for Depression (iSPOT-D), a randomized clinical trial: rationale and protocol. *Trials* **12**, 4.