Outcome measures in intervention trials for adults with autism spectrum disorders; a systematic review of assessments of core autism features and associated emotional and behavioural problems

TRAOLACH S. BRUGHA,^{1*} LUCY DOOS,² ALTHEA TEMPIER,³ STEWART EINFELD⁴ & PATRICIA HOWLIN^{3,4*}

1 Department of Health Sciences, University of Leicester, Leicester General Hospital, Leicester, UK

2 Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, Edgbaston, Birmingham, UK

3 Department of Psychology, Institute of Psychiatry, King's College London, London, UK

4 Centre for Disability Research and Policy, Brain & Mind Research Institute, The University Of Sydney, Camperdown, NSW, Australia

Key words

autism, systematic review, scale evaluation, clinical trials, adult

Correspondence

Traolach S. Brugha, Department of Health Sciences, University of Leicester, Leicester General Hospital, Leicester, LE5 4PW, UK. Telephone (+44) 116-252-3211 Email: tsb@le.ac.uk

*TRAOLACH S. BRUGHA and PATRICIA HOWLIN are joint main authors.

Received 20 August 2014;
revised 17 December 2014;
accepted 20 February 2015

Abstract

A systematic review was conducted of outcome measures used in treatment trials for older adolescents and adults with autism spectrum disorders (ASDs). Of 818 titles only 30 articles (19 of which involved pharmacological treatments) were identified that met inclusion criteria (sample size > 5; mean age of group > 15 years; mean IQ > 30; ASD diagnosis confirmed; use of objective ASD outcome measures; focus on symptoms core to or typically associated with ASDs). Selected studies included randomized and placebo-controlled trials, retrospective assessment studies, case series and open label or case-control trials. Use of outcome measures varied with frequent use of non-standardized assessments, very little use of measures designed specifically for individuals with ASD or of instruments focusing on core ASD deficits, such as communication or social functioning. Most commonly used were the Clinical Global Impression (CGI) rating scale and the Yale–Brown Obsessive Compulsive Scale (Y-BOCS). The strengths or deficiencies of the outcome measures used were not systematically evaluated.

Although there are now many well controlled treatment trials for children with ASDs, adult intervention research is very limited. The lack of valid and reliable outcome measures for adults with ASDs compromises attempts at treatment evaluation. Copyright © 2015 John Wiley & Sons, Ltd.

Introduction

Autism spectrum disorders (ASDs) affect adults as frequently as children (Brugha *et al.*, 2011), often lead to life-long impairments (Howlin *et al.*, 2004, 2013), and are associated with high rates of behavioural and emotional problems. However, although there is considerable research focusing on the identification, assessment and treatment of children with ASD, adult research lags a long way behind. Intervention research is limited both in quantity and quality and, in particular, there are almost no high quality randomized control trials of psychosocial interventions for adults [Bishop-Fitzpatrick *et al.*, 2013; Howlin and Moss, 2012; Taylor *et al.*, 2012; National Institute for Health and Clinical Excellence (NICE), 2012; Shattuck *et al.*, 2012]. Valid, reliable and sensitive outcome measures, which are fundamental to the development of an evidence base for clinical effectiveness, are also lacking. Nevertheless, even within the child ASD intervention literature, commonly used outcome measures (such as changes in IQ, adaptive behaviour and language) are often unrelated to the main focus of treatment, are non-ASD specific, and/or rely on measures developed for individuals without intellectual or developmental disabilities. For example, trials of cognitive-behaviour therapy for children and adolescents with ASD have tended to rely on standardized measures of anxiety and mood that have yet to be evaluated in ASD (Ho *et al.*, 2014).

In this review, we consider two main groups of outcome measures relevant to adults with ASD. The first group focuses on core ASD symptoms, i.e. social – communication deficits and repetitive and stereotyped behaviours (American Psychiatric Association, 2013). Second are measures that focus on commonly co-occurring symptoms such as cognitive impairment, anxiety, self-injury or disruptive behaviours. We also included measures of quality of life, as this is generally rated as poor for adults with ASD. However, we excluded measures related to non-ASD specific outcomes (e.g. as related to employment or leisure interventions). Although these real world outcomes are clearly of enormous importance to adults with ASD, this would require a different, far more broad reaching, review. Interventions meeting inclusion criteria fell essentially into two groups: pharmacological and non-pharmacological. As these can differ contextually, scientifically and in terms of clinical and treatment outcomes, both types of studies are considered and discussed.

Problems related to intervention assessment in ASD

In intervention research for mental health disorders, change in the severity of core mental disorder symptoms

is frequently chosen as the primary outcome measure. For example, in the treatment of adult depression, one of the most widely used instruments is the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967). This makes use of clinician ratings of change in severity of depressed mood, with each item corresponding closely to the diagnostic criteria for major depressive disorder [World Health Organization (WHO), 1993; American Psychiatric Association, 1994, 2013]. However, there are no parallel measures for monitoring change in ASD. Although there are several widely used measures designed specifically to aid ASD diagnosis [e.g. Autism Diagnostic Interview – Revised (ADI-R), Le Couteur *et al.* (2003); Autism Diagnostic Observation Schedule (ADOS-G), Lord *et al.* (2000)] these are predominantly child-based. Moreover, despite the fact that they have been used to assess effectiveness in child-based interventions (e.g. Dawson *et al.*, 2010; Green *et al.*, 2010) they were not constructed to be sensitive to change in response to treatment.

Assessment of ASD-specific treatment effects in adults presents additional problems that are not found with interventions for non-ASD adults with mental disorders. For instance, the severity of ASD symptoms tends to be less in adulthood than in childhood (Howlin *et al.*, 2013); developmental changes in adulthood, too, are less marked. Therefore, the scope for change on standardized ASD diagnostic instruments may be more limited than in child ASD studies and may also be less than the changes in mental health symptoms observed in acute mental disorders such as depression in adults.

Secondly, in adult trials in mental disorder, assessment of treatment response often relies on participants' own descriptions (i.e. self report) of their experiences and emotions (Robins *et al.*, 1988; Wing *et al.*, 1990; Fitzpatrick *et al.*, 1998) rather than observer ratings. This reliance on self-report can give rise to problems for adults with ASD because of their difficulties in communication, abstract thought and interpretation of emotions. Thus, even among individuals with ASD of average or above intellectual ability, there may be a need to use a combination of self and informant measures of ASD. However, most well standardized informant-based assessments of ASD symptoms (e.g. ADI-R; Le Couteur *et al.*, 2003) are based on parental information and this can be difficult to obtain for older individuals (Ferriter *et al.*, 2001). Child-focussed ASD instruments such as the CARS (Childhood Autism Rating Scale; Schopler *et al.*, 1980), which includes an observational element, may be inappropriate for adults and although a module of the ADOS-G (Lord *et al.*, 2000) is available for adolescents and adults there are no data on its sensitivity to change in adult ASD intervention studies.

Standardized assessments of associated or comorbid (non-ASD) problems also present problems as these measures have rarely been evaluated with individuals with ASD. We do not know, for example, whether adults with ASD interpret the questions on standard self-report measures (e.g. for anxiety or depression) in the same way as adults without ASD, or whether their responses accurately reflect their emotional state. There can also be problems in distinguishing some associated problems (e.g. OCD, obsessive compulsive disorder) from core autism symptoms (e.g. repetitive thoughts and behaviours; Russell, *et al.*, 2005).

There are few studies of the relative validity or reliability of measurement of ASD-related symptoms from different informant sources such as parents, self-report or expert observation. Work comparing teachers' and parents' ratings of psychiatric symptoms in children with ASD indicated relatively poor agreement (e.g. correlations for Affective disorders and Anxiety 0.08 and 0.14, respectively; Kanne *et al.*, 2009). In adults, Bishop and Seltzer (2012) also found low correlations (0.18) between self-reported autism severity ratings on the Autism Spectrum Quotient (AQ; Baron-Cohen *et al.*, 2001) and parent ratings on the ADI-R (Le Couteur *et al.*, 2003).

In an attempt to identify more effective methods for measuring outcome in intervention trials for adults with ASDs, we conducted a systematic examination of the outcome methods used in interventions studies for adults with ASDs.

Methods

Search strategy and data sources

We used automated and manual search methods to identify relevant published reports on adult ASD interventions. Electronic databases included Cochrane Library, PubMed, MEDLINE; EMBASE and PsycINFO. Other online sources included The National Autistic Society (NAS) Library (National Autistic Society, 2012), guidelines, and ASD related associations (Research Autism, 2013). Relevant online journals were also searched. Bibliographies and reference lists of key articles were searched, field experts were contacted and key journals were hand searched. The search strategy was further expanded by examining the names of authors whose papers we had already retrieved; for each author we searched through all their publications to see if they had published other relevant papers of value. Unpublished manuscripts and book chapters were not included in the search.

Study selection

Searches were limited to articles in English published between 1960 and January 2013. Only articles that included adults or adolescents with autism or Asperger syndrome were included in our review. The following search terms were used: ASD, autism (autistic) spectrum (disorder), Asperger(s) disorder (syndrome), high functioning autism, Autism Spectrum Quotient, or autism screening questionnaire combined with "adult" or "adolescent" combined with "intervention", "treatment", "therapy", "clinical trial" or "screening". For inclusion in the review studies were required to meet the following criteria:

- (1) A focus on treating core ASD symptoms and associated conditions of ASD. Studies with a focus on specific issues that are not core to autism (e.g. access to employment or vocational training) were excluded, although we did include intervention studies related to general quality of life and studies that examined aspects related to autism [e.g. theory of mind improvements (Garcia-Villamizar and Hughes, 2007)].
- (2) Participants involved adults or adolescents (studies involving children were included if the mean age of the group was ≥ 15 years).
- (3) Evaluation of therapy compared with the same treatment at a different dose or intensity, an alternative intervention and/or placebo or usual care.
- (4) Incorporating at least one standardized or quantitative outcome measure of effectiveness associated with improvement of core/associated or secondary features of ASD.

Studies were excluded if:

- (1) They primarily concerned participants with severe/profound intellectual disability. (This decision was taken since the assessment of individuals with ASD and severe/profound intellectual impairments further complicates diagnostic and measurement issues.)
- (2) They provided no indication of the IQ level of participants.
- (3) Sample size was less than five participants.

Two authors independently extracted study data (LD and AT). If there were any discrepancies in selected papers between the two reviewers, a third reviewer (TSB) independently reviewed the papers and, through discussion, obtained a consensus. Independent ratings (by PH and AT) of the reasons for exclusion for the 416 rejected articles reached 99% agreement.

Article selection, data extraction and synthesis

References were exported to a bibliographic database and duplicates eliminated electronically and manually. When there were multiple reports, the study with the most recent and complete data was selected. The full study reports were examined when citation titles or abstracts did not contain sufficient detail. Data were extracted from each study in terms of study design, participant characteristics, type of intervention, the type and method of outcome measurements (ASD specific and other general outcomes) and results. We relied on the authors' own descriptions of outcome measures, and changes on these, to obtain information about treatment effects and, if reported, about the psychometric properties of the outcome measures in ASD populations. It was our initial intention to combine the data using classic methods of meta-analysis and report summary estimates with 95% confidence intervals. However, because the trials identified were clinically heterogeneous, meta-analysis was considered inappropriate. Instead the primary outcome data extracted from the original studies were assessed qualitatively and subject to narrative review.

Results

The search identified 818 potential articles. Based on the title, 416 abstracts (or full text in the case of no abstract) were selected and read. From these we discarded abstracts ($n = 283$) that clearly did not meet inclusion criteria; subsequently we read and rejected a further 133 full-text articles. Reasons for exclusion were: no information on age/IQ or size of sample ($n = 6$); profound intellectual disability ($n = 2$); mainly child participants ($n = 96$); conference abstract or commentary only ($n = 6$); literature review ($n = 41$); sample size < 5 ($n = 4$); not in English ($n = 3$); not focused on ASD or not an intervention study ($n = 258$; of these, 16 were neither a treatment trial nor about ASD; nine were about intervention but not about ASD; the remainder was about ASD but not treatment trials). Thirty articles, describing 29 intervention studies [two articles by Hollander *et al.* (2003, 2007) involved the same participants] met inclusion criteria. The earliest article was published in 1992.

Summaries of study design, sample, methods of measuring outcome and article numbers (1–30) are presented in Table 1. The studies included in the review comprised 10 blinded and three non-blinded randomized controlled trials, five placebo-controlled or case-control studies, six open label drug trials, two retrospective assessment studies and three case series.

Only 10 studies used non-pharmacological interventions (Table 1, study numbers: 6, 10, 11, 13–16, 20, 29, 30).

Pharmacological intervention studies mainly involved the use of antidepressant or antipsychotic types of medication, although three studies assessed oxytocin (Table 1, study numbers: 1, 2, 18/19).

Core ASD symptom measures

Global measures of ASD symptoms

Overall, there was relatively little use of measures of core ASD symptoms. Apart from one exception (Gantman *et al.*, 2012), all of the studies in which a global measure of ASD was used were of pharmacological interventions.

Ritvo–Freeman Real-life Rating Scale (RF-RLRS; Freeman et al., 1986). The Ritvo–Freeman Real-life Rating Scale (RF-RLRS) was the most frequently used global measure of ASD symptoms, albeit only in pharmacological interventions and mainly by McDougle *et al.* (1996, 1998a, 1998b; Studies 22–24, Table 1). The RF-RLRS is an observational measure of symptoms of ASD. It includes sensory motor behaviours (Subscale I: hand-flapping, rocking, pacing, etc.), social relationships (Subscale II: appropriate responses to interaction attempts, initiating appropriate physical interactions), affectual reactions (Subscale III: abrupt changes in affect, crying, temper outbursts), sensory responses (Subscale IV: agitated by noises, rubbing surfaces, sniffing self or objects), and language (Subscale V: communicative use of language, initiating appropriate verbal communication). The items reflect behaviours commonly observed in adults with ASD across the ability range and are scored on a four-point scale: (0 = never; 3 = almost always). An overall score is derived from the mean value of each of the five subscale scores (range for overall score, 0.42 to 2.58). As an observational scale the RF-RLRS can only be used to evaluate change in study participants who are repeatedly observed by trained raters. However, Arnold *et al.* (2000) found that even with extensive and repeated staff training and standardization of probes during direct observations they were unable to obtain cross-site reliability. They also stated that they were working on a more reliable system for future studies. Meanwhile they were “using the Ritvo–Freeman scale (in adapted format) as a parent rating scale supplementing other parent scales”.

Although changes were reported on various subscales of the RF-RLRS (Studies 22–24 and 26; Table 1) these differed from study to study and most subscales did not show change, even in open label studies. Whether the failure to

Table 1. Study design, sample and outcome in 30 articles on 29 reviewed studies

Article (number and first author)	Design	Intervention	Number of participants (N)/main diagnosis/age/IQ	Outcome measures (for measures not referenced in the bibliography see article cited in column 1)	Outcome in treated versus control groups
1. (Anagnostou <i>et al.</i> , 2012)	Randomized, placebo controlled, double blind	Oxytocin	N = 19 Autism spectrum disorder (ASD) Mean age 33 years (SD 13.3 years) Mean FIQ 107 (SD 24.0)	Analysis of Non-verbal Accuracy. Repetitive Behavior Scale (RBS) Social Responsiveness Scale (SRS) Yale–Brown Obsessive Compulsive Scale (Y-BOCS) Reading-the-Mind-in-the-Eyes Test Emotional/social subscales of WHO Quality of Life Non-standard tasks. Positive and Negative Affect Schedule (PANAS) Mood assessment	Some improvements in social cognition, repetitive behaviours and emotional well-being
2. (Andari <i>et al.</i> , 2010)	Randomized, placebo controlled, double blind	Oxytocin	N = 13, AS or HFA (+controls). Mean age 26 years (range 17–39 years) Mean IQ “above average” N = 9 ASD (+controls) Mean age 29 years (SD 9.9 years) Mean IQ 111.1 (SD 10.9)		Improved affect and social responses to facial stimuli
3. (Beversdorf <i>et al.</i> , 2008)	Randomized, double-blind Placebo	Propranolol	N = 14 HFA (+controls) Mean age 19 years (SD 2.9 years) Mean IQ 103.9 (SD 12.3)	Hemodynamic response rates. Observations of anagram latency scores	Improved performance on anagram latency verbal tasks; Decrease in heart rate; blood pressure (systolic and diastolic)
4. (Beversdorf <i>et al.</i> , 2011)	Randomized, double-blind Placebo	Propranolol	N = 14 ASD (+controls) Mean age 18.9 years (SD 2.9 years) Mean FSIQ 103.9 (SD 12.3)	Hemodynamic response rates. Observations of words produced	Improved performance on category fluency task but not on letter fluency; decrease in heart rate and blood pressure (systolic and diastolic) compared to controls
5. (Bodner <i>et al.</i> , 2012)	Double-blind, Counter-balanced trial	Propranolol	N = 14 ASD (+controls) Mean age 18.9 years (SD 2.9 years) Mean FSIQ 103.9 (SD 12.3)	AX continuous performance test (AX-CPT)	Significant improvement in working memory performance for individuals with ASD
6. (Bolte <i>et al.</i> , 2002)	RCT (non-blind)	Computerized facial recognition training	N = 10, high functioning autism Mean age 27.2 years (SD 7.0 years) Mean IQ 104.2 (SD 17.1)	Recognition of facial expressions	Improved facial recognition.

