

# Autism diagnosis in children and young people

Evidence Update April 2013

A summary of selected new evidence relevant to NICE clinical guideline 128 'Autism – recognition, referral and diagnosis of children and young people on the autism spectrum' (2011)

**Evidence Update 40**



Evidence Updates provide a summary of selected new evidence published since the literature search was last conducted for the accredited guidance they relate to. They reduce the need for individuals, managers and commissioners to search for new evidence. Evidence Updates highlight key points from the new evidence and provide a commentary describing its strengths and weaknesses. They also indicate whether the new evidence may have a potential impact on current guidance. For contextual information, this Evidence Update should be read in conjunction with the relevant clinical guideline, available from the NICE Evidence Services topic page for [autism](#).

**Evidence Updates do not replace current accredited guidance and do not provide formal practice recommendations.**

NICE Evidence Services are a suite of services that provide online access to high quality, authoritative evidence and best practice.

### **National Institute for Health and Care Excellence**

Level 1A  
City Tower  
Piccadilly Plaza  
Manchester M1 4BT  
[www.nice.org.uk](http://www.nice.org.uk)


© National Institute for Health and Care Excellence, 2013. All rights reserved. This material may be freely reproduced for educational and not-for-profit purposes. No reproduction by or for commercial organisations, or for commercial purposes, is allowed without the express written permission of NICE.

# Contents

Introduction .....	4
Key points .....	5
1      Commentary on new evidence .....	8
1.1   Local pathway for recognition, referral and diagnostic assessment of possible autism .....	8
1.2   Recognising children and young people with possible autism .....	8
1.3   Referring children and young people to the autism team .....	14
1.4   After referral to the autism team .....	19
1.5   Autism diagnostic assessment for children and young people .....	19
1.6   After the autism diagnostic assessment .....	29
1.7   Medical investigations .....	29
1.8   Communicating the results from the autism diagnostic assessment .....	30
1.9   Information and support for families and carers .....	31
Areas not currently covered by NICE guidance .....	31
2      New evidence uncertainties .....	32
Appendix A: Methodology .....	33
Appendix B: The Evidence Update Advisory Group and Evidence Update project team .....	35

## Introduction

This Evidence Update identifies new evidence that is relevant to, and may have a potential impact on, the following reference guidance:

<sup>1</sup>  [Autism– recognition, referral and diagnosis of children and young people on the autism spectrum](#). NICE clinical guideline 128 (2011).

A search was conducted for new evidence from 17 August 2009 to 29 October 2012. A total of 3626 pieces of evidence were initially identified. Following removal of duplicates and a series of automated and manual sifts, 25 items were selected for the Evidence Update (see Appendix A for details of the evidence search and selection process). An [Evidence Update Advisory Group](#), comprising topic experts, reviewed the prioritised evidence and provided a commentary.

Although the process of updating NICE guidance is distinct from the process of an Evidence Update, the relevant NICE guidance development centres have been made aware of the new evidence, which will be considered when guidance is reviewed.

NICE is developing a clinical guideline on the management of autism in children and young people. See [the NICE website](#) for further information on the development of this guidance.

NICE also has guidance on [the recognition, referral, diagnosis and management of adults on the autism spectrum](#).

## Feedback

If you have any comments you would like to make on this Evidence Update, please email [contactus@evidence.nhs.uk](mailto:contactus@evidence.nhs.uk)

---

<sup>1</sup> NICE-accredited guidance is denoted by the Accreditation Mark 

## Key points

The following table summarises what the Evidence Update Advisory Group (EUAG) decided were the key points for this Evidence Update. It also indicates the EUAG's opinion on whether the new evidence may have a potential impact on the current guidance listed in the introduction. For further details of the evidence behind these key points, please see the full commentaries.

The section headings used in the table below are taken from the guidance.

**Evidence Updates do not replace current accredited guidance and do not provide formal practice recommendations.**

Key point	Potential impact on guidance	
	Yes	No
<p><b>Recognising children and young people with autism</b></p> <ul style="list-style-type: none"> <li>Retrospective analysis suggests that differences in development between typically developing children and those with autism may be apparent by age 6 months, and the types of differences may change over the child's first 2–3 years of life.</li> <li>In California, USA, the median age of diagnosis of autism seems to have fallen, particularly since 1996. Age of diagnosis may have correlations with socioeconomic and demographic factors.</li> <li>Evidence suggests that ethnic origin, maternal social class, and mother's marital status may not be associated with a diagnosis of autism or having severe autistic traits. However, first-born children may be more likely to have autism or autistic traits than subsequent children.</li> <li>In a study from the Netherlands, young people who have been arrested seemed to have higher levels of symptoms of autism than children in the general population, but lower levels of symptoms than children with diagnosed autism.</li> <li>Girls may have different symptoms of autism from boys, including less repetitive stereotyped behaviour and fine motor impairment, and higher levels of emotional problems. Boys may be more likely than girls to show aggressive behaviour and hyperactivity.</li> <li>Symptoms of autism may change over time in young children; this may be improvement, worsening, or persistence of symptoms.</li> </ul>		<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> </ul>
<p><b>Referring children and young people to the autism team</b></p> <ul style="list-style-type: none"> <li>The sensitivity and specificity of autism diagnostic tools may be improved by alternative scoring algorithms using the items most associated with a diagnosis of autism and selecting optimum cut-off scores. However, no single tool alone seems to have adequate sensitivity and specificity for diagnosis of autism.</li> <li>Children with autism may have differing developmental trajectories: regression, plateau, and no regression or plateau. Those with regression may have worse symptoms.</li> </ul>		<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> </ul>

Key point	Potential impact on guidance	
	Yes	No
<ul style="list-style-type: none"> <li>Children with autism and regression may have a period of rapid early language development before regression, however all children with autism seem to have lower social-communication behaviour than typically developing children by age 24 months.</li> </ul>		✓
<p><b>Autism diagnostic assessment for children and young people</b></p> <ul style="list-style-type: none"> <li>Criteria for diagnosing autism according to the Diagnostic and Statistical Manual (DSM)-IV-TR (text revision) may not have adequate sensitivity and specificity for diagnosing autism in children with intellectual disabilities.</li> <li>Proposed criteria for diagnosing autism in DSM-5 may have lower sensitivity but better specificity for diagnosing autism than DSM-IV-TR criteria.</li> <li>Diagnosis of subgroups of autism spectrum disorder may vary in different centres. Children with social communication problems and low intelligence quotient (IQ) may be more likely to be diagnosed with autism than with other autism spectrum disorders.</li> <li>Mothers with depression may report higher levels of symptoms of autism in their children on the Social Responsiveness Scale; however, smaller effects of maternal depression may be seen with the Autism Diagnostic Interview – Revised (ADI-R). Autism Diagnostic Observational Schedule (ADOS) results may not be affected.</li> <li>The sensitivity and specificity of ADOS in detecting autism may differ according to the child’s age and verbal development.</li> <li>New algorithms for scoring ADI-R results in diagnosing autism in children younger than 4 years may improve specificity without large reductions in sensitivity.</li> <li>Using results of both ADOS and ADI-R to diagnose autism may result in optimum sensitivity and specificity compared with use of either tool alone.</li> <li>Results of a preliminary study of the Child Symptom Inventory suggest that it may not have adequate sensitivity or specificity for discriminating between children with autism and those with attention deficit and hyperactivity disorder (ADHD).</li> <li>Parents may report language milestones on the ADI-R as happening later than they actually occurred because of ‘telescoping’ (that is, perceiving distant events as more recent than they are). The further after the child begins to talk that the ADI-R is administered, the greater the effects of telescoping.</li> <li>The Manchester Inventory for Playground Observation may not have adequate sensitivity and specificity for discriminating between children with autism, externalising disorders, or internalising disorders.</li> </ul>		<ul style="list-style-type: none"> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> <li>✓</li> </ul>

Key point	Potential impact on guidance	
	Yes	No
<p><b>Medical investigations</b></p> <ul style="list-style-type: none"> <li>Genetic testing of all children with autism seems to be associated with a small yield of new diagnoses of genetic disorders, and children with dysmorphic features may be more likely to have a genetic disorder than those without dysmorphic features.</li> <li>Children with autism may have differences in electroencephalogram (EEG) coherence (a measure of the connectivity between different parts of the brain) compared with typically developing children; however, further studies of EEG coherence in disorders associated with or related to autism are needed to confirm these findings.</li> </ul> <p><b>Areas not currently covered by NICE guidance</b></p> <ul style="list-style-type: none"> <li>Over time, diagnoses of autism may change including reclassification of children as having a different autism spectrum disorder than the original diagnosis, and a small proportion may no longer meet the diagnostic criteria for autism.</li> </ul>		<p>✓</p> <p>✓</p> <p>✓</p>

# 1 Commentary on new evidence

These commentaries analyse the key references identified specifically for the Evidence Update. The commentaries focus on the 'key references' (those identified through the search process and prioritised by the EUAG for inclusion in the Evidence Update), which are identified in bold text. Supporting references provide context or additional information to the commentary. Section headings are taken from the guidance.

## Glossary of abbreviations

This glossary lists selected abbreviations that are repeated throughout this Evidence Update.

<b>ADI-R</b>	Autism Diagnostic Interview – Revised
<b>ADOS</b>	Autism Diagnostic Observational Schedule
<b>DSM</b>	Diagnostic and Statistical Manual of Mental Disorders
<b>ICD</b>	International Classification of Diseases
<b>IQ</b>	Intelligence quotient
<b>PDD-NOS</b>	Pervasive developmental disorder-not otherwise specified
<b>SCQ</b>	Social Communication Questionnaire
<b>SRS</b>	Social Responsiveness Scale

### 1.1 Local pathway for recognition, referral and diagnostic assessment of possible autism

No new key evidence was found for this section.

### 1.2 Recognising children and young people with possible autism

#### **Possible early predictors of autism**

[NICE CG128](#) includes tables of possible signs and symptoms of autism for case identification. It recommends that autism should not be ruled out if the exact features described in the tables are not evident; they should be used for guidance, but do not include all possible manifestations of autism.

[Bolton et al. \(2012\)](#) reported on a retrospective analysis of data from the Avon Longitudinal Study of Parents and Children (ALSPAC), which was a prospective cohort study of 13,971 surviving children of 14,541 pregnant women in south-west England with an expected delivery date between April 1991 and December 1992. The present analysis focused on early predictors (up to age 30 months) of later autism (up to age 11 years).

Children with autism (n=86) were identified from community records in the special educational needs database. All cases were diagnosed after multidisciplinary assessment, and clinical records were reviewed by a consultant paediatrician to confirm that diagnoses met ICD-10 criteria. Identification and review of cases was blind to the ALSPAC data. The ALSPAC ethical approval did not allow researchers to recall participants with high scores on questionnaires for in-depth assessment; therefore the identification of cases was dependent on clinical diagnosis.



Seven principal factors relating to verbal ability, language acquisition, social understanding, semantic-pragmatic skills, repetitive stereotyped behaviour, articulation and social inhibition were scored using results of parents' questionnaires and standardised observational measures collected at various time points. A common autistic trait score was calculated as the mean of the scores of the principal factors.

The prevalence of autism was 62 per 10,000 children by age 11 years. At age 6 months, differences in fine motor skills and social skills and communication, and concerns about vision were associated with subsequent diagnosis of autism. Differences in hearing, vocabulary and understanding words, and in feeding difficulties and fads were apparent by age 15 months. At age 18 months, more widespread differences were associated with a subsequent diagnosis of autism: listening and responding to sounds, play and imitation, health concerns and repetitive and unusual behaviours. Temperamental traits and differences in bowel habit and stool characteristics were noticed by age 24 months, and by 30 months differences in crying and tempers were associated with autism. The authors stated that these associations remained significant after correction for multiple comparisons, although individual p values were not reported.

Further analysis to account for children's IQ suggested that some associations may be affected by general cognitive abilities. Associations that were consistently associated with the child's IQ were: play and imitation, fine motor skills, vocabulary, communication, temperamental traits, and bowel habit and stool characteristics. The authors discussed the possibility of reporting bias, in that parents of children with autism may generally view their children as having more problems, but concluded that reporting bias was not the only explanation for the results. Additionally, this study was retrospective and prospective identification of characteristics associated with autism may not be possible in very young children (for example, at 6 months old).

The characteristics associated with subsequent diagnosis of autism identified in this study are broadly consistent with the signs and symptoms of possible autism noted in current guidance; therefore no impact on [NICE CG128](#) is expected.

#### **Key reference**

Bolton PF, Golding J, Emond A et al. (2012) [Autism spectrum disorder and autistic traits in the Avon Longitudinal Study of Parents and Children: precursors and early signs](#). *Journal of the American Academy of Child & Adolescent Psychiatry* 51: 249–60

#### **Role of social and demographic factors**

[NICE CG128](#) recommends that when considering the possibility of autism, clinicians should be aware that signs and symptoms will not always have been recognised by parents, carers, children or young people themselves or by other professionals.

[Fountain et al. \(2011\)](#) reported on a retrospective cohort study investigating the individual and community-level factors that may affect the age of diagnosis of autism. Included children (n=17,185) were born in California, USA, between 1992 and 2001, had a DSM-IV diagnosis of autism between the ages of 2 and 8 years, and were enrolled with the Californian Department of Developmental Services (DDS). DDS records were matched with Californian birth records using algorithms based on name, birth date, ethnicity, area of birth (zip code) and sex. The age of diagnosis was calculated from client development evaluation reports. Individual-level variables studied were maternal ethnicity, sex, coexisting intellectual disability, maternal age, parental education level, economic status, and birth order. Community autism prevalence was calculated by zip code, and data for community educational status, median property value, and proportion of residents living below the poverty line were obtained from US census records.

The mean age of diagnosis fell over time, particularly after 1996 (from 4.9 years in 1992 to 3.6 years in 2001); the authors noted that this was mainly due to an increase in diagnosis at age 3 years. Non-white ethnicity and poverty (as indicated by births paid for by Medi-Cal) seemed to be associated with older age at diagnosis. Higher parental educational status, and higher local autism rates seemed to be associated with lower age of diagnosis. Higher local property values, and conversely, higher poverty rates also seemed to be associated with younger age of diagnosis. Children with higher communication capabilities seemed to be diagnosed at an older age.

Limitations discussed by the authors were that the results may not be generalisable outside of California, that enrolment in the DDS is voluntary so it may not include all cases, and the available data for variables was from databases that collect information for different purposes than this study. The authors noted a need for research to show whether their results could be replicated in different regions, and to determine how later age of diagnosis affects children's outcomes.

[Russell et al. \(2011\)](#) reported a study investigating whether social and demographic factors are related to diagnosis or non-diagnosis of autism. Data from ALSPAC were used for children surviving at least 1 year (n=13,981). Children with a diagnosis of autism (n=71) were identified via their medical records. Diagnosis was made by the children's clinicians who were blind to data obtained for the ALSPAC study. A control group consisted of 142 children who scored in the top 2 percentiles of the composite autism score (described below). Demographic and socioeconomic data derived from questionnaires sent to mothers when their children were young were assessed for factors that could influence diagnosis.

The composite autism score was calculated by measuring 27 autistic traits at age 2.5–4.0 years, encompassing impairment in social interaction, social communication, and repetitive behaviours or restricted interests that had been independently assessed in the ALSPAC study were identified. These traits were then used to identify children who showed them early in life but had not subsequently been diagnosed with autism. Logistic regression analysis was then used to determine which measures of autistic behaviour were most strongly associated with a diagnosis of autism for each domain. These traits were combined to give the overall composite autism score.

For impairments in social interaction, the prosocial score on the Strengths and Difficulties Questionnaire (SDQ) was most strongly associated with a clinical diagnosis of autism (odds ratio [OR]=3.77, 95% confidence interval [CI] 2.67 to 5.30). For impairments in communication, the 'enjoys pretend games' score was most associated with a diagnosis of autism (OR=2.13, 95% CI 1.72 to 2.64); and for restricted repetitive and stereotyped patterns of behaviour the 'child is afraid of new things or new situations' score was most strongly associated with a diagnosis of autism (OR=1.64, 95% CI 1.30 to 2.06).

Ethnic origin, maternal social class, and mother's marital status did not significantly predict a diagnosis of autism or having severe autistic traits. However, about 9 times more boys than girls were diagnosed with autism ( $p<0.001$ ) and significantly more boys than girls had autistic traits ( $p<0.001$ ). The mean age that mothers gave birth (30 years) to children who were subsequently diagnosed with autism was higher than the age of the overall population (27 years,  $p=0.002$ ). A first-born child was more likely to have autism ( $p=0.015$ ) or autistic traits ( $p=0.007$ ) than subsequent children.

The authors noted that the groups with diagnosed autism and autistic traits were not an exact match, and that diagnosis of autism was not standardised. There was no knowledge of alternative diagnoses made in the control group, and possible selection bias in people who participated in ALSPAC and those who did not.

The results of the US-based study by Fountain et al. (2011) may not be directly relevant to the UK because of differences in healthcare systems and policies for identification and referral of children with possible autism. However, the results from Fountain et al. (2011) and Russell et al. (2011) provide some support for recommendations in [NICE CG128](#) that signs and symptoms will not always have been recognised by parents, carers, children or young people themselves or by other professionals, in that there may be factors affecting parental recognition of symptoms.

#### **Key references**

Fountain C, King MD, Bearman PS (2011) [Age of diagnosis for autism: individual and community factors across 10 birth cohorts](#). *Journal of Epidemiology and Community Health* 65: 503–10

Russell G, Steer C, Golding J (2011) [Social and demographic factors that influence the diagnosis of autism spectrum disorders](#). *Social Psychiatry and Psychiatric Epidemiology* 46: 1283–93

#### **Children in the criminal justice system**

[NICE CG128](#) recommends that when considering the possibility of autism, clinicians should be aware that: important information about early development may not be readily available for some children and young people, for example looked-after children and those in the criminal justice system.

[Geluk et al. \(2012\)](#) reported a prospective cohort study from the Netherlands assessing the level of symptoms of autism in children aged under 12 years who had been arrested for a first offence (n=308, 86% male), compared with children from the general population. The study additionally investigated the predictive value of symptoms of autism for behaviour that may lead to arrest and possible mediating effects of coexisting externalising disorders. Children were included if they were taken to the police station or reprimanded on the spot for behaviour that could have been prosecuted if the person was older than 12 years, excluding status offences. Status offences are actions that are prohibited only to a specific group of people, for example drinking under the legal minimum age.

Participants were followed-up annually for 2 years after their first arrest; follow-up data were available for 235 children (76%). The mean age at baseline was 10.7 years (standard deviation [SD]=1.5 years). The Children's Social Behaviour Questionnaire (CSBQ) was administered at each assessment – complete data were available for: 273 children at baseline; 243 children at year 1, and 224 children at year 2. Behaviour that could lead to arrest was measured by the Observed Antisocial Behaviour Questionnaire (OABQ). Externalising disorders were assessed in arrestees with the US National Institute of Mental Health's Diagnostic Interview Schedule for Children (DISC) sections for attention deficit and hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), and conduct disorder.

At each assessment, the arrestee group was compared with a general population group matched for age and gender (n=840, n=1134, and n=1057 for baseline, year 1, and year 2 assessments respectively). The 'general population' sample was derived from participants of a study using the CSBQ in schools and from a sample of children who participated in a study about intellectual disability. A further comparison against a group of children with diagnosed autism matched for age and gender was done (n=209, n=213, and n=185 for baseline, year 1, and year 2 assessments respectively). This group consisted of children attending a specific clinic in the Netherlands and children who participated in the study of intellectual disability.

The arrestees had significantly higher CSBQ scores than the general population for most subscales: not tuned; social; orientation; understanding; and stereotyped (p<0.01 for all comparisons) with medium effect sizes at baseline and small effect sizes at follow-up. Arrestees had significantly lower CSBQ scores in all subscales including the change subscale, than children with autism (p<0.01 for all comparisons) with medium-to-large effects sizes at all assessments.

In regression analyses, all CSBQ subscales were significantly associated with future behaviour that could lead to arrest ( $p < 0.001$  for all subscales). However, after adjustment for coexisting externalising disorders (present in 33.2% of arrestees at baseline) the change subscale was no longer significantly associated with future behaviour that could lead to arrest, although a significant association remained for all other subscales.

The authors stated that the main limitation of this study was the lack of clinical diagnosis of autism in this sample, so no conclusion could be made about autism and behaviour that could lead to arrest. Additionally, they noted that the stressful experience of having a child arrested could have led to over-reporting of symptoms of autism by parents at baseline.

This study suggests that children who have been arrested may have higher levels of symptoms of autism than the general population, but lower levels of symptoms than those who have had a clinical diagnosis of autism. However, further research is needed, so no impact on [NICE CG128](#) is anticipated.

#### Key reference

Geluk CAML, Jansen LMC, Vermeiren R et al. (2012) [Autistic symptoms in childhood arrestees: longitudinal association with delinquent behaviour](#). *Journal of Child Psychology and Psychiatry* 53: 160–7

#### Differences in symptoms of autism between girls and boys

[NICE CG128](#) recommends that when considering the possibility of autism, clinicians should be aware that: autism may be under-diagnosed in girls.

[Mandy et al. \(2012\)](#) reported a cohort study of consecutive referrals of girls ( $n=52$ ) and boys ( $n=273$ ) at a specialist clinic in the UK for the assessment of high-functioning children with social communication difficulties to investigate the female phenotype of autism. Participants were seen between June 1999 and July 2009.

All children were assessed with the Developmental, Dimensional and Diagnostic Interview (3Di) administered by an experienced child psychiatrist or clinical psychologist. The 3Di uses a computerised algorithm to calculate ADI-R equivalent scores. Final clinical diagnosis was based on the 3Di plus ADOS, if available, and structured reports from the child's nursery or school. ADOS was administered in 35 girls and 154 boys, and parent and teacher reported SDQ results were available for 37 girls and 205 boys.

A diagnosis of autism ( $n=113$ ) was made if the child had 3Di scores above the standard cut-off and delay in single word or phrase speech. Asperger's disorder ( $n=94$ ) was diagnosed if the child had 3Di scores above the cut-off and no delay in speech. PDD-NOS ( $n=118$ ) was diagnosed according to DSM-IV-TR (text revision) criteria if the child had significant qualitative impairments in reciprocal social interaction and either communication or repetitive stereotyped behaviour.

Girls had significantly lower 3Di scores for repetitive stereotyped behaviour (mean=3.6, SD=2.7) than boys (mean=4.4, SD=2.6,  $p=0.03$ ). A similar result was seen for ADOS scores for repetitive stereotyped behaviour (mean for girls=14.4, SD=17.3, mean for boys=22.7, SD=22.1,  $p=0.04$ ). Fine-motor impairment was lower in girls (mean=6.4, SD=4.8) than in boys (mean for boys=8.3, SD= 4.2,  $p=0.003$ ). When repetitive and stereotyped behaviours were analysed individually, boys were more likely to meet the criterion 'large store of factual information' than girls (OR=1.44, 95% CI 1.06 to 1.95,  $p=0.006$ ). Other repetitive and stereotyped behaviours did not seem to differ between girls and boys.

Differences were also seen between SDQ scores reported by parents and teachers. Parent-reported SDQ scores showed a significantly higher level of emotional problems in girls (mean=6.0, SD=3.1) than in boys (mean=4.75, SD=2.8,  $p=0.02$ ). Teacher-reported SDQ scores showed: significantly lower overall score in girls (mean=14.0, SD=8.2) than in boys

(mean=17.2, SD=6.9,  $p=0.01$ ); significantly lower hyperactivity and inattention in girls (mean=4.3, SD=3.2) than in boys (mean= 6.5, SD=2.7,  $p<0.001$ ); and significantly higher prosocial behaviour in girls (mean=5.2, SD=3.1) than in boys (mean=3.9, SD=2.7,  $p=0.009$ ).

Author-reported limitations were: the nature of the sample (referrals to a specialist clinic), which may mean that the results are not generalisable to all people with high-functioning autism, or to people with lower-functioning autism. Additionally, ADOS and SDQ data were not available for all participants, which the authors explained was due to changes in the clinic's practice over time, rather than systematic bias in which selected participants received full assessment.

[Giarelli et al. \(2012\)](#) investigated the diagnosis of autism and behavioural features in boys compared with girls using data from the US Autism and Developmental Disabilities Monitoring (ADDM) network collected between 2001 and 2006. This analysis used data from 2002 for children who were aged 8 years ( $n=2568$ ). Participants met the ADDM definition of a surveillance case of autism which was either a documented diagnosis of autism by a qualified professional, with no conflicting diagnostic information ( $n=1497$ ), or no documented diagnosis of autism but at least 1 evaluation record from a school or clinic indicating behaviour consistent with DSM-IV-TR criteria for autism or pervasive developmental disorder ( $n=1071$ ).

Boys accounted for 81% of the sample. Girls were more likely than boys to have cognitive impairment (OR=0.76, 95% CI 0.65 to 0.97,  $p=0.024$ ). When the group that had no documented diagnosis of autism was analysed by likely diagnosis as determined by the researchers, girls were significantly more likely than boys to have a diagnosis of general developmental delay ( $p=0.02$ ), epilepsy ( $p<0.001$ ), intellectual disability ( $p=0.04$ ) or no diagnosis ( $p=0.038$ ).

For children with an autism spectrum disorder, girls were less likely than boys to show aggressive behaviour (50% of boys vs 41% of girls,  $p=0.008$ ) or hyperactivity or short attention span (85% of boys vs 78% of girls,  $p<0.001$ ), but were more likely to have seizures or seizure-like behaviour (24% of boys vs 34% of girls,  $p<0.001$ ).

Limitations reported by the authors were that this sample may not have included all children with autism because some children may not have had any assessment at school or in a clinic. Additionally, data for cognitive function were not available for all children and the age of testing varied from 4 to 8 years.

The evidence from both Mandy et al. (2012) and Giarelli et al. (2012) suggests that differences in symptoms of autism may exist between girls and boys with autism that could contribute to under-recognition of autism in girls, which is consistent with [NICE CG128](#). Girls may have less repetitive stereotyped behaviour and fine motor impairment, and higher levels of emotional problems. Boys may be more likely than girls to show aggressive behaviour, and hyperactivity.

#### **Key references**

Giarelli E, Wiggins LD, Rice CE et al. (2010) [Sex differences in the evaluation and diagnosis of autism spectrum disorders among children](#). *Disability and Health Journal* 3: 107–16

Mandy W, Chilvers R, Chowdhury U, et al. (2012) [Sex differences in autism spectrum disorder: evidence from a large sample of children and adolescents](#). *Journal of Autism and Developmental Disorders* 42: 1304–13

#### **Stability of symptoms of autism over time in young children**

[NICE CG128](#) states – do not rule out autism because of: good eye contact, smiling and showing affection to family members; reported pretend play or normal language milestones; difficulties appearing to resolve after a needs-based intervention (such as a supportive



structured learning environment); or a previous assessment that concluded that there was no autism, if new information becomes available.

A cohort study reported by [Lord et al. \(2012a\)](#) assessed symptoms of autism over time in children with possible autism compared with a control group. Children with possible autism (n=65) were recruited from consecutive referrals to a specialist clinic in the USA. A control group (n=13) consisted of children who were either typically developing or had language delay who were part of comparison groups in projects studying communication delays. Of the total sample, 77% were male.

Children were seen an average of 6 times over an average of 20 months. The first appointment was at age 12–19 months. Children were seen by the same clinician at most visits, and every 6 months each child was assessed by an independent clinical psychologist or advanced graduate student in psychology who had not met the child before and was blind to the child's previous assessments or diagnosis. The independent clinician gave a best estimate diagnosis and a probability rating of the degree of certainty in the diagnosis. After every assessment, families were given oral feedback and a brief report of the results.

At the final visit (age 34 months, SD=7 months), each child was given a final diagnosis by 2 clinicians, 1 of whom was unfamiliar with the child, but at this point clinicians had access to all information about previous assessments. These final diagnoses resulted in a population consisting of 39 children with autism, 20 children with typical development, and 19 children with other diagnoses. An analysis of ADOS scores using Bayesian information criterion scoring resulted in 4 classes of autism: severe persistent (n=16, 21%), worsening (n=16, 21%), improving (n=15, 19%), and non-spectrum (n=31, 40%).

The authors did not discuss possible limitations of their study. However, this evidence suggests that trajectories of autism differ, and that young children may have symptoms that worsen, improve or stay the same over time. This conclusion is generally consistent with the recommendation in [NICE CG128](#) not to rule out autism because a previous assessment concluded that there was no autism, if new information becomes available.

#### **Key reference**

Lord C, Luyster R, Guthrie W et al. (2012a) [Patterns of developmental trajectories in toddlers with autism spectrum disorder](#). *Journal of Consulting and Clinical Psychology* 80: 477–89

### **1.3 Referring children and young people to the autism team**

#### **Diagnostic criteria and tools**

[NICE CG128](#) recommends that clinicians should be aware that tools to identify children and young people with an increased likelihood of autism may be useful in gathering information about signs and symptoms of autism in a structured way but are not essential and should not be used to make or rule out a diagnosis of autism. Also they should be aware that: a positive score on tools to identify an increased likelihood of autism may support a decision to refer but can also be for reasons other than autism; and a negative score does not rule out autism.

#### ***The Quantitative Checklist for Autism in Toddlers (Q-CHAT) and Autism Spectrum Quotient (AQ)***

[Allison et al. \(2012\)](#) assessed the Q-CHAT and 3 versions of the AQ (child, adolescent and adult) to determine whether 10 items from each diagnostic tool showed equivalent sensitivity and specificity to the full versions. Each tool was assessed twice in participants with autism and a control group, once as a 'derivation' sample, and once as a 'validation' sample (16 groups altogether).

Parents of adolescents aged 12–15 years (n=162), children aged 4–11 years (n=432) and toddlers aged 15–47 months (n=126) with autism voluntarily registered via the study's website and completed the appropriate questionnaire (AQ-adolescent for adolescents, AQ-child for children, and Q-CHAT for toddlers). Only children and young people diagnosed by a recognised doctor or clinical psychologist at a recognised clinic using DSM-IV criteria were included. The control groups for the AQ-adolescent and AQ-child were recruited by sending a copy of the relevant questionnaire to parents of adolescents (n=475), children (n=940) who were participating in an epidemiological study of social communication skills, and only those without neurodevelopmental disorders were included. The toddler control group (n=754) was from a previous study reported by [Allison et al. \(2008\)](#).

Adults with autism (n=449) voluntarily registered via the study's website, provided details about their diagnosis and completed an online version of the AQ. Only people diagnosed by a recognised doctor or clinical psychologist at a recognised clinic using DSM-IV criteria were included. Adult control participants (n=838) were recruited via a website for people from the general population who are interested in taking part in research. Only people older than 16 years who had no neurodevelopmental diagnosis were included.

For each tool, participants in the autism and control groups were randomly assigned to 'derivation' or 'validation' groups. The scores for the AQ and the Q-CHAT were converted to a binary format for consistency of analysis across tools. A discrimination index was calculated for each item of every tool, which was the proportion of positive scores in the autism group minus the proportion of positive scores in the control group. For all versions of the AQ, the 2 items with the highest discrimination index in each of the 5 subscales were chosen and for Q-CHAT, the 10 items with the highest score were chosen. Receiver operating characteristic (ROC) curves were produced for the items with high discrimination index scores and the area under the curve (AUC) was calculated for each. The AUC for each 10-item measure was then compared with the AUC of the full version.

All short versions had an AUC of more than 0.90, which the authors described as 'excellent' validity. All scores for 10-item tools showed significant differences between the autism group and the control group (all  $p < 0.0001$ ). The magnitude of the difference was described as large for all tools. The 10-item version of each tool correlated significantly with its respective full version (all  $p < 0.0001$ ). At a cut-off score of 6: the 10-item AQ-adult had sensitivity of 0.88, specificity of 0.91, and positive predictive value of 0.85; the 10-item AQ-adolescent had sensitivity of 0.93, specificity of 0.95 and positive predictive value of 0.86; and the 10-item AQ-child had sensitivity of 0.95, specificity of 0.97 and positive predictive value of 0.94. At a cut-off score of 3, the 10-item Q-CHAT had sensitivity of 0.91, specificity of 0.89, and positive predictive value of 0.58.

The authors noted several limitations, including the retrospective nature of the analyses, the differing methods of administering the questionnaires, the lack of independent validation of diagnoses of autism, and that the subtypes of autism differed between groups. Furthermore, the self-selection used to recruit many groups may bias the results. The authors concluded that their results are the first step in developing short instruments to help health and social care professionals in the referral pathway for autism, and that further work to validate the short tools in primary care and social care was necessary. Therefore, this evidence is not likely to affect recommendations in [NICE CG128](#).

#### **Key reference**

[Allison C, Auyeung B, Baron-Cohen S \(2012\) Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. \*Journal of the American Academy of Child & Adolescent Psychiatry\* 51: 202–12](#)

### Supporting reference

Allison C, Baron-Cohen S, Wheelwright S, et al. (2008) [The Q-CHAT \(Quantitative Checklist for Autism in Toddlers\): a normally distributed quantitative measure of autistic traits at 18-24 months of age: preliminary report](#). *Journal of Autism and Developmental Disorders*

### ***The Ghuman-Folstein Screen for Social Interaction (SSI)***

[Ghuman et al. \(2011\)](#) reported a preliminary study of the validity of the SSI in children (n=350) aged 24–61 months. Children were recruited during appointments in psychiatry, autism and developmental clinics, or specialty research clinics for genetic and developmental disorders. This 'high-risk' group (n=182) included children with autism (n=66), PDD-NOS, (n=48) and those determined as not having autism spectrum disorder (n=68). A control group, of children (n=168) who were developing typically, was recruited during routine visits to primary care. Analyses were split by age (24–42 months and 43–61 months) and ROC curves were produced.

For the 'high-risk' group, the SSI was completed before the child's diagnostic assessment. Autism was assessed via expert clinical diagnosis. Diagnoses for children who did not have autism spectrum disorder included communication disorders, intellectual impairment and other developmental disorders; 17% had psychiatric diagnoses only.

The SSI differentiated between each of the diagnostic groups for both the older and younger groups (both  $p < 0.001$ ). The autism groups had the lowest SSI scores, the PPD-NOS groups' scores were higher, the scores for the group with diagnoses other than autism were higher again, and the control group had the highest scores. No effect or interaction between verbal ability or age and test scores was seen (statistical data not reported).

Regression analysis was used to determine the critical items on the SSI. These results were used to refine the instrument in both younger and older age groups, resulting in 2 separate instruments. Items with lower discrimination between autism and the other disorders were removed. The optimum cut-off score was 45 for the SSI for younger children with a positive predictive value of 0.87 and negative predictive value of 0.70. The optimum cut-off score for the SSI for older children was 37 with a positive predictive value of 0.78 and negative predictive value of 0.76.

Limitations recognised by the authors were that standard assessments were not conducted across all children. A small number of participants had severe developmental delays so the results may not be generalisable to that population. The small number of children with autism or PDD-NOS (n=114) relative to the 54 items on the SSI, and lack of a matched control group may have affected the ability to determine critical items. The authors described the results as preliminary and, as such, no impact is expected on [NICE CG128](#).

### Key reference

Ghuman JK, Leone SL, Lecavalier L et al. (2011) [The screen for social interaction \(SSI\): a screening measure for autism spectrum disorders in preschoolers](#). *Research in Developmental Disabilities* 32: 2519–29

### ***The SCQ as second-level screening tool***

[Oosterling et al. \(2010\)](#) investigated the use of the SCQ as a second-level screening tool in children aged 20–40 months (n=208, 82% male) referred to a child and adolescent psychiatry clinic in the Netherlands for suspected autism. This study additionally aimed to assess the sensitivity and specificity of the SCQ and ADI-R against clinical diagnosis with or without ADOS results, and to determine whether cut-off scores need to vary for different age groups.

Most children (94%) referred to the clinic had a positive result on the Early Screening of Autistic Traits (ESAT) questionnaire, and the remainder had negative ESAT scores but parents or healthcare professionals were concerned about their social-communicative development.



Assessment consisted of: unstructured developmental interviewing, psychiatric evaluation and parent-child play observation; administration of SCQ, ADOS, ADI-R and the Communication and Symbolic Behavior Scales-Developmental Profile; and psychometric testing of cognition and language. A consensus-based diagnosis was made by at least 2 experienced professionals and was explained to parents in a feedback session.

An autism spectrum disorder was diagnosed in 143 of the children: 92 had autism, 49 had PDD-NOS, and 2 had Asperger's disorder. Of those who did not have an autism spectrum disorder, 21 had language disorders, 17 had externalising disorders, 13 had 'other' developmental disorders, 10 had intellectual disabilities without autism, 3 had internalising disorders, and 1 had normal development.

SCQ scores were significantly higher in the children with autism ( $p < 0.001$ ) or autism spectrum disorder ( $p < 0.05$ ), than in children who did not have an autism spectrum disorder. For detecting autism, SCQ had sensitivity of 0.76 and specificity of 0.62, and for detecting autism spectrum disorders it had a sensitivity of 0.66 and specificity of 0.62. No cut-off score with acceptable sensitivity and specificity could be identified.

The authors concluded that the SCQ would be likely to result in a high number of false positives as a second-level screening instrument. The correlation with ADI-R scores was noted to be strong ( $r = 0.70$ ,  $p < 0.01$ ), however the authors noted that correlation was not useful if a child is assessed as having autism on 1 but not both tools.

Limitations identified by the authors were: that ADOS and ADI-R scores were administered by the same clinician; the use of ADOS and ADI-R scores may have influenced the best clinical estimate diagnosis; and that the sample was defined as at high risk of autism (most scored positively on ESAT) and would be expected to score highly on SCQ, resulting in possible false-positives.

The SCQ is not mentioned specifically in [NICE CG128](#); however, this evidence is consistent with the recommendation that tools may be useful for information gathering, but that they should not be used to make or rule out a diagnosis of autism.

#### **Key reference**

Oosterling I, Rommelse N, de Jonge M et al. (2010) [How useful is the Social Communication Questionnaire in toddlers at risk of autism spectrum disorder?](#) *Journal of Child Psychology and Psychiatry* 51: 1260–8

#### **Developmental trajectories in autism**

[NICE CG128](#) recommends referring children younger than 3 years to the autism team if there is regression in language or social skills.

[Kalb et al. \(2010\)](#) reported a study investigating development characteristics and outcomes in children with autism, grouped by 3 patterns of symptom onset: regression, plateau, and no regression or plateau. Data were obtained from the US Interactive Autism Network, an internet-based registry for families with at least 1 child with autism. Parents of children with autism can register with the network and complete questionnaires about their child. Data were extracted for participants aged 3–17 years ( $n = 2720$ , 83% male) who had completed primary history questionnaires and autism symptom measures (SRS and SCQ) with a score of 15 or more on the SCQ. Children were excluded if they had: fragile X syndrome or tuberous sclerosis; developmental concerns; or loss of skill reported after the age of 3 years. Differences in age and ethnicity were accounted for as covariates in analyses.

The mean age of onset of plateau was 24.1 months and of regression was 19.5 months. First concerns about autism occurred more than 2 months later for children who had regression or plateau than for children with no regression or plateau ( $p < 0.001$ ). Children with regression had higher SRS scores (mean=113, SD=28) than those with plateau (mean=109, SD=29,

$p < 0.001$ ) or those with no regression or plateau (mean=108, SD=27,  $p < 0.001$ ). SCQ scores were higher in children with regression (mean=26, SD=6) than those with plateau (mean=25, SD=6,  $p < 0.001$ ) and those with no regression or plateau (mean=24, SD=6,  $p < 0.001$ ).

The authors noted that although no differences in cognitive ability were found, this outcome may be particularly affected by reporting bias. Limitations identified by the authors were: possible recall bias and 'forward telescoping' by parents (that is, perceiving distant events as more recent than they are) due to the cross-sectional retrospective study design; possible bias from data based solely on parental report including lack of validation of diagnosis; and concerns about quality of online data.

This evidence is not likely to affect [NICE CG128](#) because the finding that regression is associated with more severe symptoms is broadly consistent with the recommendation to refer to the autism team if regression occurs in children younger than 3 years.

#### Key reference

Kalb LG, Law JK, Landa R et al. (2010) [Onset patterns prior to 36 months in autism spectrum disorders](#). *Journal of Autism and Developmental Disorders* 40: 1389–402

#### Social-communication behaviour in different trajectories of autism

[NICE CG128](#) recommends referring children younger than 3 years to the autism team if there is regression in language or social skills.

[Ozonoff et al. \(2011\)](#) reported a study examining home videos of children with autism ( $n=52$ ) or typical development ( $n=23$ ). The group of children with autism had been diagnosed with autism or PDD-NOS in the community before joining the study (at age 23–59 months), and were subsequently assessed as having autism according to DSM-IV criteria by study staff. They did not have other medical conditions such as seizures or fragile X syndrome, and had no hearing or vision impairments. Typically developing children were aged 12–42 months at the time of participating in the study and had no medical or developmental concerns, or hearing or vision impairments.

All available home videos of participants recorded at ages 6–24 months were requested from families. Only segments of footage that included that child and at least 1 other person were scored, and scoring was done by a researcher blind to whether the child was in the autism or typically developing groups. Four social communication behaviours lasting 0.5 seconds or longer were counted: looks at people, smiles at people, language and joint attention. Any findings inconsistent with typical development (such as phrase speech at 6 months) or were extreme outliers (more than 3 SD of the mean), were checked by inspecting the video.

Latent class analysis was used to characterise social-communication behaviour. Bayesian information criterion was used to compare competing models to select the best fitting model. The final model had 3 trajectories: early onset ( $n=20$ ), regression ( $n=20$ ) and plateau ( $n=12$ ). Children with early onset autism had the lowest social-communication behaviours at 6 months, with a small decline over the next 18 months. Children with a plateau trajectory of autism had about the same level of social communication as typically developing children at 6 months, but had a slight decline over the next 18 months. Children with regressive autism had significantly higher social-communication than typically developing children at 6 months ( $p$  value not reported), with a rapid decline over time.

By 24 months all children with autism had significantly lower social-communication behaviour than typically developing children ( $p$  value not reported). The early-onset and regressive autism groups had similar social-communication behaviour, but the plateau trajectory group had significantly higher social-communication behaviour than the other autism groups ( $p$  values not reported). Parent-reported regression on ADI-R corresponded poorly with the video analysis: of 20 children whose video analysis showed early-onset, only 8 were determined to have early-onset autism by ADI-R. For the 20 children whose video analysis

showed regression, 11 had regression on ADI-R, and for the 12 whose video analysis showed plateau, only 3 had a plateau trajectory according to the ADI-R.

The authors noted that the finding of rapid early language development in children with regression needs to be replicated before it can be interpreted. Additionally, the small number of participants may have limited the number of trajectories found. About 10% of children did not clearly belong to a single trajectory, having similar likelihood of belonging to multiple trajectories.

This evidence is not likely to affect [NICE CG128](#) because the finding that regression is associated with poorer social communication is broadly consistent with the recommendation to refer to the autism team if regression occurs in children younger than 3 years.

#### **Key reference**

Ozonoff S, Iosif A-M, Young GS et al. (2011) [Onset patterns in autism: correspondence between home video and parent report](#). *Journal of the American Academy of Child & Adolescent Psychiatry* 50: 796–806

## **1.4 [After referral to the autism team](#)**

No new key evidence was found for this section.

## **1.5 [Autism diagnostic assessment for children and young people](#)**

### **Sensitivity and specificity of DSM-IV-TR criteria in children with intellectual disabilities**

[NICE CG128](#) recommends that every autism diagnostic assessment should include a developmental history, focusing on developmental and behavioural features consistent with ICD-10 or DSM-IV criteria (an autism-specific tool to gather this information should be considered). For the entire list of all items that should be included in every autism diagnostic assessment please see the full recommendation in [NICE CG128](#). Additionally, intellectual disabilities should be considered as part of differential diagnosis or as a possible coexisting condition, which may need specific assessments and interpretation of the autism history.

[Hartley and Sikora \(2010\)](#) reported a study investigating DSM-IV-TR criteria in children aged 6–15 years with intellectual disabilities referred to an autism clinic in the USA (n=89, 71% male). Intellectual disability was diagnosed on the basis of DSM-IV-TR criteria for mental retardation. All children were administered a standardised test of intellectual ability appropriate to their age, the ADOS, and the Oral and Written Language Scale (OWLS). Parents completed a measure of adaptive behaviour: either the Vineland Adaptive Behaviour Scale or the Adaptive Behaviour Assessment System, and underwent an interview of DSM-IV-TR criteria for autism. A multidisciplinary team was involved in all assessments of autism, and clinicians often had access to medical and school records. Clinical diagnosis was made by team consensus using all available information. Autism was diagnosed in 31 children and 58 were found not to have autism.

The sensitivity of DSM-IV-TR criteria in children with intellectual disabilities ranged from 33% to 74% and specificity ranged from 45% to 88%. All criteria in the social relatedness domain were seen significantly more in children with autism and intellectual disability compared with children with intellectual disability only (all  $p \leq 0.01$ ). In the communication domain, delay or lack of speech and impaired conversational ability criteria did not differ significantly between children with intellectual disabilities and autism and those who had intellectual disabilities only ( $p$  values not reported), however significantly more children with autism showed stereotyped repetitive or idiosyncratic language and lack of make-believe or imaginative play (both  $p \leq 0.001$ ). In the restricted repetitive or stereotyped patterns domain, only the stereotyped or

restricted pattern of interest criterion was seen significantly more in children with intellectual disability and autism ( $p \leq 0.001$ ).

The authors recognised methodological limitations of their study, including use of parental interviews to endorse DSM-IV-TR criteria rather than directly observing the child, and the use of differing measures of cognitive functioning and adaptive behaviour. Using a binary score of DSM-IV-TR criteria may have missed differences in types or severity of behaviours. The children with intellectual disabilities but found not to have autism by the study had been referred to an autism clinic and therefore probably had more autistic behaviours than other children with intellectual disability, and so may not be representative of the broader population with intellectual disabilities. This may have resulted in underestimation of the usefulness of DSM-IV-TR criteria in detecting autism in children with intellectual disabilities.

This evidence is generally consistent with [NICE CG128](#), which recommends use of criteria such as DSM-IV as part of an extensive autism diagnostic assessment, and supports the use of additional assessments or interpretation of autism history if intellectual disability is a possible differential diagnosis or coexisting condition.

#### **Key reference**

Hartley SL, Sikora DM (2010) [Detecting autism spectrum disorder in children with intellectual disability: which DSM-IV-TR criteria are most useful? Focus on Autism and Other Developmental Disorders 25: 85–97](#)

#### **Validation of DSM-5**

[NICE CG128](#) recommends that every autism diagnostic assessment should include: a developmental history, focusing on developmental and behavioural features consistent with ICD-10 or DSM-IV criteria (an autism-specific tool to gather this information should be considered). For the entire list of all items that should be included in every autism diagnostic assessment please see the full recommendation in [NICE CG128](#).

[Frazier et al. \(2012\)](#) reported results from an exploratory study investigating the use of proposed DSM-5 criteria for classifying autism symptoms by comparing dimensional, categorical and hybrid models. The proposed DSM-5 criteria are based on 2 symptom domains: social communication and interaction and restrictive, repetitive behaviour whereas 3 symptom domains were used in DSM-IV. DSM-V includes 1 diagnosis of autism spectrum disorder instead of the 3 specific diagnoses used in DSM-IV (autism, Asperger's disorder and PDD-NOS). The dimensional model proposed that no distinct category of autism exists, with only degrees of difference between autism symptoms and typical development. The categorical model proposed that a person either has autism or does not. The hybrid model combined dimensional and categorical models.

Data from the US Interactive Autism Network included 14,774 siblings, (8911 with autism spectrum disorder and 5863 without autism). Most clinical diagnoses of autism (93%) were made by a qualified healthcare professional or team. Scores for siblings who did not have autism ( $n=5863$ ) were obtained as a comparison group. The group with autism had a higher proportion of males (82%), but the comparison group, which consisted of siblings without autism had a lower proportion of males (46%).

Primary analyses focused on the SRS ( $n=6949$ ) and sensitivity analyses were done with the SCQ ( $n=14,200$ ). The autism symptoms from the SRS and SCQ were mapped to the DSM-IV-TR and proposed DSM-5 criteria to assess classification accuracy of both DSM editions. A hybrid model including a categorical dimension (autism versus not autism) and 2 symptom dimensions (social communication and interaction, and restricted and repetitive behaviours) was the optimum model for autism, which agreed with about 90% of diagnoses of autism.

DSM-5 criteria had lower sensitivity than DSM-IV-TR (0.81 vs 0.95 respectively) but better specificity (0.97 vs 0.86 respectively). Reducing symptom criteria by 1 gave DSM-5 sensitivity of 0.93 and specificity of 0.95.

The authors recognised the use of a self-selecting population as a potential limitation of this study. The use of siblings without autism as comparators was not considered to be equivalent to an independent control group. Mapping of caregiver reports of symptoms to diagnostic criteria was also identified as a limitation. The authors concluded that despite the limitations of their study, they had generated several promising leads for future studies of the DSM-5 algorithm.

The results of this study suggest that criteria for diagnosis of autism in DSM-5 result in broadly similar diagnoses of autism to those in DSM-IV-TR, so have no effect on recommendations in [NICE CG128](#).

#### **Key reference**

Frazier TW, Youngstrom EA, Speer L et al. (2012) [Validation of proposed DSM-5 criteria for autism spectrum disorder](#). *Journal of the American Academy of Child & Adolescent Psychiatry* 51: 28–40 [\[NIH Public Access author manuscript – full text\]](#)

#### **Early behaviours predicting later diagnosis of autism**

[NICE CG128](#) recommends that clinicians should use information from all sources, together with clinical judgment, to diagnose autism based on ICD-10 or DSM-IV criteria.

[Lord et al. \(2012b\)](#) reported an observational study of the clinical diagnosis of autism spectrum disorders in children and young people aged 4–18 years (n=2102, 86% male) over 12 locations in North America. Each child was assessed with the ADOS, and their parents were interviewed with the ADI-R and the Vineland Adaptive Behavior Scales, and completed the Aberrant Behavior Checklist. A senior clinician made a best estimate clinical diagnosis of autism, PDD-NOS or Asperger's disorder after reviewing all the information and observing the child either in person or on video.

Included children met ADOS or ADI-R criteria for autism, had a non-verbal mental age of at least 18 months and had a best estimate clinical diagnosis of autism, PDD-NOS or Asperger's disorder. Children were excluded: if they had hearing, vision or motor disabilities that were likely to affect interpretation of behavioural data; if they had a known relative (up to third degree) with autism; if a sibling had language or psychological problems related to autism; if they had fragile X syndrome, tuberous sclerosis, Down's syndrome or a clinically significant medical history such as very low birth weight.

The proportion of children assigned each of the diagnoses (autism, PDD-NOS or Asperger's disorder) varied significantly between the 12 sites ( $p < 0.001$ ). However, autism classification based on standardised assessments did not differ significantly between sites (p value not reported). Although the core features of autism varied substantially between individuals, the distributions were similar across sites, with only 1 site outside a 99.5% tolerance band.

Using a sequential model-fitting strategy, classification and regression trees were produced for diagnosis using different sets of predictors. The strongest predictor of a diagnosis of autism was ADOS-measured social communication: 61% of children had moderate-to-severe social communication problems, and they were mainly diagnosed with autism. The remaining 39% of children with milder social communication problems included most of the children with a diagnosis of PDD-NOS or Asperger's disorder and about a third were diagnosed with autism.

More experienced clinicians and those with a doctor of medicine or master's degree diagnosed autism more often; clinicians with a doctor of philosophy degree diagnosed PDD-NOS more often. When site was added to the classification model, it accounted for more



variance than any factor other than ADOS social communication, and the effects of clinician differences mostly disappeared. When demographic, developmental and specific behavioural characteristics were added to the model, verbal IQ was the most important factor: 93% of children with an ADOS social communication score of 12 or more and verbal IQ of 85 or lower were diagnosed with autism.

The authors noted concerns about the robustness of results from classification and regression tree models, which they described as tools for discovery rather than for hypothesis testing and inference. Additionally, some sites had small samples, which might have affected the results. The authors also noted that giving clinicians clearer decision rules or standard training in autism diagnosis might have been useful.

The results of this study suggest that best estimate clinical diagnosis of the subgroups of autism spectrum disorder may vary across locations. Although this study may not be directly applicable to the UK, it provides some support to recommendations in [NICE CG128](#) about using information from all sources together with clinical judgement to diagnose autism.

#### **Key reference**

Lord C, Petkova E, Hus V, et al. (2012b) [A multisite study of the clinical diagnosis of different autism spectrum disorders](#). *Archives of General Psychiatry* 69:306–313

#### **Accuracy of autism diagnostic tools**

[NICE CG128](#) states: do not rely on any autism-specific diagnostic tool alone to diagnose autism.

#### ***Effects of maternal depression on reporting of autism symptoms***

[Bennett et al. \(2012\)](#) reported on a multicentre Canadian longitudinal study of children (n=214, mean age=37.8 months, SD 7.6 months, 85% male) with a previous diagnosis of autism spectrum disorder aged 2–4 years whose mothers completed the Symptom Checklist (SCL)-90. Children meeting ADOS and ADI-R criteria were diagnosed by a multidisciplinary team and were assessed by the SRS.

Children were excluded if they had a known genetic syndrome or neurological basis for autism, and if they did not understand English or French. Maternal depression was defined as depression scores in the 90th percentile on the general symptom index or interpersonal sensitivity or somatising subscales of the SCL-90.

Confirmatory factor analysis was used, in which a baseline model specified covariance between maternal depression and a quantitative estimate of autistic behaviours. A second model was then developed, which added covariance associations between the maternal depression latent variable and the unique variance terms for the maternal report. This model intended to test the hypothesis that maternal depression influences maternal reporting behaviour. A third model was developed to test whether any effects noted by the second model showed a gradient in strength of association between maternal depression and the unique variance in the SRS, ADI-R and ADOS results.

Among the whole sample, 83% (n=178) of children were diagnosed with autism, and 32% (n=66) of mothers met the definition of depression. Significantly higher SRS scores were seen in children whose mothers met the definition of depression (mean=81.21, SD=12.94) compared with those whose mothers did not have depression (mean=72.80, SD=11.34, p<0.001). However, no significant differences were seen for ADI-R (mean=32.09, SD=7.83 vs mean=30.51, SD=7.39 respectively, p=0.16) or ADOS scores (mean=14.68, SD=3.39 vs mean=14.43, SD=3.05 respectively, p=0.68).

Although a gradient effect was seen for depression severity and SRS ADI-R and ADOS scores, the baseline model was noted to be an unacceptable fit. The second model showed an 'excellent' fit, but the association between maternal depression and autism severity was

small and non-significant ( $p=0.67$ ). In the third model, the authors described a clear gradient effect in that the SRS was most strongly correlated with maternal depression ( $p=0.01$ ), ADI-R had a weaker correlation ( $p=0.02$ ) and ADOS was not correlated with maternal depression ( $p=0.68$ ).

The authors stated that their study should be regarded as preliminary and that limitations of this study should be addressed in future work. These included: determining the adequacy of the criterion latent variable, using indicators from 1 instrument reported by different sources (for example, parent, teacher and clinician informants), and using a larger sample size to allow inclusion of additional covariates that might affect results such as child aggression or socioeconomic status.

This study suggests that depression in mothers may affect their reporting of symptoms of autism seen in their children, which lends support to the recommendation in [NICE CG128](#) not to rely only on one tool in diagnosing autism.

#### **Key reference**

[Bennett T, Boyle M, Georgiades K et al. \(2012\) Influence of reporting effects on the association between maternal depression and child autism spectrum disorder behaviors. Journal of Child Psychology and Psychiatry 53: 89–96](#)

#### **ADOS in initial assessment**

[Molloy et al. \(2011\)](#) reported a cohort study investigating the sensitivity and specificity of ADOS when used as an initial diagnostic assessment in children with suspected developmental delay or autism in a developmental and behavioural paediatric clinic in the USA. Children were included ( $n=668$ ) if they were assessed for possible autism with ADOS between 1 January and 31 December 2008.

Children attending the clinic were assessed by a multidisciplinary team: an initial visit with a developmental paediatrician; parent interview and assessments of the child's cognitive abilities and adaptive behaviour with a clinical psychologist; and assessment of communication skills and administration of ADOS with a speech-language therapist. The clinical psychologist and speech-language therapist made written reports of their findings. The developmental paediatrician made the final diagnosis on the basis of all available information (with discussion of any disagreements within the team), then presented the results to parents.

ADOS sensitivity and specificity were reported for autism versus non-spectrum disorders, and autism spectrum disorders other than autism (defined as Asperger's disorder, PDD-NOS, autism spectrum disorder) versus non-spectrum disorders. The communication and social domains were reported together, and the social-affective domain and repetitive and restricted behaviour domain were reported together. Each analysis was reported for 5 ADOS groups defined by the module of ADOS used and their verbal or chronological age: 'module 1, no words'; 'module 1, some words'; 'module 2, less than 5 years'; 'module 2, 5 years and older'; and 'module 3'. Children assessed with ADOS module 4 ( $n=58$ ) or who were older than 12 years and assessed with modules 1 or 2 ( $n=5$ ) were excluded from analysis.

Information on final diagnosis was available for 584 children (97% of included children, 87% male) on electronic medical records. For detection of autism versus non-spectrum disorders, the sensitivity was 67–91% for communication and social domain scores and 82–94% for social affective and repetitive restricted behaviour domain scores. Specificity was 65–95% and 55–81% respectively. For detection of autism spectrum disorders other than autism versus non-spectrum disorders the sensitivity was 75–94% for communication and social domain scores and 72–100% for social affective and repetitive restricted behaviour domain scores. Specificity was 29–81% and 29–60% respectively.

The availability of test scores for cognitive, behavioural and communication assessments varied because the clinic was migrating to electronic records in the year of this study, which the authors noted was a limitation.

This evidence is consistent with the recommendation in [NICE CG128](#) that diagnosis should not rely solely on 1 tool.

#### **Key reference**

Molloy CA, Murray DS, Akers R et al. (2011) [Use of the Autism Diagnostic Schedule \(ADOS\) in a clinical setting](#). *Autism* 15: 143–62

#### **Revised scoring algorithms for ADI-R**

[Kim and Lord \(2012a\)](#) reported on a cohort study investigating the development of new algorithms for scoring the ADI-R in children younger than 4 years. Participants were mainly from 2 projects ('Early diagnosis of autism' and 'First words and toddlers'), undertaken at 2 specialist centres for autism and communication disorders in the USA. Other participants were referred to one of the specialist centres. The sample was restricted to children aged 12–47 months who had a non-verbal mental age of at least 10 months.

Analyses were conducted on 695 participants ('unique cases'), however some participants had repeated assessments, therefore 829 cases ('all cases', 77% male) were studied for some analyses. A case was defined by complete ADI-R and contemporaneous ADOS, non-verbal IQ scores, and a best-estimate clinical diagnosis. Children were grouped by age and language level into 3 groups: all children aged 12–20 months plus children aged 21–47 months who were non-verbal (hereafter referred to as the non-verbal group); children aged 21–47 months with single-word speech; and children aged 21–47 months with phrase speech.

For children participating in the 'Early diagnosis of autism' project, each caregiver was administered the ADI-R, and then ADOS and cognitive testing were done within a few days of the ADI-R. Videotapes of the ADOS, the ADI-R scores, and observations made during the testing were used by a researcher to make an independent best estimate clinical diagnosis of autism. For those participating in the 'First words and toddlers' programme, 2 clinicians used ADI-R and ADOS scores and observations to make a best estimate clinical diagnosis according to DSM-IV criteria. For children referred to the clinic, diagnosis was made by a psychologist or psychiatrist after review of all information.

The distribution of all items in the toddler and standard versions of the ADI-R was examined to identify the items that best differentiated between autism and non-spectrum disorders. The ADI-R was scored: 0 for no definite behaviour of the type specified; 1 for behaviour of the defining type probably present but defining criteria not fully met; and 2 for definite abnormal behaviour of the type described in the definition and coding. A score of 3 was occasionally used to indicate extreme severity, but was classed as a score of 2 in analyses.

Selection criteria for items differentiating between autism and non-spectrum disorders were that no more than 20% of cases of autism scored 0 and no more than 20% of cases of non-spectrum disorders scored 2. These criteria were modified for items deemed to be theoretically important. Factor analyses were done for the 3 developmental groups to verify the fit, and ROC curves were calculated to examine sensitivity and specificity of cut-off scores. An existing algorithm for diagnosing autism was used to compare its predictive validity against that of the algorithms developed in this study.

The mean age of participants was 33.3 months (SD=9.4) and 535 of the 694 participants were male (77%). Autism spectrum disorder was diagnosed in 491 participants; 136 children had non-spectrum disorders; and 67 children were typically developing.



According to the existing clinical cut-off algorithms, ADI-R had sensitivity of 97% and specificity of 43% for detecting autism versus non-spectrum disorders in the non-verbal group; for the single-word group the sensitivity was 91% and specificity was 82%; and in the phrase-speech group the sensitivity was 70% and specificity was 68%.

Initial analyses showed that items appearing in both the standard and toddler versions of the ADI-R were consistently more informative than items in only 1 version. Therefore further analysis was restricted to items appearing in both versions. The new algorithms resulted in: ADI-R sensitivity of 85% and specificity of 70% for detecting autism versus non-spectrum disorders in the non-verbal group; for the single-word group the sensitivity was 94% and specificity was 81%; and in the phrase-speech group the sensitivity was 80% and specificity was 70%.

Limitations reported by the authors included that cut-off scores for the optimum sensitivity and specificity varied by the aged groups studied: the clinical cut-off scores were 11 for the non-verbal group, 8 for the single-word group, and 13 for the phrase-speech group. Additionally different items were used in each algorithm, and participant factors, socioeconomic status and skills of the examiners could have affected the results. The authors stated that further studies are needed to replicate these findings and test the algorithms in populations with and without autism.

Although the results of this study suggest that new algorithms for young children under the age of 4 years may increase the sensitivity and specificity of ADI-R, this increased diagnostic accuracy needs to be confirmed. Therefore this evidence is unlikely to impact on [NICE CG128](#).

#### **Key reference**

Kim SH, Lord C et al. (2012a) [New diagnostic interview-revised algorithms for toddlers and young pre-schoolers from 12 to 47 months of age](#). *Journal of Autism and Developmental Disorders* 42: 82–93.

#### **ADOS plus ADI-R**

[Kim and Lord \(2012b\)](#) reported on a study investigating the combined use of the ADI-R and ADOS in children under the age of 4 years, and the agreement between these tools. Children with complete data (n=595, 79% male) for ADOS, ADI-R, non-verbal IQ and best-estimate clinical diagnosis were included. These data were combined from 2 projects ('Early diagnosis of autism' and 'First words and toddlers') undertaken at a specialist centre for autism and communication disorders in the USA. Additional participants were clinic patients at the same specialist autism and communication disorders centre.

For children participating in the 'Early diagnosis of autism' project, each caregiver was administered the ADI-R, and then ADOS and cognitive testing were done within a few days of the ADI-R. Videotapes of the ADOS, the ADI-R scores, and observations made during the testing were used by a researcher to make an independent best estimate clinical diagnosis of autism. For those participating in the 'First words and toddlers' programme, 2 clinicians used ADI-R and ADOS scores and observations to make a best estimate clinical diagnosis according to DSM-IV criteria. For children referred to the clinic, diagnosis was made by a psychologist or psychiatrist after review of all information.

The mean age of participants was 31.8 months (SD=9.6 months, range 12–47 months). Autism spectrum disorder was diagnosed in 435 children, 113 had non-spectrum disorders, and 47 children had typical development. Children were grouped by age and language level: all children aged 12–20 months plus children aged 21–47 months who were non-verbal (hereafter referred to as the non-verbal group); children aged 21–47 months with single-word speech; and children aged 21–47 months with phrase speech.

The authors analysed the sensitivity, specificity and positive likelihood ratio for each group of children (non-verbal, single-word speech, and phrase speech) for 7 combinations of ADI-R

and ADOS results using both clinical and research cut-offs for ADI-R, giving 21 sensitivity scores for autism, 21 sensitivity scores for Asperger's disorder, 21 specificity scores, and 21 positive likelihood ratios.

The authors noted that the best strategy was using the ADI-R clinical cut-off score and the ADOS together, which had sensitivity of 90–98% and specificity of 80–92% across the groups of children analysed. Sensitivity for detecting autism was best when either ADI-R or ADOS were used with sensitivity of 99–100%, but specificity was lower at 45–85%. If a child's ADI-R and ADOS scores were both judged to be in the range of concern, the odds ratio of having a best estimate clinical diagnosis of autism was 56.19 ( $p < 0.001$ ).

Generally, sensitivity for detecting Asperger's disorder was lower than that for autism, with the ADI-R clinical cut-off plus ADOS showing sensitivity of 82–92%, and either ADI-R or ADOS showing specificity of 97–99%. The positive likelihood ratio seemed to be best for the ADI-R research cut-off plus ADOS (range 6–19). The correlation between ADI-R and ADOS scores was highest in the non-verbal group ( $r = 0.75$ ) compared with the single-word group ( $r = 0.47$ ,  $p < 0.01$ ) and the phrase speech group ( $r = 0.59$ ,  $p \leq 0.01$ ).

The authors noted that the sample size and possible recruitment bias were potential limitations. The lack of blinding of clinicians, who sometimes administered both the ADI-R and then the ADOS, may have affected the ADOS results. Finally, the sample consisted of both clinic and research participants, so the results may not be generalisable to practice in developmental disorders clinics.

These results, showing that 2 diagnostic tools give optimum sensitivity and specificity compared with either tool alone, are consistent with recommendations in [NICE CG128](#) that diagnosis of autism should not rely on the results of 1 tool alone.

#### **Key reference**

Kim SH, Lord C (2012b) [Combining information from multiple sources for the diagnosis of autism spectrum disorder for toddlers and young pre-schoolers from 12 to 47 months of age](#). *Journal of Child Psychology and Psychiatry* 53: 143–51 [[NIH Public Access author manuscript – full text](#)]

#### ***The Child Symptom Inventory (CSI)-4***

[DeVincent and Gadow \(2009\)](#) reported a study assessing 3 algorithms for using the Child Symptom Inventory (CSI)-4 to differentiate between children with autism spectrum disorders and children with ADHD. The authors had previously developed and tested algorithms for scoring the CSI-4 tool, and this study was intended to replicate the authors' original findings.

Participants in the autism group ( $n = 186$ ) were children aged 6–12 years (mean age 8.6 years,  $SD = 1.9$  years) who were consecutive referrals to a developmental disabilities specialty clinic in the USA, and met DSM-IV criteria for autism ( $n = 54$ , 29%), PDD-NOS ( $n = 85$ , 46%) or Asperger's disorder ( $n = 47$ , 25%). Participants with ADHD ( $n = 251$ , mean age 8.0 years,  $SD = 1.4$  years) were recruited from several sources including an outpatient clinic, parent support groups, local media advertisements, and referrals from school and other professionals assessed as part of a diagnostic and follow-up study. The ADHD group included children with chronic multiple tic disorder ( $n = 80$ ). Parents and teachers completed the CSI-4 for all participants.

Autism spectrum disorders were diagnosed using multiple sources of information; ADOS was added to the assessment schedule during the study and was used as part of diagnosis for 53% of children with autism.

Three algorithms for scoring CSI-4 were assessed: an algorithm for teacher scores and 2 algorithms for parent scores. For the teacher algorithm, the original cut-off score of 26 was not optimum. In this study, the optimum cut-off was 27 with sensitivity of 0.84, specificity of 0.72, positive predictive value of 0.74, and negative predictive value of 0.82. The first parent algorithm (generated in a previous study) used a cut-off score of 38 for differentiating between

autism and not autism, which was not optimum for sensitivity and specificity in the current study. In this study, the optimum cut-off score for differentiating between autism and ADHD or not autism was 40, with sensitivity of 0.86, specificity of 0.79, positive predictive value of 0.81 and negative predictive value of 0.85. To improve the specificity and sensitivity, a second algorithm was developed for parent scores with an optimum cut-off score for differentiating between autism and ADHD or non-autism of 60 giving sensitivity of 0.91, specificity of 0.96, positive predictive value of 0.99 and negative predictive value of 0.92.

The authors described their results as preliminary and noted that samples with different characteristics would result in different values for clinical utility, and that coexisting conditions would additionally influence results.

This evidence is unlikely to affect [NICE CG128](#) because the new algorithms with increased sensitivity and specificity for discriminating between the diagnoses of autism and ADHD need assessment in further studies.

#### Key reference

DeVincent CJ, Gadow KD (2009) [Relative clinical utility of three Child Symptom Inventory-4 algorithms for differentiating children with autism spectrum disorder vs. attention-deficit hyperactivity disorder](#). *Autism Research* 2: 312–21

#### Telescoping of caregiver reports on ADI-R

[NICE CG128](#) recommends including in every autism diagnostic assessment: a developmental history, focusing on developmental and behavioural features consistent with ICD-10 or DSM-IV criteria (an autism-specific tool to gather this information should be considered). For the entire list of all items that should be included in every autism diagnostic assessment please see the full recommendation in [NICE CG128](#).

[Hus et al. \(2011\)](#) reported a cohort study investigating the potential effects of 'telescoping' (that is, perceiving distant events as more recent than they are) on the age of developmental milestones reported by carers of children referred for possible autism or developmental delay before age 3 years.

The data were originally collected as part of a longitudinal study of autism in the USA that included 192 consecutive referrals and a control group of 22 children with developmental delay who had not been diagnosed with or referred for assessment of autism. Each participant was assessed with ADI-R, ADOS, and the Vineland Adaptive Behavior Scales and had an assessment of cognitive function. At years 2, 5, and 9 of the study, a best estimate diagnosis was made by a psychologist or child psychiatrist blind to previous diagnoses. The ADI-R had a section that asked the carer to estimate the age at which symptoms first manifested.

82% of the group with autism were male, compared with 44% of the control group. 20 children were excluded because carers did not report any concerns until after the child was referred; 107 children were included in the analysis. The age of first reported concern (measured at age 2 years) did not differ significantly between the group with autism (estimated marginal mean [EMM]=13.9 months, standard error [SE] 0.7 months) and the control group (EMM=12.0 months, SE=1.9 months p value not reported). The age of first reported concern did not differ significantly between time points for either the autism group (EMM at 2 years=13.9 months, EMM at 5 years=14.4 months, and EMM at 9 years=13.1 months) or the control group (EMM at 2 years=12.0 months, EMM at 5 years=12.4 months, and EMM at 9 years=11.2 months, p values not reported).

The reported age of first word increased over time for both the autism group (EMM at 2 years=14.9 months, EMM at 5 years=20.2 months, and EMM at 9 years=30.2 months) and the control group (EMM at 2 years=20.5 months, EMM at 5 years=25.8 months, and EMM at

9 years=35.8 months, p values not reported). The reported age of first phrases increased over time for the autism group (EMM at 3 years=34.41 months, EMM at 5 years=39.5 months, and EMM at 9 years=42.5 months) but data for the control group were not reported. The authors did not discuss possible limitations of their study.

The results suggest that ADI-R scores may be affected by telescoping effects of parents' memories, which clinicians should be aware of when taking developmental histories, but this evidence is unlikely to affect [NICE CG128](#).

#### Key reference

Hus V, Taylor A, Lord C (2011) [Telescoping of caregiver report on the Autism Diagnostic Interview – Revised](#). *Journal of Child Psychology and Psychiatry* 52: 753–60 [\[NIH Public Access author manuscript – full text\]](#)

### The Manchester Inventory for Playground Observation (MIPO)

[NICE CG128](#) recommends including in every autism diagnostic assessment: assessment (through interaction with and observation of the child or young person) of social and communication skills and behaviours, focusing on features consistent with ICD-10 or DSM-IV criteria (an autism-specific tool to gather this information should be considered). For the entire list of all items that should be included in every autism diagnostic assessment please see the full recommendation in [NICE CG128](#). Additionally, if there are discrepancies during the autism diagnostic assessment between reported signs or symptoms and the findings of the autism observation in the clinical setting, clinicians should consider: gathering additional information from other sources or carrying out further autism-specific observations in different settings, such as the school, nursery, other social setting or at home.

[Gibson et al. \(2011\)](#) investigated the reliability and validity of the MIPO tool. Participants were recruited from child and adolescent mental health services (CAMHS) and speech and language therapy services in the Greater Manchester area. 4 child psychiatrists from 2 CAMHS teams made sequential unselected referrals of children with formal clinical diagnoses of externalising disorders (including conduct disorder, ODD and ADHD, n=44), internalising disorders (including primary depression and anxiety, n=19) and autism spectrum disorders (n=39). Children referred from speech and language therapy services (n=42) had no emotional or behavioural difficulties. Overall, 84% of participants were male and the mean age was 8.8 years (SD 1.7 years).

All participants were in mainstream schools, and observations took place during the regular playground sessions at the school. Researchers blind to the child's specific diagnosis, observed unobtrusively from the edge of the play area, but close enough to observe subtle interactions and hear conversation. If school protocols permitted, a control child engaged in similar play activity, matched for age, gender and ethnicity was identified and observed. No additional background data were obtained for the control group, but each child was rated on the Achenbach System of Empirically Based Assessment observational schedule for school-age children.

Inter-observer reliability was assessed by comparing the scores of 2 researchers simultaneously observing and independently scoring a child (in 27% of participants, kappa coefficient=0.70). Test-retest reliability was assessed by comparing the scores of the same researcher observing a child twice in 7–14 days if possible at the same playtime and day of the week (in 14% of participants, kappa coefficient=0.47). Cronbach's alpha for internal consistency was 0.92.

The optimum cut-off score for discriminating between cases and controls was 13, with sensitivity of 0.75, specificity of 0.88, and AUC of 0.90. Classification accuracy was 69% for autism spectrum disorders, 75% for externalising disorders, 81% for speech or language disorders, and 0% for internalising disorders.

The authors noted that the result for discriminating between cases and controls should be interpreted with caution because the observer was not blind to the status of children in the control group. Further limitations identified by the authors included the small sample size and heterogeneity within the different clinical groups, and the lack of permission to assess control children at some sites.

This evidence is consistent with the recommendation in [NICE CG128](#) to include assessment of social and communication skills and behaviours as part of every autism diagnostic assessment, and to observe the child in different settings, such as the school, if the diagnosis is uncertain

#### **Key reference**

Gibson J, Hussain J, Holsgrove S et al. (2011) [Quantifying peer interactions for research and clinical use: The Manchester Inventory for Playground Observation](#). *Research in Developmental Disabilities* 32: 2458–66

## **1.6 [After the autism diagnostic assessment](#)**

No new key evidence was found for this section.

## **1.7 [Medical investigations](#)**

[NICE CG128](#) states – do not routinely perform any medical investigations as part of an autism diagnostic assessment, but consider the following in individual circumstances and based on physical examination, clinical judgment and the child or young person's profile: genetic tests, as recommended by your regional genetics centre, if there are specific dysmorphic features, congenital anomalies and/or evidence of intellectual disability; and electroencephalography (EEG) if there is suspicion of epilepsy.

### **Genetic testing in children with autism**

[Roesser \(2011\)](#) reported a cohort study of genetic testing for causes of autism in a population referred to a developmental service centre in the USA. Medical records for all children referred to the centre with an initial diagnosis of autism were reviewed (n=507, 86% male) and data were extracted including: whether DSM-IV criteria were met and any standardised tests administered, such as ADOS; dysmorphic features; and any genetic testing.

The median age of diagnosis was 4.5 years (range 18 months to 15 years). Most children were diagnosed with PDD-NOS (53%); 36% were diagnosed with autism; and 11% were diagnosed with Asperger's disorder. Genetic testing was undertaken in 207 children. Overall, 13 children (6%) had genetic disorders, but of these only 7 were newly identified, the others had known Down's syndrome or partial trisomy 21.

All tests were normal in 96% of boys. However, significantly fewer girls had normal genetic test results (82%,  $p=0.007$ ). Significant differences were also noted for dysmorphic features, with 97% of tests normal if no dysmorphic features were present compared with 80% of those with dysmorphic features ( $p=0.0018$ ).

The author noted that the retrospective nature of the study, relying on medical records which did not document history in a standard way, and that examinations were not standardised were potential limitations. Additionally, the author noted that an important question to be answered in larger trials was whether any factors such as medical and behavioural history, examination for dysmorphic features or laboratory protocols could increase the diagnostic yield of genetic testing.



The results of this study are consistent with the recommendations in [NICE CG128](#) that genetic testing should be considered if clinically indicated, but not undertaken routinely, in children with autism, and should be undertaken in accordance with local genetic practice.

#### Key reference

Roesser J (2011) [Diagnostic yield of genetic testing in children diagnosed with autism spectrum disorders at a regional referral centre](#). *Clinical Pediatrics* 50: 834–43

#### Electroencephalograms in children with autism

[Duffy and Als \(2012\)](#) reported on a study of EEG coherence (a measure of the connectivity between parts of the brain) in children aged 2–12 years with autism (n=430, 84% male) compared with neurotypical children (n=554, 88% male). Data were obtained from a developmental neurophysiology laboratory in the USA, which maintains a database of EEGs from patients and research participants. Patients are usually referred to rule out epilepsy or sensory processing anomalies by EEG and evoked potentials.

Artefacts were managed in unprocessed EEG signals, by correlating observations during the EEG recording (for example, movement or blinking) with the corresponding part of the EEG recording, which was then excluded from analysis. Further artefact management was done using the source component technique in the BESA 3.5 software.

About 8–20 minutes of awake-state EEG data were used per person, and transformed by the BESA 3.5 software to compute scalp Laplacian or current source density estimates. Spectral coherence was calculated, and a further process to manage artefacts was undertaken. A linear regression model was used in which the dependent variables were targeted for artefact reduction and the independent variables were representative of remaining artefacts subsequently, analysis was done for the targeted data with artefacts removed. Principal component analysis was then done to reduce the number of variables. A 2-group discriminant function analysis was used to determine the significance of a group separation and summarise the classification of each participant.

Forty factors were found to account for 51% of variation. When all 40 coherence factors were used in the primary discriminant function, there was a significant difference between the autism and control groups ( $p < 0.0001$ ).

The authors noted that further studies covering diagnoses associated with or related to autism would be needed to determine whether other conditions could be correctly classified before any conclusions about the use of EEG coherence analysis in diagnosis of autism could be drawn. However, they did not discuss potential limitations of their study.

This evidence suggests that differences may exist in EEG coherence between children with autism and children with typical neurodevelopment. However, the current evidence does not show clinical findings that could be easily used in practice, so no impact is expected on [NICE CG128](#).

#### Key reference

Duffy AH, Als H (2012) [A stable pattern of EEG spectral coherence distinguishes children with autism from neuro-typical controls – a large case control study](#). *BMC Medicine* 10: doi:10.1186/1741-7015-10-64

## 1.8 [Communicating the results from the autism diagnostic assessment](#)

No new key evidence was found for this section.

## 1.9 [Information and support for families and carers](#)

No new key evidence was found for this section.

### Areas not currently covered by NICE guidance

#### Change in diagnosis of autism over time

[NICE CG128](#) does not contain any recommendations covering whether a child who has been diagnosed with autism will remain on the autism spectrum in the future. During development of the guideline this subject was investigated (as documented in the [full version of the guideline](#)), but no robust evidence was available to make recommendations in this area.

[Woolfenden et al. \(2012\)](#) did a systematic review of the stability of diagnosis of autism including 23 studies (n=1363) of at least 30 participants who had been diagnosed with autism or pervasive developmental disorder diagnosed with a standardised diagnostic instrument or standardised diagnostic criteria including DSM-IV-TR, ICD-9 or ICD-10 at both baseline or follow-up. Studies additionally had to have duration of observation of at least 12 months. Diagnostic tools included ADI-R, Childhood Autism Rating Scale, ADOS, the Diagnostic Interview for Social Communication Disorders and the 3Di.

In 11 studies of children aged under 3 years (n=653), the proportion of children with a diagnosis at follow up was 53–100%. Across studies, up to 30% of children either moved from a diagnosis of autism to another autism spectrum disorder or moved off the spectrum. In 5 studies of children aged 3–5 years (n=219), 73–100% of children still had a diagnosis of autism at follow-up. Up to 15% moved from a diagnosis of autism to another autism spectrum disorder and up to 20% of children moved off the spectrum. In 7 studies of children older than 5 years (n=496), 81–100% had a diagnosis at follow-up, up to 17% moved from autism to another autism spectrum disorder, and up to 16% moved off the spectrum.

Risk of bias was variable, with only 2 studies deemed to be at low risk of bias for all 5 quality criteria. The main risks of bias related to use of clinic-based samples and lack of blinding of outcome assessors to baseline measures. Although no formal assessment of heterogeneity between studies was done, the authors noted that there were differences in diagnostic practices, age at baseline diagnosis, duration of follow-up, and cognitive impairments of participants. Additionally, data could not be extracted to allow comparison of predictors of change or stability of diagnosis.

This evidence suggests that over time, children may show different symptoms of autism that could change their diagnosis, which is consistent with the recommendation to keep children under review if there is uncertainty about diagnosis. The finding that children may have future developments that mean they move off the spectrum is unlikely to affect [NICE CG128](#) because of limitations in the current evidence base.

#### Key reference

[Woolfenden S, Sarkozy V, Ridley G et al. \(2012\) A systematic review of the diagnostic stability of autism spectrum disorder. Research in Autism Spectrum Disorders 6: 345–54](#)

## 2 New evidence uncertainties

No new evidence uncertainties were identified during the Evidence Update process, however current uncertainties for autism can be found in the [UK Database of Uncertainties about the Effects of Treatments](#) (DUETs) at and in the [NICE research recommendations database](#).

UK DUETs was established to publish uncertainties about the effects of treatments that cannot currently be answered by referring to reliable up-to-date systematic reviews of existing research evidence.



# Appendix A: Methodology

## Scope

The scope of this Evidence Update is taken from the scope of the reference guidance:

- [Autism diagnosis in children and young people](#). NICE clinical guideline 128 (2011).

## Searches

The literature was searched to identify studies and reviews relevant to the scope. Searches were conducted of the following databases, covering the dates 17 August 2009 (the end of the search period of NICE clinical guideline 128) to 29 October 2012:

- CDSR (Cochrane Database of Systematic Reviews)
- CENTRAL (Cochrane Central Register of Controlled Trials)
- CINAHL (Cumulative Index to Nursing and Allied Health Literature)
- DARE (Database of Abstracts of Reviews of Effects)
- EMBASE (Excerpta Medica database)
- HTA (Health Technology Assessment) database
- MEDLINE (Medical Literature Analysis and Retrieval System Online)
- NHS EED (Economic Evaluation Database)
- PsycINFO

Table 1 provides details of the MEDLINE search strategy used, which was adapted to search the other databases listed above. The abbreviation ASD (autism spectrum disorder) was removed from the search for this Evidence Update because of a high number of irrelevant results (for example ASD meaning atrial septal defect). The search strategy was used in conjunction with validated Scottish Intercollegiate Guidelines Network [search filters for RCTs, systematic reviews, observational studies and diagnostic test accuracy studies](#).

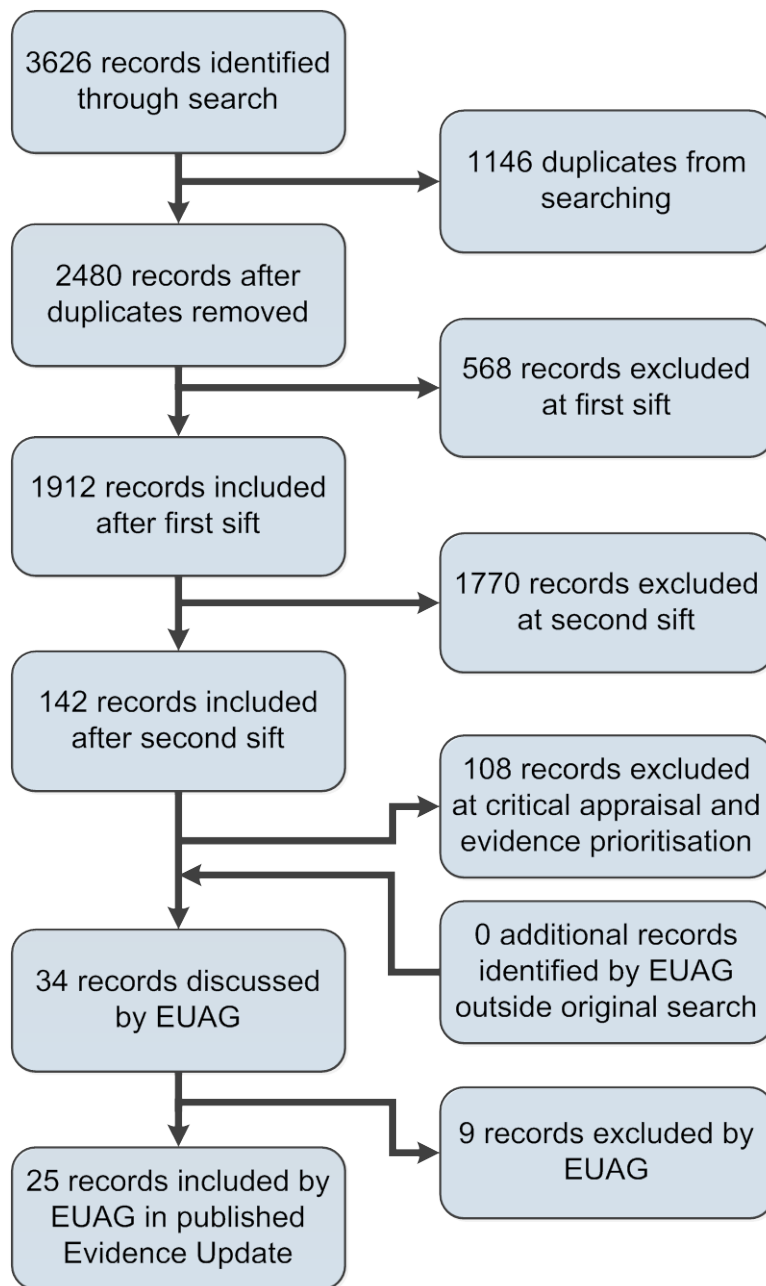
Figure 1 provides details of the evidence selection process. The long list of evidence excluded after review by the Chair of the EUAG, and the full search strategies, are available on request from [contactus@evidence.nhs.uk](mailto:contactus@evidence.nhs.uk)

There is more information about [how NICE Evidence Updates are developed](#) on the NICE Evidence Services website.

**Table 1 MEDLINE search strategy (adapted for individual databases)**

1	Autistic disorder/	5	Pervasive developmental disorder\$.ti,ab.
2	Kanner.ti,ab.	6	Pdd.ti,ab.
3	(Autistic or autism or asperger\$.ti,ab.	7	Pdd-nos.ti,ab.
4	Child development disorders, pervasive/ or asperger syndrome/	8	Or/1-7

**Figure 1 Flow chart of the evidence selection process**



EUAG – Evidence Update Advisory Group

# