NETSCC, HTA

28 March 2012

MeASURe: Measurement in autism spectrum disorder under review

HTA no 11/22 Questions:
What is the validity of tools and outcome measures used in measuring and monitoring autism spectrum disorder?
How well do these reflect and measure issues of importance for patients and carers?

Planned Investigation
Objectives

1. To identify the tools and outcome measures reported in literature on quantitative research involving young children with autism spectrum disorder (ASD).
2. To identify priorities for child outcomes as valued by parents, young people with ASD and professionals.
3. To conduct detailed systematic reviews of measurement properties of tools within the major domains of development and functioning.
4. To synthesise evidence regarding the most robust and useful tools in these different domains.
5. To recommend a potential battery of tools and outcome measures for use in research and clinical practice.
6. To identify gaps for future development of appropriate outcome measures.

Autism spectrum disorders (ASD) are neurodevelopmental, lifelong conditions diagnosed using a set of behavioural criteria (ICD-10 1992 and DSM-IV 1994), and characterised by problems in social interaction and communication, along with restricted and repetitive behaviours. Children with ASD are increasingly diagnosed as early as 2 years of age; around 1% of school age children are recognized to have ASD (Baird et al 2006). Children's abilities and severity of difficulties vary greatly, and four times as many boys as girls have ASD.

In the light of increased awareness about the prevalence of ASD, and the emphasis on early identification and diagnosis, it is important that health, education and social care services provide evidence-based interventions and early support for individuals with ASD, and their families, carers and teachers. In the past decade there has been an increase in ASD intervention research, with recent improvement in the quality of studies (Charman 2011). The ASD early intervention literature is largely focussed on promotion of social communication skills, with less emphasis on interventions for restricted and repetitive behaviours; it also includes interventions (including bio-medical interventions) focused on the high rates of co-occurring behaviours and problems (e.g. sleep, faddiness about food, aggression to others, toileting difficulties). One problem for the interpretation of research findings is the multitude of different measurement tools which have been used in collecting evidence of progress and outcomes. Furthermore, longitudinal studies highlight the variation in individual developmental pathways (Charman et al., 2005; Anderson et al., 2007). The literature thus presents a large set of measures, inconsistently used, of varying relevance and with variable or indeed no evidence of their psychometric properties. A review of the quality and appropriateness of measures currently used with children up to the age of about 6 years in monitoring their progress and outcomes is needed, in order to identify a set of robust and appropriate measures to be used in the future by researchers and assist in sound decision making by service providers and service users.

Policy background
Autism was once considered a relatively rare condition; now ASD is known to affect at least 1% of the child and adult population (Baird et al., 2006; Baron-Cohen et al., 2009; Brugha et al., 2011). There is wide variation in the progress made by individuals with ASD, so that many individuals have significant lifelong needs for support. The burden and cost to the individual, family and broader society are very high, with the economic costs in the UK being estimated to be £28bn per year (Knapp et al., 2009). ASD has become a

priority area for national government, local education, social care, health and mental health services and charitable organisations. This is reflected in:

- the National Audit Office report (2009),
- the Autism Act (2009) leading to the development of an Adult Autism Strategy by the Department of Health (2010), and
- forthcoming publication of NICE guidelines on recognition, referral and diagnosis of
  children and young people with ASD (due September 2011),
  adults with an ASD (due 2012), and
  management of children and young people with an ASD (due 2013).

National policy and practice guidelines are available for parts of the UK (ASD Strategic Action Plan for Wales: http://wales.gov.uk/topics/childrenyoungpeople/publications/autisticspectrumdisorderplan/?lang=en; Scottish Autism Strategy: http://www.scotland.gov.uk/Publications/2010/09/07141141/0; SIGN, 2007; ASD Strategic Action Plan for Northern Ireland) and internationally. The publication of NICE guidelines for children and young people with an ASD will ensure that early identification and provision of health and education services in the preschool and early school years will continue to be a priority for both families and services over the coming decade.

What should be measured?

There are several ways to consider the question of what to measure: what the NHS needs in order to measure progress and outcomes, what matters to parents and individuals with ASD, and the theoretical basis of ASD which has implications regarding important domains to measure. There is currently consultation about the NHS Outcomes Framework 2011/12 (2010), part of a strategy that aims to deliver 'the outcomes that matter most to people', using 'patient-reported outcome measures' (PROMs). The Kennedy report 'Getting it right for children and young people' (2010) highlighted the need to identify a common vision between families and professionals for what services are seeking to achieve for children. Measuring outcomes that are valued by families is central to that vision, which in turn will influence what services are provided and how, and potentially what services and interventions are prioritised for research evaluation. However, there is currently no agreement on how best to measure valued outcomes for children with neurodevelopmental disorders in general, let alone the heterogeneous group of children with ASD. The project will address this problem by bringing together views of parents, individuals with ASD and professionals, alongside a comprehensive search of current measurement tools, informed by a strong conceptual framework. One important influence on the framework will be the International Classification of Functioning, Disability and Health for Children and Youth (ICF-CY, WHO 2007), so that what is measured can be 'mapped' against domains of functioning (e.g. learning, communication, self-care) and participation (e.g. relationships, community life).

The conceptual framework will also be influenced by an understanding of ASD. The behavioural characteristics of ASD, on which the current diagnostic classifications ICD10 and DSM-IV are based, are underpinned by genetic, brain structure, and neuropsychological differences from typical development (Tager-Flusberg & Joseph 2003; Happé & Ronald 2008). For example, many studies suggest that the impairments in ASD are best considered within two groupings, social-communication difficulties and restricted/repetitive interests and behaviours. The underpinnings for each of these broad groupings of impairments may be different. The internationally agreed classification systems (ICD10 and DSM-IV) define ASD sub-categories such as 'atypical autism' and Asperger syndrome. However, the recent consultation process for DSM-V (January 2011; final publication expected May 2013) has removed sub-categories, emphasising the characterisation of ASD as a dimensional disorder, with considerable heterogeneity. Therefore, to ensure that this project is relevant for the future conceptualisation of ASD, subcategories will not be part of the conceptual framework.

One of the key features of the profile of children with ASD is their social communication difficulties, involving a range of different linguistic and social skills from receptive vocabulary, through expressive syntax to

pragmatic language skills, prosody (intonation) and social interaction. It will be important to consider measures that tap into these skills partly as outcomes in their own right, but also because of the potential role they play in mediating other outcomes. For example, there is evidence that pragmatic skills (such as social timing in interaction) are closely associated with particular types of behavioural profiles (Gilmour et al 2004; Ketelaars et al 2009). A child with ASD with good receptive or expressive language skills may be better able to respond to the type of behavioural interventions most commonly available in early years' settings.

Another aspect of complexity in the understanding of ASD is that its measurement is affected by developmental considerations, that is, children's impairments may look very different over time. For example, we cannot measure children's ability to make and keep friendships until they are of an age where that might be expected in typical development. However, there is general agreement on the core early impairments which may be observed, and this knowledge has been enhanced by recent studies of the early development of baby siblings of children with autism (who have an increased chance of themselves developing ASD). For example, Zwaigenbaum and colleagues (2005, 2009) reported unusual eye contact, a lack of visual attention, orienting to name, imitation, social interest and emotional affect, and heightened sensory-oriented behaviours. This combination of deficits has obvious consequences for development of relationships, early language and play. The complexity of understanding ASD is made even greater when considering the interaction between domains of development, and how deficits in one may impact upon another; for example, visual sensory overload may lead to avoidance which reduces opportunities for visual learning and social experience, leading to poor social skills. Thus the conceptual framework for the project will consider both the measurement of individual areas of functioning, which are likely to change over time, and also measures which bring together these separate areas into a more holistic assessment.