(Continues)

Table 1. (Continued)

Article (number and first author)	Design	Intervention	Number of participants (N)/main diagnosis/age/IQ	Outcome measures (for measures not referenced in the bibliography see article cited in column 1)	Outcome in treated versus control groups
7. (Brodtkin <i>et al.</i> , 1997)	Open label	Clomipramine	N = 35 autism, AS PDD Mean age 30 years (SD 7 years) Mean IQ 64.6 (SD 27.2; range 35–112)	Autism Behavior Checklist. Clinical Global Impression scale (CGI)	Reduction in repetitive thoughts/ behaviors; aggression; improved eye contact and verbal ability
8. (Buchsbaum <i>et al.</i> , 2001)	Placebo-controlled crossover	Fluoxetine	N = 6, AS or autism Mean age 30.5 years (SD 8.6 years) Mean IQ 95 (53–119) 23, autism (+ controls)	CGI Y-BOCS	Improved scores on (Y-BOCS), Hamilton Rating Scale for Anxiety
9. (Cook <i>et al.</i> , 1992)	Open label	Fluoxetine	Mean age 16 years (7–28 years) IQ range – profound ID to normal IQ N = 5 HFA (+controls)	Hamilton Rating Scale for Anxiety CGI	CGI Significant improvement in repetitive behavior and aggression
10. (Faja <i>et al.</i> , 2008)	Case-control comparison	Computerized face-processing training programme	Mean age 17 years (12–32 years) Mean IQ 99.4 (SD 13.9)	Benton Test of Facial Recognition. Wechsler Memory Scales for faces. Self report	No group differences on Benton test; improvements on Wechsler memory for faces
11. (Faja <i>et al.</i> , 2012)	Randomized non-blind case-control comparison	Computerized face versus object processing training programme	N = 9 HFA (+ controls) Mean age 22 years (SD 4.4 years) Mean IQ 116.3 (SD 16.3)	(as in study 10 above)	Both groups improved on Benton and Wechsler tests; Face training group showed more change in electro-physiological measures
12. (Fatemi <i>et al.</i> , 1998)	Retrospective chart review	Fluoxetine	N = 7 autistic disorder Mean age 16 years (9–20 years) IQ: three individuals moderate ID N = 17 (autism, Asperger and PDD-numbers) Mean age 20.4 (SD 1.62)	Aberrant Behavior Checklist (ABC)	Significant decrease on lethargy subscale
13. (Gantman <i>et al.</i> , 2012)	Randomized control trial; (treatment and delayed treatment control groups)	Social skills intervention, (PEERS)		SRS Social Skills Rating System (SSRS) Social and Emotional Loneliness Scale for Adults (SELSA) Quality of Socialization Questionnaire (QSQ) Social Skills Inventory (SSI) Empathy Quotient (EQ) Test of Young Adult Social Skills Knowledge, (TYASSK)	Statistically significant improvements in loneliness; social skills frequency of social contacts social responsiveness & empathy

Table 1. (Continued)

Article (number and first author)	Design	Intervention	Number of participants (N)/main diagnosis/age/IQ	Outcome measures (for measures not referenced in the bibliography see article cited in column 1)	Outcome in treated versus control groups
14. (Garcia-Villamisar and Hughes, 2007)	Randomized control	Effects of supported employment programme on executive function	Total N = 44 Autism Mean age 25.5 years (SD 3.3 years) IQ: > 35th percentile on Standard Progressive Matrices (SPM)	Cambridge Automated Battery (CANTAB) Trail Making Test Matching Familiar Figures	Improvements on: CANTAB (Spatial Span; Spatial Working Memory (Stockings of Cambridge) Trail Making Test Matching Familiar Figures test
15. (Golan and Baron-Cohen, 2006)	Case-control trial with multiple repeated measures	"Mind Reading"-interactive systematic guide to emotions	N = 19 AS/HFA (+controls) Mean age 30.5 years (17–48 years) Mean IQ 108 (SD 13.3)	Cambridge Mind-Reading Face Voice Battery	Significant improvement in the ability to recognize complex emotions and mental states from faces and voices.
16. (Hillier <i>et al.</i> , 2012)	Pilot study, no control	Music therapy	N = 22, Autism spectrum Mean age 18 years (range 13–29) IQ – "high functioning" N = 14 autism, Asperger disorder or PDD Mean age 18 years (15–40 years) Mean IQ 169 (SD 20.7)	Index of Peer Relations (IPR). Rosenberg Self-Esteem Scale (SES). State-Trait Anxiety Inventory (STAI) CGI	Increase in self-esteem, reduced anxiety, and more positive attitudes toward peers. 10 showed improvement in core autism symptoms and associated features of affective instability, impulsivity and aggression. Improvements in repetitive behaviour and affective speech comprehension.
18/19. (Hollander <i>et al.</i> , 2003; Hollander <i>et al.</i> , 2007)	Randomized, placebo-controlled, double-blind crossover design;	Oxytocin	N = 15 Autism, Asperger Mean age 32.9 years (19–56 years) Mean IQ 90.3 (SD 9.9)	Y BOCS Ability to match affective speech to emotional expression and sentences	Improvement in conversational and social skills demonstrated in enacted social scenarios
20. (Howlin and Yates, 1999)	One year follow up case series	Social group skills	N = 10 HFA Asperger Mean age 28 years (19–44 years) Mean IQ 109 (86–138) N = 5 ASD Age 20–35 years IQ within normal range	Video ratings of social skills in artificial scenarios. Self- and parent-report ratings CGI	Four significantly improved on CGI-I, one 'much worse'
21. (Jordan <i>et al.</i> , 2012)	Case series	Aripiprazole for challenging behaviour			

(Continues)

Table 1. (Continued)

Article (number and first author)	Design	Intervention	Number of participants (N)/main diagnosis/age/IQ	Outcome measures (for measures not referenced in the bibliography see article cited in column 1)	Outcome in treated versus control groups
22. (McDougle <i>et al.</i> , 1996)	Randomized double-blind placebo controlled trial	Fluvoxamine	N = 15 autism (+controls) Mean age 30 years (18–53 years) Mean IQ 79.9 (SD 29.7)	CGI Y-BOCS Vineland Maladaptive Behavior Scale (VMBS) Ritvo-Freeman Real Life Rating Scale (RF-RLRS)	Y-BOCS: Significant improvement in repetitive behaviour and aggression; VMBS: Significant reduction in maladaptive behaviours RF-RLRS: Significant improvement in language usage but not in other subscales.
23. (McDougle <i>et al.</i> , 1998a)	Prospective open label	Sertraline	N = 42 autism, AS, PDD-numbers Mean age 26 years (18–39 years) Mean IQ 60.5 (SD 22.7)	ADI & ADOS (diagnosis only) CGI RF-RLRS VMBS Y-BOCS, Self-Injury Behaviour Questionnaire (SIB-Q)	Statistically significant overall improvement in repetitive behaviour and aggression (CGI) but not on other outcome measures
24. (McDougle <i>et al.</i> , 1998b)	12 week double-blind placebo controlled trial	Risperidone	N = 31, autism and PDD Mean age = 28 years (18–43 years) Mean IQ 54.6 (range 24–113)	CGI RF-RLRS SIB-Q Modified version of Y-BOCS Visual Analogue Scales (VAS)	Significant improvement in autism symptoms [repetitive behaviours (Y-BOCS), aggression (SIB-Q), anxiety or nervousness, depression and irritability (VAS)] and on RF-RLRS scales I and III (sensorimotor behaviour and affectual reactions)
25. (Narayanan <i>et al.</i> , 2010)	Open treatment trial	Propanolol	N = 10 ASD Mean age = 24.3 years (SD 6.4 years) Mean IQ 102–9 (SD 14.0)	Phonological processing task MRI measures of functional connectivity	Improvements in both phonological task and MRI measures

Table 1. (Continued)

Article (number and first author)	Design	Intervention	Number of participants (N)/main diagnosis/age/IQ	Outcome measures (for measures not referenced in the bibliography see article cited in column 1)	Outcome in treated versus control groups
26. (Potenza <i>et al.</i> , 1999)	Open treatment trial	Olanzapine	N = 8 (four children/adolescents four adults) autism or PDD Mean age = 20.9 years Mean IQ (?) but not severely impaired	CGI, Y-BOCS RF-RLRS SIB-Q VAS VMBS	Significant improvement in symptoms of autism on RF-RLRS: motor restlessness or hyperactivity, social relatedness; sensory motor behaviours, affectual reactions; language usage, self injurious behaviours, aggression, irritability, anger, anxiety and depression Improvement in aggressive behaviour; 3 = significant changes in overall social interaction; 4 = significant changes in speech.
27. (Ratey <i>et al.</i> , 1987)	Open clinical trial	Beta blockers	N = 8, autism + extreme aggression or self-injurious behaviour Mean age 33 (25–50 years). Mean IQ (profound MR to > 120) N = 36, autism	Observational counts of target behaviours by nursing staff	
28. (Remington <i>et al.</i> , 2001)	3x Crossover RCT (double blind)	Clomipramine	Mean age 16 years (10–36 years) Mixed IQ; scores not reported N = 20 (+controls) ASD Mean age 44.4 years (SD 11.1)	Aberrant Behavior Checklist (ABC) Childhood Autism Rating Scale (CARS) Symptom Checklist-90-revised (SCL-90-R) (measure of psychological distress) Rumination-Reflection Questionnaire (RRQ) Dutch Global Mood Scale (GMS) Face Emotion Identification	Haloperidol > effective than clomipramine on autism symptoms
29. (Spek <i>et al.</i> , 2013)	Randomized control, double-blind	Haloperidol Placebo Effect of modified MBT for anxiety and depression	Mean age 42.5 years (25–55 years) Mean IQ 113 (SD 20)	Questionnaire (RRQ) Dutch Global Mood Scale (GMS) Face Emotion Identification Test (FEIT), Hinting task (ToM), Social Communication Skills Questionnaire (SCSQ), Social Skill Performance Assessment (SSPA); satisfaction ratings	Statistically significant reduction in depression, anxiety and rumination
30. (Turner-Brown <i>et al.</i> , 2008)	Quasi-experimental pilot study	Social Cognition and Interaction training	IQ: "high functioning" N = 6 HFA (+controls) Mean age 42.5 years (25–55 years) Mean IQ 113 (SD 20)		Significant improvements in social cognition and perceived social functioning

demonstrate change reflects poor sensitivity, or is a valid finding is unclear.

Autism Behavior Checklist (Krug *et al.*, 1980). The Autism Behavior Checklist is a 57-item checklist, originally designed as a diagnostic screening instrument for children. It has five scales or factors (Sensory, Relating, Body/object use, Language, Social/self help) and also provides an overall severity score. Although there are concerns about its use as a diagnostic measure (e.g. Volkmar *et al.*, 1988), Marteleto and Pedromonico (2005) suggest it can be useful in identifying children with ASD traits, especially when increased cutoff levels are employed. Its value for monitoring treatment effects in adult trials has not been established and it was used here in just one study (Brodtkin *et al.*, 1997) to assess its correlation with a pharmacological treatment outcome and not specifically as a measure of change.

Social Responsiveness Scale (SRS; Constantino and Gruber, 2005). The Social Responsiveness Scale (SRS) is a 65 item rating scale of ASD traits that focuses on social impairments (social awareness, social information processing, reciprocal communication, social avoidance and autistic mannerisms). It has been widely used in genetic studies (Constantino *et al.*, 2010) and is increasingly used as a screening measure for ASD in clinical and educational settings (Charman *et al.*, 2007). Studies of its use with children indicate good inter-rater reliability and it correlates well with the ADI-R (Constantino *et al.*, 2003; Wilkinson, 2011). The SRS was used in two of the studies reviewed here, i.e. Anagnostou *et al.* (2012) on oxytocin and Gantman *et al.* (2012) on social skills training, both of which reported improvements in scores.

Childhood Autism Rating Scale (CARS). The CARS was the only other specific measure of core ASD used in the studies reviewed. It is designed to observe and rate 15 items: relationship to people, imitation, emotional response, body use, object use, adaptation to change, visual response, listening response, taste–smell–touch response and use, fear and nervousness, verbal communication, non-verbal communication, activity level, level and consistency of intellectual response, and general impressions. The CARS was employed by Remington *et al.* (2001) to assess change in a trial of Haloperidol and Clomipramine.

Other global diagnostic measures. None of the reviewed papers made use of detailed diagnostic interviews (e.g. ADI-R; Le Couteur *et al.*, 2003), DISCO (Diagnostic Interview for Social and Communication Disorders; Wing *et*

al., 2002), ADOS (Lord *et al.*, 2000), to measure change in ASD, although they were used in some studies to confirm diagnosis. However, it should be noted that the ADOS does not appear to be particularly sensitive to change even in studies of successful treatment of children (Green *et al.*, 2010; Dawson *et al.*, 2010).

Measures of specific autism domains

Social-communication skills. There was no consistent use of standardized measures designed specifically to assess social competence in ASD. Gantman *et al.* (2012) used a number of different instruments to assess response to social skills training. These included the Social Skills Rating System (SSRS; Gresham and Elliott, 1990); the Social and Emotional Loneliness Scale for Adults (SELSA; DiTommaso and Spinner, 1993); the Empathy Quotient (EQ; Baron-Cohen and Wheelwright, 2004); the Social Skills Inventory (SSI, Riggio *et al.*, 2005); the Quality of Socialization Questionnaire (QSQ; adapted from Frankel *et al.*, 2010) and the Test of Young Adult Social Skills Knowledge, (TYASSK; Laugeson *et al.*, 2009).

Assessing the impact of social skills training (Anagnostou *et al.*, 2012), used the “Reading the Mind in the Eyes Test” (Baron-Cohen *et al.*, 2001a); Hillier *et al.* (2012) used the Index of Peer Relationships (Hudson, 1993) and Golan and Baron-Cohen used the Cambridge Mind-Reading Face Voice Battery (Golan *et al.*, 2006). Faja *et al.* (2008, 2012) used the Benton test of facial recognition (Benton, 1990) and the Wechsler Memory Scales to assess the effectiveness of a face processing training programme. Other studies employed non-standardized tasks to assess social outcome (e.g. Studies 1, 18, 19, 20, 29).

Communication skills were indirectly assessed using word processing tasks in three ASD intervention studies (Studies 3, 4, 25) although there is no indication how well these relate to social communication more generally

Repetitive behaviours. Yale–Brown Obsessive Compulsive Scale [Y-BOCS (full or modified forms); Goodman *et al.*, 1989]. Each of the 10 items on the Y-BOCS is scored on a five-point scale (0 = “least symptomatic” to 4 = “most symptomatic”; maximum score = 40). Five items assess the severity of repetitive thoughts and five assess the severity of repetitive behaviour. It was used as an outcome measure in six studies – all of which were pharmacological (Studies 8, 18/19, 22–24). However, because many individuals with ASD have limited verbal skills and may have difficulties providing information about obsessional thoughts, McDougle *et al.* (1998b) assessed only repetitive

behaviours, reporting a significant reduction in Y-BOCS scores over time.

The Repetitive Behavior Scale (RBS; Bodfish et al., 1999) Although a well established measure of repetitive behaviours in ASD (Turner-Brown *et al.*, 2011), the Repetitive Behavior Scale (RBS) was used in only one (pharmacological) study (Anagnostou *et al.*, 2012).

Non-autism specific measures

Measures of global functioning

The Clinical Global Impression (CGI) rating scale (Guy, 1976). The CGI rating scale has been widely used in clinical trials in psychiatry to assess symptom severity and response to treatment. The Severity (CGI-S) scale requires the clinician to rate the severity of the patient's illness at the time of assessment on a seven-point scale (1 = not at all ill; 7 = extremely ill). The Improvement (CGI-I) scale is used to assess how much the patient has improved or worsened following intervention (1 = very much improved; 4 = no change; 7 = very much worse). Aman *et al.* (2004) specifically recommend the use of the CGI rating scale for autism intervention trials. However, although it has generally good reliability and validity data when used by trained clinicians in adult mental health services, it is a subjective measure and thus open to bias, especially when used by untrained clinicians (Beneke and Rasmus, 1992; Berk *et al.*, 2008). Moreover, Arnold *et al.* (2000) highlight the problems of using the CGI for people with ASD because the level of disability typically found in clinical samples leads to most individuals being coded as extremely severe. As a solution (at least for children with ASD), they recommend that a CGI score of three (mildly ill) should be a baseline score applied to all participants meeting basic diagnostic criteria for autism. Higher scores should be used when significant secondary behaviours are observed such as irritability, rapid mood changes, anxiety, aggression, agitation, tantrums, self-injurious behaviour (SIB). However, they do not discuss how this modification could be applied in intervention trials involving adults. The CGI rating scale was used in nine ASD intervention studies reviewed here, all pharmacological (Studies 7–9, 17, 21–24, 26).

Global measures of “maladaptive” behaviours

Aberrant Behaviour Checklist (ABC; Aman et al., 1985). The Aberrant Behaviour Checklist (ABC) was developed to assess drug and other treatment effects mainly in individuals with severe intellectual disability. It comprises a five-factor scale: (I) irritability, agitation, crying; (II)

lethargy, social withdrawal; (III) stereotypic behaviour; (IV) hyperactivity, non-compliance; (V) inappropriate speech. The ABC has been widely used in studies of individuals including adults with intellectual disability and/or ASD and has been translated into many different languages (Aman, 2012).

The ABC was used in two studies identified in the present review, both pharmacological (Studies 12, 28) and was reported as showing improvements in at least some factor subscales.

Maladaptive subscale of the Vineland Adaptive Behavior Scale (VABS; Sparrow et al., 1984). The Vineland Adaptive Behavior Scale (VABS) was used in three of the reviewed studies, all pharmacological and conducted by McDougle's group (Studies 22, 23, 26). McDougle *et al.* (1996) found a significant reduction in the total Maladaptive score but in other ASD intervention studies overall scores did not change significantly.

Aggression and self injury

Two pharmacological studies (Studies 22, 23) used the SIB-Q, an unpublished 25-item clinician-rated instrument that assesses self-injurious behaviour, physical aggression toward others, destruction to property, and other maladaptive behaviour.

Mood states

Emotional functioning was hardly considered in the reviewed intervention studies, possibly because adults with ASD typically have difficulty in describing feelings and emotions. The Hamilton Anxiety Rating scale (Hamilton, 1959) was used in the small ($n = 6$) trial of fluoxetine conducted by Buchsbaum *et al.* (2001) and indicated significant improvements post-intervention. Hillier *et al.* (2012) used the State-Trait Anxiety Inventory (Spielberger *et al.*, 1970). The Positive and Negative Affect Schedule (PANAS) scales (Watson *et al.*, 1998) were used in the Oxytocin trial of Andari *et al.* (2010) and the Symptom Checklist-90-R (SCL-90-R; Derogatis and Spencer, 1982) in the Spek *et al.* (2008) trial of modified Mindfulness based therapy. All reported improvements in mood.

Other measures

Self-esteem and quality of life

These were issues largely ignored in the ASD intervention studies reviewed. Only one (pharmacological) study (Anagnostou *et al.*, 2012) examined any aspect of quality

of life; Hillier *et al.* (2012) monitored self-esteem. Both studies reported some improvements in the domains assessed.

Cognitive change

Two studies (Studies 5, 14) assessed change in cognitive ability. Bodner *et al.* (2012) reported improvements in working memory following a trial of propranolol. Executive function improved in the supported employment trial of Garcia-Villamizar and Hughes (2007). One further study (Narayanan *et al.*, 2010) using magnetic imaging to assess response to propranolol showed some improvement in brain connectivity.

None of the ASD intervention studies identified in the review provided information on sensitivity to change of the outcome measures used in adult autism populations.

Discussion

The aim of the present review was to explore the use of evidence based measures of outcome suitable for use in intervention trials for adults with ASD and, where provided, to consider evidence of their validity and reliability and the purpose for which they were used. However, of the 29 ASD intervention studies included in our review, most were based on small samples and often lacked any comparison groups. Thirteen treatment trials (all but four pharmacological) involved randomized control studies. None of the studies reviewed specifically assessed the sensitivity, reliability, validity, or even general utility of the measures they used for adults with ASD. The only publication we identified that discusses these issues in detail was a review paper on child pharmacological intervention studies by Arnold *et al.* (2000).

On the whole, this review indicated a surprising lack of attention to core ASD symptoms. The RF-RLRS was used in four studies (all pharmacological) and of the five subscales, the social relatedness subscale appeared to be the most promising in reflecting change. Although problems in achieving inter-rater reliability across different centres have been reported (Arnold *et al.*, 2000) the RF-RLRS may have some value in single-site pharmacological studies with facilities for repeated direct observation and monitoring of rater consistency; its value in non-pharmacological studies has not been tested. The SRS (Constantino and Gruber, 2005) was found to reflect improvements in social skills in two studies. However, the version used was not specifically designed for adults and although there is an adult version available (SRS-A; Bolte, 2012) the relatively small number of items is likely to limit sensitivity.

We could not identify any adult study that assessed change using detailed interview or observational diagnostic instruments such as the DISCO, ADI-R or ADOS, although work has been conducted on the measurement of change in autism severity in children based on the ADOS (Shumway *et al.*, 2011). One small randomized treatment trial (White *et al.*, 2014) also provides data on children up to age 17, showing sensitivity to change on the Developmental Disability – Child Global Assessment Scale. Given that diagnostic approaches have been used to evaluate change in most other forms of mental and behavioural disorder in adults (Hamilton, 1967; Beck and Steer, 1990), the value of such approaches to adult ASD intervention research should not be overlooked. However, it is important to note that these instruments were primarily designed to identify stable diagnostic traits and were not initially developed for use in outcome assessment; any such use would first need to undergo objective testing and evaluation. A more detailed breakdown of individual item scores on the ADOS or ADI-R, rather than reliance solely on algorithm items, for example, might be more sensitive to improvement following treatment. Recent work on a childhood version of the ADOS that is specifically designed to measure change (BOSCC; Nordahl-Hansen *et al.*, 2014) is also of potential value, especially if this can be modified for use with adults. There is also a version of the Y-BOCS for measuring change but only in children (Scahill *et al.*, 2006). A rating scale that assesses DSM-IV Pervasive Developmental Disorder (PDD) criteria “in the past two weeks” that might be of potential value in evaluating ASD intervention outcome (OSU Research Unit on Pediatric Psychopharmacology, 2005) does not appear to have been used in any published trials on adults.

The most commonly used outcome measures in the papers reviewed were not specific to ASD. The CGI scale (Guy, 1976) was the most frequently used (nine studies), followed by the Y-BOCS (either full or modified; six studies). The CGI rating scale has two features that particularly recommend it for ASD trials: there is scope for cross comparison with the many other trials in which the CGI rating scale has been used elsewhere in psychiatry, and it appears to reflect general clinical improvement even in quite small ASD studies (although it is unclear if raters were sufficiently blinded in between-group comparisons). The Y-BOCS (used mostly by McDougle and his collaborators) appears to be sensitive to change in repetitive behaviours following pharmacological intervention for ASD and has also been reported to differentiate ASD specific repetitive behaviours from obsessive compulsive behaviours (McDougle *et al.*, 1995); a finding that needs to be independently replicated.

The maladaptive subscale of the VABS was used in a number of trials although scores did not always change post-treatment. Surprisingly, the VABS, which provide information on overall levels of communication, social, self-help and daily living skills, was not used in any of the studies reviewed.

There was little evidence of the use in ASD intervention studies of the low cost self-report methods (Brugha and Meltzer, 2008; Fitzpatrick *et al.*, 1998) that are typically used in intervention trials for psychiatric disorders. For example, although the AQ (Baron-Cohen *et al.*, 2001b) is frequently used as a self-report screening measure for higher ability individuals with ASD, it was not designed as a measure of change and was not used in any of the studies reviewed. There are also concerns about its validity (Bishop and Seltzer, 2012; Brugha *et al.*, 2012). The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R; Ritvo *et al.*, 2011) is a longer (80 item) self-report scale, also designed to assist in the diagnosis of adults with ASD that might prove a more sensitive alternative, but there have been no trials of its use as a measure of response to treatment. The shorter (14 item) form of the RAADS (Eriksson *et al.*, 2013) has been validated only as a screening measure.

On the whole, most intervention studies reviewed used measures designed for different purposes, for different age or IQ groups, or non-ASDs, without attempting to validate their use with adults with ASD. The Autism Treatment Evaluation Checklist (ATEC; Rimland and Edelson, 2000), for example, has been shown to have some potential for assessing change following treatment (Magiati *et al.*, 2011) but we could find no examples of its use with adults. Other potentially useful measures, such as the adult version of the Developmental Behaviour Checklist (DBC-A; Mohr *et al.*, 2005) have not been used to assess treatment outcomes in ASDs although the child DBC has been found to be sensitive to change (Clarke *et al.*, 2003).

Finally, there was no use of individualized approaches to outcome assessment, such as Goal Attainment Scaling (Kiresuk *et al.*, 1994), which allows for the evaluation of individual-specific outcome goals as an adjunct to group standardized measures (Ruble *et al.*, 2012).

The Autism Research Unit on Pediatric Psychopharmacology (RUPP Autism Network; Arnold *et al.*, 2000), although primarily focusing on childhood treatment studies, noted that heterogeneity of ASD calls for “outcome measures sensitive to individual change over a wide spectrum of treatment response and side effects”. Arnold *et al.* (2000) also noted the importance of measuring side effects and clinical physiological responses to drug trials

in children with ASD, factors that deserve more consideration in adult trials as well.

When considering broad differences in outcome assessments between pharmacological and non-pharmacological trials, the following scales were only used in pharmacological studies in ASD (see earlier for citations): RF-RLRS, Y-BOCS, CGI, ABC, VABS, SIB-Q, and these need to be evaluated within a broader range of intervention research. Generally there was greater heterogeneity in the outcome measures used in non-pharmacological ASD intervention studies, making between study comparisons even more difficult. This indicates the need for researchers to reach a consensus on the use of “core” outcome measures in both pharmacological and non-pharmacological ASD intervention studies.

Information on the psychometric properties of measures in adult ASD populations was not provided including, crucially, data on sensitivity to change (Arnold *et al.*, 2000). It would have been particularly helpful to have such information across multiple trials and most of all within adequately powered randomized controlled trials in order to ensure objectivity, validity and reliability (Wagner *et al.*, 2007).

Limitations

Our review considered publications over almost five decades but did not include a search for unpublished reports or “grey” literature. Thus, it is possible that we overlooked dissertation studies that include worthwhile evaluations of outcome measures although it is unlikely that these would involve trials of sufficient size and power to affect our overall findings.

Conclusions

Very few methods for measuring change (either ASD specific or non-specific) following interventions for adults with ASD were identified in this systematic review of outcome assessments. On the basis of existing data there is no single assessment instrument or set of instruments that can be considered to be the standard measure of outcome in the field of ASD treatment evaluation. We could find few data on the validity or reliability of the measures when used with individuals with ASD and no data on sensitivity to change; similarly there was little discussion of the rationale for choosing particular approaches to assessing outcomes, the problems encountered, or the possible advantages of any measures.

Measures of change suitable for use in the evaluation of ASD treatments need to be developed with some urgency as the lack of standardized adult measures is also reflected

in the paucity of adult intervention research generally (Bishop-Fitzpatrick *et al.*, 2013; Henninger and Taylor, 2013; Howlin and Moss, 2012; Shattuck *et al.*, 2012).

Work is needed, too, to adapt existing instruments and/or develop and evaluate new ones for assessing ASD specific or related outcomes in adult intervention studies. There is a particular need for low cost, valid and reliable

self-report measures for use in large-scale randomized controlled trials. Teams conducting current and future treatment trials should make use of the trial to evaluate and report on the sensitivity and reliability of the measures employed (cf. Arnold *et al.*, 2000). Careful attention to training and reliability in multi-site studies is also necessary.

References

- Aman M.G.. (2012). Annotated Bibliography on the Aberrant Behavior Checklist (ABC), p. 264, Columbus, OH, The Nisonger Center, Ohio State University.
- Aman M.G., Novotny S., Samango-Sprouse C., Lecavalier L., Gadow K.D., King B.H., Pearson D.A., Gernsbacher M.A., Chez M.. (2004) Outcome measures for clinical drug trials in autism. *CNS Spectrums*, **9**(1), 36–47.
- Aman M.G., Singh N.N., Stewart A.W., Field C.J.. (1985) The aberrant behavior checklist: a behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency*, **89**(5), 485–491.
- American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders, 4th edn, Washington DC: American Psychiatric Association.
- American Psychiatric Association. (2013) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5™), Washington DC: American Psychiatric Press.
- Anagnostou E., Soorya L., Chaplin W., Bartz J., Halpern D., Wasserman S., Wang A.T., Pepa L., Tanel N., Kushki A. (2012) Intranasal oxytocin versus placebo in the treatment of adults with autism spectrum disorders: a randomized controlled trial. *Molecular Autism*, **3**(1), 1–9.
- Andari E., Duhamel J.R., Zalla T., Herbrecht E., Leboyer M., Sirigu A. (2010) Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(9), 4389–4394.
- Arnold L., Aman M., Martin A., Collier-Crespin A., Vitiello B., Tierney E., Asarnow R., Bell-Bradshaw F., Freeman B., Gates-Ulanet B., Klin A., McCracken J., McDougle C., McGough J., Posey D., Scahill L., Swiezy N., Ritz L., Volkmar F. (2000) Assessment in multisite randomized clinical trials of patients with autistic disorder: the autism RUPP network. *Journal of Autism and Developmental Disorder*, **30**(2), 99–111.
- Baron-Cohen S., Wheelwright S. (2004) The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, **34**(2), 163–175.
- Baron-Cohen S., Wheelwright S., Hill J., Raste Y., Plumb I. (2001a) The “Reading the Mind in the Eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, **42**(2), 241–251.
- Baron-Cohen S., Wheelwright S., Skinner R., Martin J. (2001b) The Autism Spectrum Quotient (AQ): evidence from Asperger syndrome/ high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorder*, **31**(1), 5–17.
- Beck A.T., Steer R.A. (1990) Manual for the Beck Anxiety Inventory, San Antonio: TX, Psychological Corporation.
- Beneke M., Rasmus W. (1992) “Clinical Global Impressions” (ECDEU): some critical comments. *Pharmacopsychiatry*, **25**(4), 171.
- Benton A.L. (1990) Facial recognition. *Cortex*, **26**, 491–499.
- Berk M., Ng F., Dodd S., Callaly T., Campbell S., Bernardo M., Trauer T.. (2008) The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. *Journal of Evaluation in Clinical Practice*, **14**(6), 979–983.
- Beversdorf D.Q., Carpenter A.L., Miller R.F., Cios J.S., Hillier A. (2008) Effect of propranolol on verbal problem solving in autism spectrum disorder. *Neurocase*, **14**(4), 378–383.
- Beversdorf D.Q., Saklayen S., Higgins K.F., Bodner K.E., Kanne S.M., Christ S.E. (2011) Effect of propranolol on word fluency in autism. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, **24**(1), 11–17.
- Bishop S.L., Seltzer M.M. (2012) Self-reported autism symptoms in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **42**(11), 2354–2363.
- Bishop-Fitzpatrick L., Minshew N.J., Eack S.M. (2013) A systematic review of psychosocial interventions for adults with autism spectrum disorders. *Journal of Autism and Developmental Disorder*, **43**(3), 687–694.
- Bodfish J.W., Symons F.J., Lewis M.H. (1999) The Repetitive Behavior Scale (RBS), Morganton NC: Western Carolina Center Research Reports.
- Bodner K.E., Beversdorf D.Q., Saklayen S.S., Christ S.E. (2012) Noradrenergic moderation of working memory impairments in adults with autism spectrum disorder. *Journal of the International Neuropsychological Society*, **18**(3), 556.
- Bolte S.. (2012) Brief report: the Social Responsiveness Scale for Adults (SRS-A): initial results in a German cohort. *Journal of Autism and Developmental Disorders*, **42**(9), 1998–1999.
- Bolte S., Feineis-Matthews S., Leber S., Dierks T., Hubl D., Poustka F. (2002) The development and evaluation of a computer-based program to test and to teach the recognition of facial affect. *International Journal of Circumpolar Health*, **61**(Supplement 2), 61–68.
- Brodin E.S., McDougle C.J., Naylor S.T., Cohen D.J., Price L.H. (1997) Clomipramine in adults with pervasive developmental disorders: a prospective open-label investigation. *Journal of Child and Adolescent Psychopharmacology*, **7**(2), 109–121.
- Brugha T.S., McManus S., Bankart J., Scott F., Purdon S., Smith J., Bebbington P., Jenkins R., Meltzer H. (2011) Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry*, **68**(5), 459–466.
- Brugha T.S., McManus S., Smith J., Scott F.J., Meltzer H., Purdon S., Berney T., Tantam D., Robinson J., Radley J., Bankart J. (2012) Validating two survey methods for identifying cases of autism spectrum disorder among adults in the community. *Psychological Medicine*, **42**(3), 647–656.
- Brugha T.S., Meltzer H. (2008) Measurement of psychiatric and psychological disorders and outcomes in populations. In Heggenhougen K., Quah S. (eds) International Encyclopedia

- of Public Health, pp. 261–272, San Diego CA: Academic Press.
- Buchsbaum M., Hollander E., Haznedar M., Tang C., Spiegel-Cohen J., Wei T., Solimando A., Buchsbaum B., Robins D., Bienstock C., Cartwright C., Mosovich S. (2001) Effect of fluoxetine on regional cerebral metabolism in autistic spectrum disorders: a pilot study. *International Journal of Neuropsychopharmacology*, **4**(2), 119–125.
- Charman T., Baird G., Simonoff E., Loucas T., Chandler S., Meldrum D., Pickles A. (2007) Efficacy of three screening instruments in the identification of autistic-spectrum disorders. *British Journal of Psychiatry*, **191**, 554–559.
- Clarke A.R., Tonge B.J., Einfeld S.L., Mackinnon A. (2003) Assessment of change with the Developmental Behaviour Checklist. *Journal of Intellectual Disability Research*, **47**(Part 3), 210–212.
- Constantino J.N., Davis S.A., Todd R.D., Schindler M.K., Gross M.M., Brophy S.L., Metzger L.M., Shoushtari C.S., Splinter R., Reich W. (2003) Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorder*, **33**(4), 427–433.
- Constantino J.N., Gruber C.P. (2005) Social Responsiveness Scale. Los Angeles CA: Western Psychological Services.
- Constantino J.N., Zhang Y., Frazier T., Abbacchi A. M., Law P. (2010) Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry*, **167**(11), 1349–1356.
- Cook E., Rowlett R., Jaseleskis C., Leventhal B. (1992) Fluoxetine treatment of children and adults with autism disorder and mental retardation. *Journal of the American Academy of Child and Adolescent Psychiatry*, **31**, 739–745.
- Dawson G., Rogers S., Munson J., Smith M., Winter J., Greenon J., Donaldson A., Varley J. (2010) Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, **125**(1), e17–e23.
- Dent A.K., Carr T., Leitman S., Lord C.E. (2012) Effects of Parent-mediated Early Intervention on Child Behavioral Outcomes in an Underserved Population. 2012 International Meeting for Autism Research: IMFAR, Toronto, Canada.
- Derogatis L., Spencer P. (1982) Administration and Procedures: BSI Manual. Baltimore: MD, John Hopkins University Press.
- DiTommaso E., Spinner B. (1993) The development and initial validation of the Social and Emotional Loneliness Scale for Adults (SELSA). *Personality and Individual Differences*, **14**(1), 127–134.
- Eriksson J.M., Andersen L.M.J., Bejerot S. (2013) RAADS-14 screen: validity of a screening tool for autism spectrum disorder in an adult psychiatric population. *Molecular Autism*. DOI: 10.1186/2040-2392-4-49
- Faja S., Aylward E., Bernier R., Dawson G. (2008) Becoming a face expert: a computerized face-training program for high-functioning individuals with autism spectrum disorders. *Developmental Neuropsychology*, **33**(1), 1–24.
- Faja S., Webb S.J., Jones E., Merkle K., Kamara D., Bavaro J., Aylward E., Dawson G. (2012) The effects of face expertise training on the behavioral performance and brain activity of adults with high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **42**(2), 278–293.
- Fatemi S., Realmuto G., Khan L., Thursa P. (1998) Fluoxetine in treatment of adolescent patients with autism: longitudinal open trial. *Journal of Autism and Developmental Disorder*, **28**(4), 303–307.
- Ferriter M., Hare D., Bendall P., Cordess C., Wlliot K.I.H., Humpston R., Souflas P., Taylor M. (2001) Brief report: assessment of a screening tool for autism spectrum disorder in adult population. *Journal of Autism and Developmental Disorder*, **31**(3), 351–353.
- Fitzpatrick R., Davey C., Buxton M.J., Jones D.R. (1998) Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment*, **2**(14), 1–74.
- Frankel F., Myatt R., Sugar C., Whitham C., Gorospe C.M., Laugeson E. (2010) A randomized controlled study of parent-assisted children friendship training with children having autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **40**(7), 827–842.
- Freeman B.J., Ritvo E.R., Yokota A., Ritvo A. (1986) A scale for rating symptoms of patients with the syndrome of autism in real life settings. *Journal of the American Academy of Child and Adolescent Psychiatry*, **25**(1), 130–136.
- Gantman A., Kapp S.K., Orenski K. Laugeson E.A. (2012) Social skills training for young adults with high-functioning autism spectrum disorders: a randomized controlled pilot study. *Journal of Autism and Developmental Disorders*, **42**(6), 1094–1103.
- Garcia-Villamisar D., Hughes C. (2007) Supported employment improves cognitive performance in adults with Autism. *Journal of Intellectual Disability Research*, **51**(2), 142–150.
- Golan O., Baron-Cohen S. (2006) Systemizing empathy: teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, **18**, 591–617.
- Golan O., Baron-Cohen S., Hill J. (2006) The Cambridge Mindreading (CAM) Face-Voice Battery: testing complex emotion recognition in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorder*, **36**(2), 169–183.
- Goodman W.K., Price L.H., Rasmussen S.A., Mazure C., Delgado P., Heninger G.R., Charney D.S. (1989) The Yale–Brown Obsessive Compulsive Scale. II. Validity. *Archives of General Psychiatry*, **46**(11), 1012–1016.
- Green J., Charman T., McConachie H., Aldred C., Slonims V., Howlin P., Le C.A., Leadbitter K., Hudry K., Byford S., Barrett B., Temple K., Macdonald W., Pickles A. (2010) Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *Lancet*, **375**(9732), 2152–2160.
- Gresham F.M., Elliott S.N. (1990) Social Skills Rating System. Circle Pines, MN, American Guidance Service.
- Guy W. (1976) CGI (Clinical Global Impression) ECDEU Assessment Manual for Psychopharmacology (NIMH Publication 76-338), pp. 218–222, Washington, DC, DHEW, NIMH.
- Hamilton M. (1959) The assessment of anxiety states by rating. *British Journal of Medical Psychology*, **32**(1), 50–55.
- Hamilton M. (1967) The development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, **6**, 278–296.
- Henninger N.A., Taylor J.L. (2013) Outcomes in adults with autism spectrum disorders: a historical perspective. *Autism*, **17**(1), 103–116.
- Hillier A., Greher G., Poto N., Dougherty M. (2012) Positive outcomes following participation in a music intervention for adolescents and young adults on the autism spectrum. *Psychology of Music*, **40**(2), 201–215.
- Ho B., Carter M., Stephenson J. (2014) Cognitive-behavioral approach for children with autism spectrum disorders: a meta-analysis. *Review Journal of Autism and Developmental Disorders*, **1**, 18–33, DOI: 10.1007/s40489-013-0002-5
- Hollander E., Bartz J., Chaplin W., Phillips A., Sumner J., Soorya L., Anagnostou E., Wasserman S. (2007) Oxytocin increases

- retention of social cognition in autism. *Biological Psychiatry*, **61**(4), 498–503.
- Hollander E., Dolgoff-Kaspar R., Cartwright C., Rawitt R., Novotny S. (2001) An open trial of divalproex sodium in autism spectrum disorders. *Journal of Clinical Psychiatry*, **62**(7), 530–534.
- Hollander E., Novotny S., Hanratty M., Yaffe R., DeCaria C.M., Aronowitz B.R., Mosovich S. (2003) Oxytocin infusion reduces repetitive behaviors in adults with autistic and Asperger's disorders. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, **28**(1), 193–198.
- Howlin P., Goode S., Hutton J., Rutter M. (2004) Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, **45**(2), 212–229.
- Howlin P., Moss P. (2012) Adults with autism spectrum disorders. *Canadian Journal of Psychiatry*, **57**(5), 275–283.
- Howlin P., Moss P., Savage S., Rutter M. (2013) Social outcomes in mid to later adulthood among individuals diagnosed with autism and average non-verbal IQ as children. *Journal of the American Academy of Child and Adolescent Psychiatry*, **52**(6), 572–581.
- Howlin P., Yates P. (1999) The potential effectiveness of social skills groups for adults with autism. *Autism*, **3**, 299–307.
- Hudson W.W. (1993) Index of Peer Relations (IPR), Tallahassee, FL, Walmyr Publishing Company.
- Jordan I., Robertson D., Catani M., Craig M., Murphy D. (2012) Aripiprazole in the treatment of challenging behaviour in adults with autism spectrum disorder. *Psychopharmacology*, **223**(3), 357–360.
- Kanne S.M., Abbacchi A.M., Constantino J.N. (2009) Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: the importance of environmental context. *Journal of Autism and Developmental Disorders*, **39**, 856–864, DOI: 10.1007/s10803-009-0694-7.
- Kiresuk T.J., Smith A., Cardillo J.E. (1994) Goal Attainment Scaling: Applications, Theory, and Measurement, Hillsdale, NJ, Lawrence Erlbaum.
- Krug D.A., Arick J., Almond P. (1980) Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **21**(3), 221–229.
- Laugeson E.A., Frankel F., Mogil C., Dillon A.R. (2009) Parent-assisted social skills training to improve friendships in teens with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **39**(4), 596–606.
- Le Couteur A., Lord C., Rutter M. (2003) Autism Diagnostic Interview – Revised. Los Angeles, CA, Western Psychological Services.
- Lord C., Rutter M., DiLavore P.C., Risi S. (2000) Autism Diagnostic Observation Schedule (ADOS), Los Angeles, CA, Western Psychological Services.
- Lounds T.J., Dove D., Veenstra-VanderWeele J., Sathe N.A., McPheeters M.L., Jerome R.N., Warren Z. (2012) Interventions for Adolescents and Young Adults With Autism Spectrum Disorders, Comparative Effectiveness Review No. 65, AHRQ Publication No. 12-EHC063-EF, Rockville, MD, Agency for Healthcare Research and Quality.
- Magiati I., Moss J., Yates R., Charman T., Howlin P. (2011) Is the Autism Treatment Evaluation Checklist a useful tool for monitoring progress in children with autism spectrum disorders? *Journal of Intellectual Disability Research*, **55**(3), 302–312.
- Marteletto M.R., Pedromonico M.R. (2005) Validity of Autism Behavior Checklist (ABC): preliminary study. *Revista brasileira de psiquiatria (Sao Paulo, Brazil: 1999)*, **27**(4), 295–301.
- McDougle C., Brodtkin E., Naylor S., Carlson D., Cohen D., Price L. (1998a) Sertraline in adults with pervasive developmental disorders: a prospective open-label investigation. *Journal of Clinical Psychopharmacology*, **81**(1), 62–66.
- McDougle C., Holmes J., Carlson D., Pelton G., Cohen D., Price L. (1998b) A double-blind, placebo-controlled study of Risperidone in adults with autistic disorder and other pervasive developmental disorders. *Archives of General Psychiatry*, **55**, 633–641.
- McDougle C., Naylor S., Cohen D., Volkmar F., Heninger G., Price L. (1996) A double-blind, placebo-controlled study of fluvoxamine in adults with autistic disorder. *Archives of General Psychiatry*, **53**, 1001–1008.
- McDougle C.J., Kresch L.E., Goodman W.K., Naylor S.T., Volkmar F.R., Cohen D.J., Price L.H. (1995) A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. *American Journal of Psychiatry*, **152**(5), 772–777.
- Mohr C.D., Tonge B.J., Einfeld S.L. (2005) The development of a new measure for the assessment of psychopathology in adults with intellectual disability. *Journal of Intellectual Disability Research*, **49**(7), 469–480.
- Narayanan A., White C.A., Saklayen S., Scaduto M. J., Carpenter A.L., Abduljalil A., Schmalbrock P., Beversdorf D.Q. (2010) Effect of propranolol on functional connectivity in autism spectrum disorder. *Brain Imaging and Behavior*, **4**(2), 189–197.
- National Autistic Society. (2012) National Autistic Society (NAS) library. <http://www.nas.org.uk/> [August 2008].
- National Institute for Health and Clinical Excellence (Great Britain) (NICE). (2012) Autism: Recognition, Referral, Diagnosis and Management of Adults on the Autism Spectrum, London: NICE.
- Nordahl-Hansen A., Blanc R., Kaale A., McConachie H., Fletcher-Watson S. (2014) A review of the Brief Observation of Social Communication Change (BOSCC). Published Poster Presentation. Enhancing the Scientific Study of Early Autism. EU COST conference. September 2014. France, Toulouse.
- Potenza M., Holmes J., Kanes S., McDougle C. (1999) Olanzapine treatment of children, adolescents, and adults with pervasive developmental disorders: an open-label pilot study. *Journal of Clinical Psychopharmacology*, **19**(1), 37–44.
- Ratey J.R., Bemporad J., Sorgi P., Bick P., Polakoff S., O'Driscoll G., Mikkelsen E. (1987) Brief report: Open trial effects of beta-blockers on speech and social behaviors in 8 autistic adults. *Journal of Autism and Developmental Disorders*, **17**(3), 439–446.
- Remington G., Sloman L., Konstantareas M., Parker K., Gow R. (2001) Clomipramine versus Haloperidol in the treatment of autistic disorder: a double-blind, placebo-controlled, crossover study. *Journal of Clinical Psychopharmacology*, **21**(4), 440–444.
- Research Autism. (2013). Research autism improving the quality of life. http://www.researchautism.net/pages/About_Us/index
- Riggio R.E., Tucker J., Coffaro D. (2005) Social Skills Inventory. A Measure of Verbal, Non-verbal Social Competence and Emotional Intelligence. Menlo Park, CA, Mind Garden Inc.
- Rimland B., Edelson S. (2000) Autism Treatment Evaluation Checklist: Statistical Analyses, San Diego, CA: Autism Research Institute.
- Ritvo R.A., Ritvo E.R., Guthrie D., Ritvo M.J., Hufnagel D.H., McMahon W., Tonge B., Mataix-Cols D., Jassi A., Attwood T., Eloff J. (2011) The Ritvo Autism Asperger Diagnostic Scale-Revised (RAADS-R): a scale to assist the diagnosis of autism spectrum disorder in adults: an international validation study. *Journal of Autism and Developmental Disorders*, **41**(8), 1076–1089.

- Robins L.N., Wing J., Wittchen H.U., Helzer J.E., Babor T.F., Burke J., Farmer A., Jablensky A., Pickens R., Regier D.A., Sartorius N., Towle M.S. (1988) The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, **45**(12), 1069–1077.
- Ruble L., McGrew J.H., Toland M.D. (2012) Goal attainment scaling as an outcome measure in randomized controlled trials of psychosocial interventions in autism. *Journal of Autism and Developmental Disorders*, **42**(9), 1974–1983.
- Russell A.J., Mataix-Cols D., Anson M., Murphy D. G. (2005) Obsessions and compulsions in Asperger syndrome and high-functioning autism. *British Journal of Psychiatry*, **186**, 525–528.
- 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., Vitiello B. (2006) Children's Yale–Brown Obsessive Compulsive Scale modified for pervasive developmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, **45**(9), 1114–1123.
- Schopler E., Reichler R.J., DeVellis R.F., Daly K. (1980) Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, **10**(1), 91–103.
- Shattuck P.T., Roux A.M., Hudson L.E., Taylor J. L., Maenner M.J., Trani J.F. (2012) Services for adults with an autism spectrum disorder. *Canadian Journal of Psychiatry*, **57**(5), 284–291.
- Shumway S., Thurm A., Swedo S.E., Deprey L., Barnett L.A., Amaral D.G., Rogers S.J., Ozonoff S. (2011) Brief report: symptom onset patterns and functional outcomes in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, **41**(12), 1727–1732.
- Sparrow S.S., Balla D.A., Cicchetti D.V. (1984) Vineland Adaptive Behavior Scales: Interview Edition Survey Form, Circle Pines, MN: American Guidance Service.
- Spek A., Scholte E., Berckelaer-Onnes I. (2008) Brief report: the use of WAIS-III in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorder*, **38**, 782–787.
- Spek A.A., van Ham N.C., Nykceck I. (2013) Mindfulness-based therapy in adults with an autism spectrum disorder: a randomized controlled trial. *Research in Developmental Disabilities*, **34**(1), 246–253.
- Spielberger C., Gorsuch R., Lushene R. (1970) Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire), Palo Alto: CA, Consulting Psychologists Press.
- The OSU Research Unit on Pediatric Psychopharmacology. (2005) OSU Autism Rating Scale—DSM-IV (OARS-4), Columbus, OH: The OSU Research Unit on Pediatric Psychopharmacology.
- Turner-Brown L.M., Lam K.S., Holtzclaw T.N., Dichter G.S., Bodfish J.W. (2011) Phenomenology and measurement of circumscribed interests in autism spectrum disorders. *Autism*, **15**(4), 437–456.
- Turner-Brown L.M., Perry T.D., Dichter G.S., Bodfish J.W., Penn D.L. (2008) Brief report: feasibility of social cognition and interaction training for adults with high functioning autism. *Journal of Autism and Developmental Disorders*, **38**(9), 1777–1784.
- Volkmar F.R., Cicchetti D.V., Dykens E., Sparrow S.S., Leckman J.F., Cohen D.J. (1988) An evaluation of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, **18**(1), 81–97.
- Wagner A., Lecavalier L., Arnold L.E., Aman M.G., Scahill L., Stigler K.A., Johnson C.R., McDougle C.J., Vitiello B. (2007) Developmental disabilities modification of the Children's Global Assessment Scale. *Biological Psychiatry*, **61**(4), 504–511.
- Watson D., Clark L.A., Tellegen A. (1998) Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, **54**, 1063–1070.
- White S.W., Smith L.A., Schry A.R. (2014) Assessment of global functioning in adolescents with autism spectrum disorders: utility of the Developmental Disability – Child Global Assessment Scale. *Autism*, **18**(4), 362–369.
- Wilkinson L.A. (2011) Identifying students with autism spectrum disorders: a review of selected screening tools. *Communique*, **40**(2), 31–33.
- Wing J.K., Babor T., Brugha T., Burke J., Cooper J. E., Giel R., Jablensky A., Regier D., Sartorius N. (1990) SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Archives of General Psychiatry*, **47**(6), 589–593.
- Wing L., Leekam S.R., Libby S.J., Gould J., Larcombe M. (2002) The Diagnostic Interview for Social and Communication Disorders: background, inter-rater reliability and clinical use. *Journal of Child Psychology and Psychiatry*, **43**(3), 307–325.
- World Health Organization (WHO) (1993) The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research, WHO: Geneva.