These challenges are particularly evident in educational assessment. In the Early Years Foundation Stage (to the end of Reception, introduced by DfES in 2007) the curriculum is organised into six areas of learning and development: personal, social and emotional development; communication, language and literacy; problem solving, reasoning and numeracy; knowledge and understanding of the world; physical development; and creative development. These areas are reflected in the Early Years Profile, an observational measure carried out on all children in England in their first year in primary school. It is possible to link children's performance on this assessment and the Key stage 1 and Key stage 2 National Curriculum assessments (SATS) at seven and eleven years, examining trajectories as well as outcomes and comparing profiles of children with ASD with those of other children. The tests and assessments used by those working in schools vary greatly from one local authority to another, and tools chosen depend on the training of those making the assessment (Luiselli et al 2001). For children with special educational needs, the P scales are a set of descriptions for recording achievement of pupils who are working towards the first level of the National Curriculum, and so offer an extended scale for the monitoring of educational achievement of children with ASD (Ndaji & Tymms 2010). In practice, teachers find it hard to distinguish between the different levels, and reliability is poor. It is important for this review to consider whether there are recommendations that can be made on tools that are relevant for monitoring progress of young children with ASD from pre-school and into the early years of their school career.

It is also important to detail other associated difficulties which are not unique to ASD but which nevertheless can often play a major part in children's development and the burden of care for families. These commonly include feeding and eating difficulties (resistance to lumps, faddiness about types and colours of food, etc), behaviour and sleeping problems. Children who lack adequate nutrition and sleep are likely to be bad-tempered and even more inflexible. Finally, as young children's development is intimately affected by their environment including the skills and resilience of parents and carers, it will be important to include consideration of the impact on the family.

Other research issues to be considered in the review must include variability in the focus and conduct of early intervention research in ASD and in the measurement of outcomes. These issues can be divided into

four main areas. First, there is the question of what should be the primary goal of intervention - a focus on reducing particular ASD impairments or overall severity, improving child functional outcomes, or quality of life for child and/or family? Each of these goals implies different conceptual and practical considerations and different targeted outcomes reflecting the competing priorities. Second, there is the question of the merits of measuring 'proximal' versus 'distal' outcomes. The value of proximal outcomes (i.e. close to the target of a particular intervention approach) is that they generally show better intervention effects than distal (i.e. more functional and generalised) outcomes. However an emphasis on proximal outcomes can be misleading in relation to the actual effectiveness of interventions. The third area for consideration has to do with external validity. The dilemma here is that subjective (particularly family-reported) measures are those with the greatest external validity, since it is the experience of children and families that interventions most want to improve; however, such ratings are prone to expectation and placebo effects within interventions. A final challenge is to define outcome measures that are responsive to change but also stable over the different settings that children experience. Aggregated measures (e.g. combining parent and teacher report with direct observation) can increase stability and reduce correlated measurement error. Sensitivity in measures is often limited in studies involving heterogeneous samples of children by floor and ceiling effects (floor effects when children are not ready to change, and ceiling effects when they have already mastered the skill). For all these reasons, the explicit definition of conceptually clear and metrically valid measures is a major challenge which will involve linking potentially different priorities and perceptions into a common purpose.

Unsurprisingly, no core set of outcome measures has yet been adopted in autism and registered with the Core Outcome Measures in Effectiveness Trials (COMET) initiative funded by the MRC Network of Hubs for Trials Methodology Research. The only review involving children with autism currently on the COMET database was authored by Professor K Williams, collaborator (Intravenous secretin for autism spectrum disorder, *Williams, K. W., Wray, J. J. & Wheeler, D. M.*, 2005). We would liaise closely with the COMET initiative, and believe that MeASURe will go a long way to enabling a core set of measures to be proposed for early intervention studies.

Types of measurement in use

The project will focus on tools for monitoring children's developmental progress, and outcome measures for evaluation of different types of intervention. For brevity, these will all be called 'tools' (and separately as monitoring tools, or outcome measures, as appropriate).

There are many types of tools currently in use, involving face to face assessment, observation or report, with varied measurement properties and levels of evidence for their reliability and validity.

Standardised norm-referenced assessment has the advantage of comparison with children of the same age, but may be misleading in assessment of young children with ASD. Their abilities may be underestimated by lack of cooperation with standardized testing, and they may have profiles which are dissimilar to typical development; for example, Hudry et al (2010) have shown that young children with ASD have an atypical language profile, where reported spoken language may not be underpinned by a level of understanding of language that would be expected in typical development. Some batteries of assessment have been developed specifically for children with ASD; for example, the Social Cognitive Evaluation Battery (Thiebaut et al 2010) is a direct psychological assessment, developed from existing tools and focused on skills in children with ASD and learning disability up to a developmental age of 24 months.

Direct observation has been used primarily to measure social interaction, especially parent-child interaction. This has the advantage of in-depth understanding of patterns of responsiveness, which may have long-term effects on language development (e.g. Siller & Sigman 2002, 2008). However, disadvantages include the short time-frame for the observation. As well as consequent questions of validity, there are almost as many coding schedules as studies, depending on the focus of interest.

Interviews have been used in the characterization of children's early development and current ASD characteristics (e.g. the Autism Diagnostic Interview-Revised), and in broad measurement of adaptive behaviour (e.g. Vineland Adaptive Behavior Scales), but in addition have been used to gather information on

additional difficulties such as behaviour problems, anxiety and sleep. Problems of measurement include a paucity of tools focused on behavior which are specifically validated for ASD (e.g. the Autism Comorbidity Interview-Present and Lifetime version (ACI-PL) is one such tool but validated only from 5 years of age, Leyfer et al 2006).

There are very many questionnaires used in studies of children with ASD, completed by parents, teacher and clinicians; again many have not been specifically validated for use in ASD and contain assumptions about patterns of typical development (e.g. standard Quality of Life measures do not ask about children's special skills and obsessive interests).

There are a number of summary measures which have been designed specifically to track progress in children with ASD. The Pervasive Developmental Disorder Behavioral Inventory (Cohen & Sudhalter 2005) and the Parent Interview for Autism-Clinical Version (Stone et al 2003) are comprehensive questionnaires for parents/teachers about child behaviours. The Autism Treatment Evaluation Checklist (Rimland & Edelson 1999; Magiati et al 2011) is a similarly comprehensive checklist. These tools will be of particular interest in the systematic reviews.

Pilot scoping study

In preparation for this application, an initial search of existing systematic reviews of intervention in ASD (8 Cochrane Collaboration reviews, 13 recent journal papers) was conducted (May/June 2011 by Livingstone, co-applicant) to gather information regarding tools and outcome measures that are commonly used, and to identify theoretically important gaps in the domains measured. Seventy-nine tools and measures were reported in the reviews (23 adaptive and maladaptive behaviour, 17 language/ communication, 13 ability, 8 sensory, 9 ASD-specific, 4 impact on family, 2 social interaction, 1 motor skills, 2 summary scales). (See linked documents in application form, spreadsheet of tools and numbered reviews)

Many of the reviews failed to discuss the relevance of the outcome domains, and the strengths and weakness of the included tools; those that did were relatively consistent in their recommendations for improvement. The most commonly cited outcome domain that review authors stressed was missing was quality of life (e.g. Nye & Brice, 2005; Williams, et al, 2010; Wheeler, et al, 2008). Other missing outcomes included "school readiness", independence and daily living skills (Nye & Brice, 2005), behavioural outcomes such as sleep disturbance, self-mutilation, attention and concentration problems (Williams et al., 2010), and parent satisfaction (Wheeler et al., 2008). Also mentioned was the need for qualitative research to determine which outcomes are "useful and relevant to consumers, clinicians and service providers" (Wheeler et al, 2008).

A key weakness mentioned in the reviews concerned ASD-specific tools, developed to aid diagnostic assessment, but used to monitor change even though not designed for this purpose (Williams, et al, 2005; Williams et al., 2010; Wheeler et al., 2008). Similarly, IQ has been used as a measure of change although designed to measure a 'stable' construct (Diggle & McConachie, 2002). Another unresolved question is what magnitude of change should be considered clinically relevant, for example, what parents consider as an important change (Williams, et al, 2005; Williams et al, 2010; Wheeler et al., 2008). Several review teams commented that included studies had measured outcomes using unpublished or non-standardised measures (Gold, et al, 2006; Ospina et al., 2008; Diggle & McConachie, 2002). Some reviews included studies focusing on anecdotal reports or ad hoc questionnaires created by the researchers for that specific study (Mulloy et al., 2010) and not adequately validated.

One prominent recommendation made throughout the different reviews was the need for a core shared battery of baseline assessments and outcome measures though the challenge of developing a single battery was recognised, due to the heterogeneity of children's difficulties and change with development (Elchaar et al, 2006). Others gave examples of measures that should be considered, including measures of intellectual ability, developmental abilities across domains, adaptive behaviour, communication skills,

severity of autism, play, social skills, challenging behaviours, rigidity, and other behaviours characteristic of children with autism (Reichow & Wolery, 2009; Williams et al., 2010). A few groups have in the past made recommendations of batteries of tools for measuring outcome in autism (e.g. Arnold et al 2000; Wolery & Garfinkle 2002), but these are not in use consistently across research groups and the measures considered in the first of these are not all applicable to young children (being focused on outcomes of medication trials for adults with autism and aggressive behavior). Other reviews of assessment batteries focus only on diagnosis (e.g. Ozonoff et al 2005).

Summary of steps in planned programme of reviews and evidence synthesis

Stage 1:
Identify the main constructs of interest.
Determine outcomes of importance by an initial overview of qualitative literature on parents' priorities, and by consultation with parents and individuals with ASD. Determine outcomes and issues of importance for clinicians and educators.
Stage 2:
Identify the tools that are used to measure these constructs.
Determine the range of tools used in epidemiological, longitudinal and intervention studies through a wide systematic review, and use information from stage 1 as part of the judgement of the importance of domains and tools to be studied further.
Stage 3:
Identify the tools with the best measurement properties.
Conduct a series of 'domain' specific systematic reviews, using the COSMIN protocol.
Stage 4:
Propose a battery of tools for future use in outcome studies and monitoring.
Through discussion and consensus with parents, professionals, researchers in ASD and method experts in the team, a battery of robust tools will be identified and recommendations made for further developments required.

Working groups
The work for the project will be led by three sets of working groups.
The **Methods Advisory group** will draw conclusions about the most robust measures. The group includes methods experts in reviewing and psychometrics, and ASD child research specialists who know many of the tools used in ASD studies and can judge their relevance to current child development and ASD theory, current and future research questions in ASD, and the likely feasibility of the tools for use in monitoring and outcome studies focused particularly on young children with ASD.
Led by Charman and Parr (including Terwee, McColl, Macdonald, Morris, Law, Williams, Pickles, Rodgers, Green, Simonoff).

Three **Parent Advisory groups** will be set up from the start, in Exeter, Newcastle and South London, to advise on the most important constructs to be measured, and later to consider the preliminary conclusions from the measurement reviews. Led by Morris, Garland, Le Couteur, Gringras.

A **Professional Reference group**: experts in education, psychology, paediatrics, speech and language therapy, psychiatry, and social work, to consider the constructs important in monitoring children's progress, and what batteries of measures might offer in clinical/educational practice. Led by Le Couteur and Jones (including Beresford, Baird, Gringras, Law, Macdonald).

An early face-to-face meeting of all the co-applicants will be held in order to confirm the planning of the project in detail. This will include a discussion of the likely important domains of measurement and how the

first 2 stages of the project will proceed. Thereafter, coordination between the working groups will be key. The CI, McConachie, will participate in all discussions of the Methods Advisory group and the Professional Reference group. All group meetings by teleconference will be transcribed. The reviews will be carried out by two systematic reviewers (SRs), who will cross-check all data extraction, under supervision of McConachie, Rodgers and Macdonald. Full notes will be taken in parent advisory groups, and shared with the full research team. McConachie, Garland, Le Couteur, Parr and Morris will take particular responsibility for ongoing review of themes across groups, and reconciling conflicting priorities.

Methods

Stage 1: Identify the main constructs of interest.

Tools and outcome measures must have face and content validity for both families and professionals in order to be credible and acceptable. That is, the outcomes assessed by tools and measures must address constructs that families and professionals think are pertinent and most important, and the way in which the construct is measured makes sense to them. Therefore, to capture outcomes of importance for parents and carers of children with ASD, we will conduct a rapid scoping review of qualitative research on parents' priorities, consult with young people with ASD, engage three parent advisory groups, and run a survey of professionals' views. The analysis and summary of information from these sources will be used in the quality filters applied in the data extraction from papers in Stages 2 and 3, and also will be included in the evidence synthesis in Stage 4.

Scoping review of qualitative literature (led by Beresford, Morris, Green)

Question: **What outcomes (child specific as well as parent/family outcomes) do parents of children with ASD perceive as important?**

Search strategy: A systematic search will be conducted using MEDLINE, CINAHL and PsychINFO (via OVID). The papers will focus on parents and/or carers of young children with ASD, and will be in English.

Condition under investigation/review (ASD terms) AND Qualitative study types (various) OR Terms referring to parents' experience (Petticrew & Roberts 2006). Decisions on which papers to include will be taken by the SRs in collaboration with the three co-applicants leading this piece of work.

Data extraction This will focus on identifying themes which concern parents' hopes for their children, experience of assessment of their children, and their priorities for intervention for and education of their children. The aim is not to assess the quality of the papers/studies, but instead to map important themes. The SRs will cross-check the themes identified by re-reading the papers. Themes are likely to include: maintaining self-esteem; managing personal safety; independence in self-care; making and keeping friends (Beresford et al 2007).

Data synthesis In order to present an overview to the parent advisory groups and the research team, key findings, quotes and concepts from each paper will be tabulated so that they can be explored and compared, and overarching themes identified (Popay et al 2006). This work will also build on the findings of focus groups, and the construction of a measure of outcomes of importance to parents of preschool children with autism, undertaken previously as part of the Preschool Autism Communication Trial (Green et al 2010) and led by Green and Le Couteur (co-applicants)

Consultation with young people with ASD

Children with ASD under the age of 6 years will not be able to comment directly on measurement issues. However, we will seek advice from young people with ASD who are older and able to comment at least on their own lives, and what has mattered to them in relation to their experiences, their own characteristics, and what has made them happy/unhappy. These reflections will highlight important outcomes that should be measured. In addition, we will devise a series of illustrated vignettes about younger children with ASD as a focus for discussion to try to elicit what are the main skills and difficulties that matter to young people;

vignettes will be derived from real examples given by parents from the Dasl^{n}e and ASD-UK databases, and clinical settings.
These consultations will be with (a) able young people aged 8 to 16 years who attend a Social Inclusion Project in Newcastle (facilitated by Garland, co-applicant), and (b) University students with Asperger syndrome who attend a group at Birmingham University (facilitated by Jones, co-applicant). The sessions will be audio-recorded, transcribed, and analysed for themes.

Consultation with parents and carers
The third strategy for identifying important constructs will be through direct consultation with parents. Each **Parent Advisory group** will be chaired by a parent and a professional (the parent co-leads for Exeter and South London are to be identified). Advertising for parents to join the groups will be through the PenCRU (Peninsula Cerebra Research Unit) Family Faculty (Exeter, directed by Morris), Dasl^{n}e (the Database of children with ASD living in the North East, directed by McConachie), and ASD-UK (family research database, directed by Parr). The discussions will take an open agenda, but with certain topics covered. This will include parents' priorities for their children's future development; parents' experiences and their report of their children's experiences of assessment processes (after initial diagnosis); any involvement they and their child have had in research studies; and what would make particular tools more acceptable/feasible (including multi-media presentation of tests and questionnaires). In addition, the main findings of the qualitative scoping review will be presented for discussion and validation. The group co-chairs will make a summary of the main themes, led by Garland and Le Couteur.

Consultation with professionals
The fourth approach to establishing the constructs of interest will be through a short survey of professionals' priorities. A **Professional Reference** group within the research team will construct a brief electronic survey (delivered using the package SurveyMonkey, an approach we have previously used successfully in similar populations) to be sent to the leads in all the 240 Child Development Teams in UK. This listing is held by the British Academy of Childhood Disability, the neurodisability arm of the Royal College of Paediatrics and Child Health, and was updated in 2010 by Parr (co-applicant). The survey will also be sent to the Early Years Advisors in every education authority in the country. The survey will ask structured and open questions about the most important constructs and domains of development (including function and participation) to monitor in children with autism; the properties of tools (including cost, time, personnel required) which allow regular monitoring of children by professionals; ways in which assessment processes may need to be modified for young children with autism (McConachie 1995; Koegel et al 1997; Akshoomoff 2006); and how professionals use data from monitoring children's progress and share this with parents. This consultation will be led by Parr.

Analysis of these four sources of evidence will be made by the group leads and the CI (McConachie) early in the project, through review of the findings from each source, and consensus discussion by teleconference. The domains identified and the important issues in measurement will inform the design of the data extraction forms for the reviews, as well as the synthesis in Stage 4.

Stage 2: Identify the tools that are used to measure these constructs.

Review questions:
What tools are in use for measuring and monitoring aspects of autism?
How frequently have these tools been used in identified studies?

Dates of searches

We will include studies published in the past 20 years (i.e. from 1992) which coincides with the publication of the current international classifications, ICD-10 and DSM-IV (APA 1994; WHO 1992).

Types of Studies

We will review all studies which include at least 50% children with ASD to discover the tools used to measure children's characteristics. The search will be designed not to miss any potentially relevant study at this initial stage; thus, it will include tools primarily designed for screening and tools designed for diagnostic assessment in so far as some of these have also been used to monitor progress and in outcome evaluation. Though the main focus is on tools used with children up to about 6 years old, searches will include children of primary school age in order not to miss tools used in monitoring progress after age 6, and to include tools used in monitoring educational progress.

We will include:

- all relevant randomised and quasi-randomised trials of early interventions for children with a diagnosis of ASD.
- epidemiological studies of children with ASD,
- case-control studies,
- descriptive cohort studies, including studies of baby siblings of children with autism, which provide information on tools to monitor developmental progress and follow early markers of ASD.

Types of Participants

Child participants with a 'best estimate' clinical diagnosis of an ASD, including autism, ASD, atypical autism, Asperger syndrome, and PDD-NOS, according to either ICD-10 or DSM-IV (APA 1994; WHO 1992) criteria. Use of a particular diagnostic tool such as the Autism Diagnostic Observation Schedule (ADOS) or the Autism Diagnostic Interview (ADI-R) will not be required.

Children with ASD and another medical condition, and children with ASD and co-morbid conditions will be included.

Types of Measurement included

1. Direct measurement of Child ASD symptoms by trained assessor
2. Direct measurement of developmental skills, i.e. language, cognition, fine and gross motor skills, by trained assessor
3. Observational measures of social interaction skills
4. Interview or self-completed (parent, teacher or other professional) questionnaire report of Child ASD symptoms.
5. Interview or self-completed questionnaire report of developmental skills, i.e. language (vocabulary), adaptive skills, with/by parent, teacher or other professional.
6. Interview or self-completed (parent, teacher or other professional) questionnaire report of associated problems: including behaviour, aggression, sleeping, eating, toileting, anxiety, hyperactivity and others identified through parent consultation
7. Idiographic measures focussed on particular behaviours (e.g. goal attainment scaling, target behaviours)
8. Measures of impact on parent or family.

Types of Measurement not included

1. Economic impact on home and family.
2. Experimental tasks and measures: e.g. barrier tasks, reaction time.
3. Biophysical measures; medical investigations

Search Methods for the Identification of Studies in stage 2

Based on: Cochrane Methods (Higgins & Green 2011), and Centre for Reviews and Dissemination guidance (2009). Searches conducted on title and abstract. Published and unpublished studies will be considered, with no language restrictions. Abstracts will be translated into English, and the whole paper if judged relevant.

Search strategy design:

Study type	*Disease/disorder*
Before and after study Cohort study Comparison group Comparative study Controlled trial Epidemiolog* Longitudinal study Prevalen* Quasi-experiment* RCT	ASC ASD Asperg* Atypical autism Autis* Autism/childhood childhood schizophrenia communicat* Kanner* language delay* PDD/PDD-NOS pervasive developmental disorder speech disorder* *mesh terms* Autistic Disorder Autistic Disorder/diagnosis Autistic Disorder/psychology Autistic Disorder/therapy Child Development Disorders, Pervasive Child Development Disorders, Pervasive/diagnosis Child Development Disorders, Pervasive/psychology
Population	*Treatment outcomes*
Child* Elementary school Infant* Junior school Kindergarten Nursery school Paediatrics/Pediatrics Pre-school/preschool Primary school Special needs Toddler	Behavio*r intervent* Non-verbal program* rehabilit Social interaction therap* train* treat* Verbal *mesh terms* Child Development Disorders, Pervasive/therapy Outcome Assessment (Health Care) Outcome Assessment (Health Care)/methods Outcome Assessment (Health Care)/standards Treatment Outcome

A similar search strategy was piloted in May 2011 on PubMed in preparation for this application. This search initially returned 19,545 results. Many were irrelevant to the topic of interest; for example, the abbreviations ASD and ASC, as well as referring to autism, are used to refer to "atrial septal defects" of the heart, "acute stress disorder" and others, therefore exclusion terms will be added. The first 1000 of these results were inspected, taking out studies such as post-mortem brain studies, animal models of autism, and general population studies of children on topics such as chronic pain or obesity, etc. Forty three percent of entries appeared sufficiently relevant to be inspected further. The task for this initial search will therefore be

large. We will use strategies to make the task achievable and avoid duplication, such as identifying systematic reviews in ASD from this search, from collaborators (including diagnosis and prognosis systematic reviews in progress by Williams et al) and by building on the search we have already undertaken and, where appropriate, updating those review searches; and filling in gaps for particular types of studies such as those describing cohorts that have not yet been synthesised in systematic reviews. Stage 2 can be completed in the time frame as data extraction will be relatively simple and focussed on individual tools identified (see below).

In all searches appropriate truncations will be included, and the same search terms will be used for all databases. Databases which will be searched include: PubMed, Medline, EMBASE, PsycINFO, The Cochrane Controlled Trials Register, Cochrane Library (includes DARE, NHS EED, HTA, CENTRAL, CDSR), ERIC (Educational Resources Information Centre), CINAHL (Cumulative Index to Nursing and Allied Health Literature), Dissertation Abstracts International, Social Sciences Abstracts, Sociological Abstracts, Linguistics and Language Behaviour Abstracts, National Research Register, LILACS (Latin American and Caribbean Health Sciences Literature), HELLIS (Health Science Libraries across Asia), Japanese Science and Technology database, Japanese Clinical Trials Database, Australasian Medical Index, CBM (Chinese Biomedical Literature Database), Oxford Patient Reported Outcomes Database, Applied Social Sciences Index and Abstracts (ASSIA), Autism Data, Digital Education Resource Archive (DERA), Health Management Information Consortium (HMIC), metaRegister of Controlled Trials (including ClinicalTrials.gov), PapersFirst, Proceedings (OCLC), ScienceDirect, TRIP, UK Clinical Research Network, Web of Knowledge, WorldCatDissertations, Zetoc.
Websites of prominent autism research funding charities will be searched for reports of studies: including Autism Education Trust, Autism Research Centre, Autism Research Institute, Autism Society of America, Autism Speaks, Autism-Europe, Interactive autism network , Research Autism, UK Autism Foundation, Autistica.
Other sources of information will be investigated using a hand search, including bibliographies of related review papers, reference lists of key articles, conference proceedings and the output of key journals in the field (including *Journal of Autism and Developmental Disorders*, *Autism*, *Autism Research, Autism Research and Treatment, Research in Autism Spectrum Disorders, Focus on Autism and Other Developmental Disabilities,* and *Journal of Child Psychology and Psychiatry*).
Experts in the field will be contacted by email to identify submitted, unpublished and ongoing intervention and longitudinal studies of relevance.

It is recognised that, though we will attempt to identify unpublished material, the review will inevitably be affected by outcome reporting bias (Sinha et al 2008a).

An initial check was made among the pilot search results (above) for eight key papers. It indicated that the proposed search strategy is comprehensive and thus will return the vast majority of relevant results.

Selection of Studies

All citations sourced from the search strategy will be transferred to EndNote. Initial screening of titles and abstracts by both SRs will eliminate all those citations obviously irrelevant to the topic. Guidelines for exclusion and inclusion will be drawn up, with inclusion covering terms for likely associated disorders, and likely important topics (e.g. language development, adaptive behaviours, social skills, emotional intelligence, parent-child relationships, neuropsychology, types of common interventions in ASD, etc). Thereafter, following appropriate training, the SRs will assess and select studies for data collection, from the identified group of superficially relevant studies. All papers from these searches will be retrieved. In the event of a disagreement about inclusion, a resolution will be reached in discussion with HM and GM, following inspection of the full paper.

Data Collection

The SRs will independently extract data, using web-based data extraction forms designed and piloted for the

purpose. Extracted data will consist of year of publication, study location, methods, diagnostic and age description of participants, number of participants, type of study.

For trials, focus of intervention; for observational/epidemiological study, focus of longitudinal assessment.

A further set of indicators will be noted, as a filter to enable the choice of tools at Stage 3.

For each tool:

- Was this the primary outcome for the study?
- By whom was it measured/reported?
- When/how often was it measured/reported?
- Was the tool developed ad hoc for the study?
- What domains did the tool intend to capture?

Quality indicators (Brundage et al 2011; Sinha et al 2008a,b):

- Was evidence (citation of previous publication would suffice) provided of the tool's validity in general (Yes/No)?
- Was evidence provided of the tool's validity with children with ASD (Yes/No)?
- Was evidence (citation of previous publication would suffice) provided of the tool's reliability in general (Yes/No)?
- Was evidence (citation of previous publication would suffice) provided of the tool's responsiveness to change (Yes/No)?
- Did the tool measure one or more of the priorities identified by service users (listed)? (Yes/No)?

Evidence Synthesis

Summaries of findings of numbers of tools, their apparent frequency and quality in use, and their domains, will be presented by the SRs to the **Methods Advisory** group within the research team. The group will consider this evidence, and the priorities identified in stage 1, to draw conclusions about an initial 'sifting' of domains, types and tools to be considered further in stage 3, recognizing that some tools cover multiple domains. Their preliminary report will also identify important constructs (identified in stage 1) which are not well served by existing tools.

The searches will be updated, and data extracted from any additional papers uncovered, toward the end of the first year of the research, to ensure that information about identified tools is complete for the final report to HTA.

Stage 3: Identify the tools with the best measurement properties.

Review questions:

What evidence is there for measurement properties of tools in the selected domains?

What evidence is there for tools being both developmentally appropriate for young children and sensitive enough to measure change over time?

What evidence is there for the feasibility of using particular tools, and their practical usefulness/utility in educational/clinical practice as well as in research?

A systematic review of measurement properties is intended to give a clear and comprehensive overview of the measurement properties of a number of tools. A series of such reviews, organised by key domains of function, will enable us to come to a conclusion about the best tools available for a particular purpose, based on research evidence. The focus in Stage 3 is on searching for studies on the measurement properties of tools which were identified in Stage 2 for further consideration, covering key domains in the monitoring of children with ASD over time and in measuring the outcomes of interventions.

Search strategy

In order to improve the success of searching for studies of measurement properties a search filter developed by the COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) group will be applied (Terwee et al 2009). The search strategy will be designed by our information specialist (Robalino) with Terwee (co-applicant) and will use search terms for the following

characteristics: construct of interest, and target population. The search terms will not be restricted to the type of measurement instruments because many studies do not use specific terms and there is a high risk of missing relevant articles. However, an instrument filter developed by the Oxford PROM group can also be used with PubMed for finding studies on patient-reported outcomes (Morris, personal communication, and http://phi.uhce.ox.ac.uk/newpubs.php). An additional search will be performed including the names of the instruments found in stage 2, combined using the AND conjunction with terms for the target population and measurement properties. We will also check the references of articles included in the review to search for additional relevant studies. No time limit will be set because older literature on measurement properties is still relevant.

Eligibility criteria

We will use the following inclusion criteria for each review:

1. the tools should aim to measure the domain of interest.
2. the study sample should concern the target population of interest, children with ASD aged up to 6 years (though they may also be used with older children)
3. the study should concern one or more of the types of measures specified in stage 2.
4. the aim of the study should be the development of a measurement tool or the evaluation of one or more of its measurement properties. Studies that only focus on interpretability, e.g. the determination of minimal important change, can also be included.
5. the study should be published as a full text original article.

The two SRs will independently assess titles, abstracts, selected full-text articles and reference lists of the studies retrieved, and in case of disagreement a third member of the team with most knowledge of studies in that domain will make the decision regarding inclusion of the article.

Evaluation of methodological quality

We will then assess the methodological quality of the studies on measurement properties identified using the COSMIN checklist (Terwee et al 2007). The checklist has 10 'boxes' or subscales (internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, responsiveness, interpretability) with standards for how each measurement property should be assessed. Each item is scored on a 4 point rating scale (poor to excellent) and an overall rating for the methodological quality of each study is determined.

Data extraction will include: characteristics of the tools, characteristics of the study populations, results of the measurement properties, and evidence on interpretability including acceptability (e.g. completion rates of questionnaire items). Content comparison at the item level will be conducted between similar measures, and also where appropriate in relation to standards such as the International Classification of Functioning, Disability & Health. Particular note will be taken of where studies differentiate the measurement properties of tools between different ages, or abilities of children. Wherever possible published manuals for tools will also be obtained to add to information on psychometric properties. The time, energy, costs, personnel or other resources required will be noted in the data extraction as indicators of feasibility (Fitzpatrick et al 1998); as well as where tools have the potential to direct subsequent intervention through being arranged in developmental steps.

Data synthesis

The next step is to evaluate the quality of individual tools, through combining the overall evidence for their measurement properties. Combining results of different studies on a measurement property of a tool is only possible when the studies are sufficiently similar with regard to study population, setting etc. The evaluation uses established criteria for good measurement properties (rated on a 5 point scale, from unknown or conflicting evidence through to strong) (Terwee et al 2007)

Evidence synthesis

Summaries from the reviews will be presented by the SRs regularly throughout the process to the **Methods**

Advisory group. At this stage of evidence synthesis, the group will consider the quality of the tools which emerge for potential recommendation of the strongest, using the levels of evidence determined by the reviews. The number of studies, the quality of the studies, the results of the studies and the consistency of the results will all be taken into account.

We will register the reviews in stage 3 with PROSPERO. PROSPERO was set up by NIHR in 2011 as an online facility to register systematic reviews for research about health and social care from around the world. Access is free and open to the public.

Stage 4: Propose a battery of tools for future use in outcome studies and monitoring.

Based on the evidence, what battery of tools/measures should be recommended to and monitor progress and measure outcome in children with ASD aged up to 6 years?

Evidence synthesis will take place throughout the process, as outlined above. In stage 4, each of the three working groups will take an overview of the evidence, and preliminary conclusions about potential tools to recommend for a battery. One important consideration will be how to reach a balance in recommendations between robust psychometric aspects of measurement, and importance and usefulness to parents, teachers and other professionals. The synthesis will help identify the characteristics of a good tool to be used for monitoring or outcome research in childhood ASD, even if such tools do not yet exist in all domains.

A final day **consensus meeting** will be held with all co-applicants and representatives from the parent advisory groups. The aim is to draw conclusions about the suitability of tools for use in monitoring progress of children with ASD, and in evaluating trials of intervention. In preparation for the day, a grid organised as domains by type of measure will be produced by the CI and one SR, drawing on all the group discussions, with identified robust tools entered, in order to map where a battery could be constructed. The summary information will include the focus and age range of each robust tool. This grid will also highlight where domains that should be measured are ill-served. Several grids may be constructed, i.e. different batteries for studies with particular objectives (e.g. sleep problems, play skills, etc) with a common core.

The format of the day will depend somewhat on the findings of the reviews. If many tools are available, we plan to develop consensus by using nominal group technique (Jones & Hunter 1999), facilitated by Morris (co-applicant), with small groups considering sets of information, and then ranking choices, for discussion in the larger group. For this we will utilize Turning Point software, and voting devices, available within Newcastle University.

The output of the consensus and of the project will be recommendations of a battery of robust and appropriate tools, presented by domain (e.g. social interaction, sensory symptoms) and type (e.g. standardized assessment, parent questionnaire) with additional recommendations for particular topics (e.g. sleep). Finally, recommendations for future research will also be made, to include ways of improving the feasibility and acceptability of tools as well as important domains for further development.

Dissemination

Our strategy has several strands:

a) *Engagement with parents of young children* through summaries (prepared with the parents who contribute to the project) for websites and publications accessed by parents of young children with autism and ASD support organisations. Summaries of evidence will be prepared for Research Autism (website which summarises research findings for parents and others) and for other websites consulted by parents. An article

will be submitted for the NAS 'Communication' magazine and the members of each parent advisory group will be invited to contribute.

b) *Engagement with the community of clinical and educational professionals* through publications in relevant journals and on websites where good practice and research are promoted. Individual co-applicants are active in clinical networks in various parts of the country, and we envisage a number of audit projects and updates to existing practice guidelines both in UK and internationally arising from the project. Furthermore, several individuals involved in the project are part of the development group for the NICE guideline on management of autism in children and young people (with an intended publication date of November 2013) allowing mutual information-sharing and influence. Following the report for HTA, papers will be prepared for professional peer reviewed journals summarising the findings concerning the reviews of measurement properties. Research digests will also be prepared for websites and journals consulted by clinicians and educators (e.g. the British Academy of Childhood Disability; Good Autism Practice).

c) Engagement with the international community of researchers*. By publication on COMET and PROSPERO the work will receive wide recognition.* Presentations will be made at the US International Meeting for Autism Research and at Autism Europe (2013), as well as at parent and professional meetings within UK. We would propose not only presentations, but also discussion fora at the two conferences in order to explain the merits of adopting a core data-set. We would publish the findings of the main reviews in peer reviewed journals. Co-applicants are active in a number of international consortia, including an EU COST action to improve science in autism early intervention.

Thus, the aim of the dissemination is to see the conclusions of the project put into action, through active promotion by the team of co-applicants.

Project timetable and milestones (see linked document in application form, GANTT chart)

April - July 2012

Initial planning meeting to confirm strategies and questions

Recruit parents to the parent advisory groups. Hold groups in early July.

Survey of professionals' views and priorities

Consult with young people who have ASD.

Conduct searches and summarise review on parents' priorities (stage 1)

Conduct searches in stage 2, exclude irrelevant articles, obtain full text

Devise and test data extraction tool for stage 2, informed by stage 1 conclusions

August - October 2012

Data extraction for stage 2

Present review of tools in use to Methods Advisory group.

Decision on number and structure of reviews for stage 3

November 2012 - June 2013

Searches and data extraction for stage 3 reviews

Methods Advisory group considers reports on a bi-monthly basis.

Update searches and data extraction for stage 2; add to stage 3 if required.

Initial synthesis of evidence from stage 3.

July - September 2013

Parent groups and Professional Reference group consider initial synthesis.

Consensus day conference to consider recommendations on suitability of tools for monitoring children's development, and for outcome research.

September - November 2013

Final report for HTA

Papers and articles prepared

Expertise

This is a uniquely strong team to conduct this important review, deliberately a large team with international representation, in order to achieve a fully authoritative conclusion for HTA. McConachie (CI), Macdonald, Livingstone, Law, McColl and Williams (collaborator) are Cochrane Collaboration systematic reviewers, with Macdonald leading the Developmental, Psychosocial and Learning Problems Review group. Williams has conducted several systematic reviews in child ASD. Jones has conducted 3 major reviews of educational services for children with ASD for UK government, Parr has reviewed intervention studies in autism for BMJ Clinical Evidence. McColl and McConachie will supervise the information specialist designing and conducting the searches. McConachie and Rodgers (who teaches systematic reviewing), and Macdonald will supervise the two SRs (one being Livingstone) conducting the detailed reviews. Terwee, McColl, Beresford and Morris are experts in reviews of measurement properties, particularly patient-reported outcome measures, and Terwee is the originator of the COSMIN review methodology for assessment of measurement properties. Pickles is the leading international statistician in the field of ASD, particularly in understanding the measurement properties of tools. Le Couteur, Green, Simonoff are child psychiatrists; Baird (collaborator), Parr, Gringras, and Williams are paediatricians; Law is a speech and language therapist; Jones is a teacher and educational psychologist; and Charman and McConachie are clinical psychologists, all involved in providing professional services to children with ASD and their families, as well as leading research in child ASD. Garland, Le Couteur, Parr and McConachie (supported by McColl) work together on a number of studies in ASD in the North East of England. Many of the co-applicants have been involved in the leading UK child ASD studies of the past decade: for example, Green led the Preschool Autism Communication Trial (published 2010) which involved Le Couteur, Charman, Pickles and McConachie as PIs; Charman, Baird, Simonoff and Pickles led the SNAP study of prevalence of ASD (published 2006) and subsequent follow-up of the children. Parr and Le Couteur now lead the UK-based International Molecular Genetic Study of Autism Consortium, which includes Pickles, Green and McConachie. Le Couteur led the National Autism Plan for Children (2003), and Baird is the chair of two NICE child guidelines groups (diagnosis, and management).

Current related studies by this team include: QUEST: comparing two measures of emotional and behavior problems in children with autism (Baird); Cochrane review of diagnostic tests for ASD in preschool children (Williams); Cochrane review update of parent-mediated early intervention in ASD (McConachie); NIHR HSR Programme: 'What outcomes of NHS care should be measured for children with neurodisability?' (Morris, Beresford, shortlisted & decision pending on full application)

In addition the proposed programme of reviews will have synergy with a number of initiatives for which the team are responsible: COST ESF action: 'Enhancing the scientific study of early autism: A network to improve research, services and outcomes' (Charman, McConachie); BASIS: Studies of the baby siblings of children with autism, to identify precursors of autism characteristics in development (Charman – who is the BASIS representative on the US Baby Sibs Research Consortium, with Parr, Le Couteur, Green – who leads the current BASIS intervention study)

Service Users

Parents of young children with ASD are involved throughout the proposed programme, and we will also consult with young people who have ASD. The format of the proposal has been influenced in particular by the experience of Morris and Beresford in designing studies with parents' and young peoples' groups of the properties and key topics for outcome measures in neurodisability.

Garland and Jones will facilitate consultation with young people with ASD who are able to comment on the desired outcomes that matter to them now and in the past.

Recruitment of parents to advisory groups will be facilitated by existing partnerships led by McConachie, Morris and Parr. Garland is a parent of a boy with ASD, and has wide connections with networks of parents through running the NAS Resource Centre in Newcastle. She has also served on the steering group of a recent evaluation of intervention, and is a co-applicant on a survey of parents' views about genetic testing. Three parent advisory groups will be involved from the start of the project, to advise on the priority issues in assessment and measurement of young children with ASD, including impact on the family. Each group will be co-led by a parent and a professional, and the parent leads will also attend the consensus day to be held at the end of the project.

Thus the outcomes will be a considered synthesis of robust measurement and service users' priorities for appropriate monitoring of children's progress.

Justification of support required

The likely size and number of the reviews requires that two trained and experienced systematic reviewers are employed for 12 months full time(one senior at point 32 and one more junior at point 24/25),with one for a further 5 months to help write up the reviews for publication. Each requires a computer, and external hard drive to store documents. In addition, an information scientist will be employed part time for 3 months to design and run the searches. The co-applicant costs are mainly at 1% (2.5 working days) or at 4% (9 working days) for those co-leading advisory groups or supervising staff. The CI is costed at 10% as she is involved in all aspects of the research. The team will make use of technology to minimise impact on the environment. Papers to be stored for data extraction will be saved on external hard drives, and printed/photocopied as little as is practicable. The survey of professionals will be done by email and SurveyMonkey. Travel is kept to a minimum, except where team-working and fruitful discussion will be enhanced by a face to face meeting. The parent advisory groups and consultation with young people with ASD will be in widely dispersed areas of the country, but will each have minimal travel costs. The consultations with young people will be supported by additional funds from the Service User and Carer budget of the Mental Health Research Network. There will be a one-day planning meeting of all co-applicants at the start to enable efficient and effective team working, and a consensus day meeting at the end to conclude the work. Full use will be made during the project of email communication, web-based applications, teleconferencing and videoconferencing to enable the advisory groups to work together.

Funds are included for inter-library loans especially for unpublished documents. Where the manuals for tools are not already in the possession of the co-applicants, they will require to be purchased. The conference attendance will allow discussion of the findings and recommended batteries of tools with the leading ASD researchers in US and Europe.

Linked documents in application form

Spreadsheet of tools identified in pilot scoping study of 21 reviews in child ASD, with numbered table of review titles.

GANTT chart to show timing of stages and working group activity.

References

Akshoomoff, N. Use of the Mullen Scales of Early Learning for the assessment of young children with Autism Spectrum Disorders. *Child Neuropsychol*, 2006; 12: 269-277.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P., Thurm, A., Shulman, C., et al. Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology,* 2007;*75*(4): 594–604.

APA *Diagnostic and statistical manual of mental disorders*. 1994. Washington: American Psychiatric Association.

Arnold, L.E., Aman, M.G., Martin, A., Collier-Crespin, A., Vitiello, B., Tierney, E., Asarnow, R., et al. Assessment in multisite randomized clinical trials of patients with autistic disorder: the Autism RUPP Network. Research Units on Pediatric Psychopharmacology. *J Autism Dev Disord.*2000; 30(2):99-111.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D. and Charman, T. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The special needs and autism project (SNAP). *Lancet* 2006; 368(9531): 210-215.

Baron-Cohen, S., Scott, F.J., Allison, C., Williams, J., Bolton, P., Matthews, F.E., & Brayne, C. Prevalence of autism-spectrum conditions: UK school-based population study. *Br J Psychiatry* 2009*;* 194(6):500-9.

Beresford, B., Tozer, R., Rabiee, P. & Sloper, P. Desired outcomes for children and adolescents with autistic spectrum disorders. *Children & Society*, 2007; 21: 4-16.

Brugha, T.S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., Bebbington, P., Jenkins, R., & Meltzer, H. Epidemiology of Autism Spectrum Disorders in adults in the community in England. *Archives of General Psychiatry.* 2011; 68(5):459-465

Brundage, M., Bass, B., Davidson, J., Queenan, J., Bezjak, A., Ringash, J., Wilkinson, A., & Feldman-Stewart, D. Patterns of reporting health-related quality of life outcomes in randomized clinical trials: implications for clinicians and quality of life researchers. *Quality of Life Research,* 2011*;* 20: 653-664.

Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care.* University of York, 2008

Charman, T. 'Glass half full or half empty? Testing social communication interventions for young children with autism (invited commentary)', *J Child Psychol Psychiat* 2011; 52: 22-23.

Charman, T., Taylor, E., Drew, A., Cockerill, H., Brown, J.-A. & Baird, G. Outcome at 7 years of children diagnosed with autism at age 2: predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *J Child Psychol Psychiat*, 2005; 46: 500–513.

Cohen, I. L., & Sudhalter, V. *The PDD Behavior Inventory*. 2005. Lutz, FL: Psychological Assessment Resources, Inc.

Diggle, T.J. & McConachie, H. Parent-mediated early intervention for young children with autism spectrum disorder. *Cochrane Database of Systematic Reviews* 2002, Issue 2. Art. No.: CD003496. DOI: 10.1002/14651858.CD003496.

Elchaar, G. M., Maisch, N. M., Augusto,L. M., Wehring, H. J. Efficacy and safety of naltrexone use in pediatric patients with autistic disorder. *Annals of Pharmacotherapy*. 2006; 40(6):1086-1095.

Fitzpatrick, R., Davey, C., Buxton, M.J., & Jones, D.R. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess.* 1998; 2(14):1-74.

Gilmour, J., Hill, B., Place, M. & Skuse, D.H. Social communication deficits in conduct disorder: a clinical and community survey. *J Child Psychol Psychiat* 2004; 45: 967-978.

Gold, C., Wigram, T., & Elefant, C. Music therapy for autistic spectrum disorder. *Cochrane Database of Systematic Reviews* 2006, Issue 2. Art. No.: CD004381. DOI: 10.1002/14651858.CD004381.pub2.

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P., Le Couteur, A., Leadbitter, K., Hudry, K., Byford, S., Barrett, B., Temple, K., MacDonald, W., Pickles, A. and the PACT Consortium. (2010) Parent-mediated communication-focused treatment for preschool children with autism (MRC PACT): a randomized controlled trial. *The Lancet,* 375, 2152-2160.

Happé, F., & Ronald, A. The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychol Rev.* 2008; 18(4):287-304.

Higgins, J. P. T. & Green, S. Cochrane handbook for systematic reviews of interventions 5.1.0. *The Cochrane Library*. 2011. Chichester, UK: John Wiley and Sons.
Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., Charman, T. & the PACT Consortium. Preschoolers with autism show greater impairment in receptive compared to expressive language abilities, *Int J Lang Communication Dis* 2010; 45: 681-690.
Jones, J. & Hunter, D. Using the Delphi and nominal group technique in health services research. In Mays N & Pope C, eds. *Qualitative Research in Health Care.* London: BMJ Books; 1999.
Ketelaars, M.P., Cuperus, J., Jansonius, K., & Verhoeven, L. Pragmatic language impairment and associated behavioural problems. *Int J Lang Communication Dis.* 2009; 28: 1-17
Koegel, L.K., Koegel, R.L. & Smith, A. Variables related to differences in standardised test outcomes for children with autism. *J Autism Dev Disord* 1997; 27: 233-403.
Knapp, M., Romeo, R., & Beecham , J. Economic cost of autism in the UK. *Autism,* 2009; 13 (3): 317-336.
Leyfer, O., Folstein, S., Bacalman, S., Davis, N., Dinh, E., Morgan, J., Tager-Flusberg, H. & Lainhart, J. Comorbid psychiatric disorders in children with autism: interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, 2006; 36: 849-861.
Luiselli, J.K., Campbell, S., Cannon, B., DiPietro, E., Ellis, JT, Taras, M. & Lifter, K. Assessment instruments used in the education and treatment of persons with autism: brief
report of a survey of national service centers. *Research in Developmental Disabilities* 2001; 22: 389–398.
Magiati, I., Moss, J., Yates, R., Charman, T. & Howlin, P. Is the Autism Treatment Evaluation Checklist a useful tool for monitoring progress in children with autism spectrum disorders? *J Intell Disab Res.,* 2011; 55: 302-312.
McConachie, H. Critique of current practices in assessment of children. In P. Zinkin & H. McConachie, *Disabled Children and Developing Countries,* Clinics in Developmental Medicine no. 136, 1995, MacKeith Press, Cambridge.
Mulloy, A., Lang, R., O'Reilly, M., Sigafoos, J., Lancioni, G., & Rispoli, M. Gluten-free and casein-free diets in the treatment of autism spectrum disorders: a systematic review. *Research in Autism Spectrum Disorders*, 2010; 4(3):328-339.
Nadji, F & Tymms, P. The P scales: how well are they working? *British Journal of Special Education* 2010; 37: 198-208.
Nye, C., & Brice, A. Combined vitamin B6-magnesium treatment in autism spectrum disorder. *Cochrane Database of Systematic Reviews* 2005, Issue 4. Art. No.: CD003497. DOI: 10.1002/14651858.CD003497.pub2.
Ospina, M.B., Seida, J.K., Clark, B., Karkhaneh, M., Hartling, L., Tjosvold, L., Vandermeer, B., & Smith, V. Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *PLoS ONE*. 2008; 3(11):e3755.
Ozonoff, S., Goodlin-Jones, B.L. & Solomon, M. Evidence-based assessment of Autism Spectrum Disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology,* 2005; 34: 523-540.
Petticrew, M. & Roberts, H. *Systematic Reviews in the Social Sciences.* 2006. Oxford: Blackwell Publishing.
Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. *Guidance on the conduct of narrative synthesis in systematic reviews. ESRC Methods Programme*, 2006
Reichow, B. & Wolery, M. Comprehensive synthesis of early intensive behavioral interventions for young children with autism based on the UCLA young autism project model. *J Autism Dev Disord*. 2009; 39: 23-41.
Rimland, B. & Edelson, M. *The Autism Treatment Evaluation Checklist (ATEC).* Autism Research Institute, San Diego, CA, 1999.
Siller, M. & Sigman, M. The behaviours of parents of children with autism predict the subsequent development of their children's communication. *Journal of Autism and Developmental Disorders*, 2002; 32: 77-89.
Siller, M. & Sigman, M. Modeling longitudinal change in the language abilities of children with autism: parent behaviours and child characteristics as predictors of change. *Developmental Psychology*, 2008; 44: 1691-1704.
Sinha, I. P., Williamson, P.R. & Smyth, R.L. Outcomes in clinical trials of inhaled corticosteroids for children with asthma are narrowly focussed on short term disease activity. *PLoS One*, 2008(a); 4 (7): e6276.

Sinha, I.P., Jones, L., Smyth, R.L. & Williamson, P.R. A systematic review of studies tha aim to determine which outcomes to measure in clinical trials in children. *PLoS Medicine*, 2008 (b); 5(4):e96.
Stone, W.L., Coonrod, E.E., Pozdol, S.L., & Turner, L.M. The Parent Interview for Autism–Clinical Version (PIA-CV): A measure of behavioral change for young children with autism. *Autism,* 2003; 7: 9–30.
Tager-Flusberg, H. & Joseph, R.M. Identifying neurocognitive phenotypes in autism. *Philos Trans R Soc Lond B Biol Sci*, 2003; 28, 358(1430): 303-314.
Terwee, C.B., Bot, S.D.M., de Boer, M.R., van der Windt, D.A.W.M., Knol, D.L., Dekker, J., Bouter, L.M. & de Vet, H.C.W. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 2007; 60: 34-42.
Terwee, C. B., Jansma, E. P., Riphagen, I. I., & de Vet, H.C.W. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of Life Research*, 2009; 18: 1115-1123.
Thiebaut, E., Adrien, J-L., Blanc, R. & Barthelemy, C. The Social Cognitive Evaluation Battery for children with autism: a new tool for the assessment of cognitive and social development in children with autism spectrum disorders. *Autism Research and Treatment*, 2010; doi:10.1155/2010/875037.
Wheeler, D., Williams, K., Seida, J. & Ospina, M. The Cochrane Library and Autism Spectrum Disorder: An Overview of Reviews. *Evidence-Based Child Health 2008;* 3: 3–15.
Williams, K,. Wheeler, D.M., Silove, N., & Hazell, P. Selective serotonin reuptake inhibitors (SSRIs) for autism spectrum disorders (ASD). *Cochrane Database of Systematic Reviews* 2010, Issue 8. Art. No.: CD004677. DOI: 10.1002/14651858.CD004677.pub2.
Williams, K.J., Wray, J.J., & Wheeler, D.M. Intravenous secretin for autism spectrum disorder. *Cochrane Database of Systematic Reviews* 2005, Issue 3. Art. No.: CD003495. DOI: 10.1002/14651858.CD003495.pub2.
Wolery, M. & Garfinkle, A.N. Measures in intervention research with young children who have autism. *J Autism Dev Disord*. 2002; 32(5):463-78.
World Health Organisation. *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. 1992. Geneva, World Health Organisation.
World Health Organisation. *International Classification of Functioning, Disability and Health for Children and Youth.* 2007. Geneva, World Health Organisation.
Zwaigenbaum, L., Bryson, S., Rogers ,T., Roberts, W., Brian, J., & Szatmari, P. Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosc* 2005; 23 (2-3): 143-152
Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A.S., Carver, L., Chawarska, K., et al. Clinical assessment and management of toddlers with suspected autism spectrum disorder: Insights from studies of high-risk infants. *Pediatrics,* 2009; 123: 1383–1391.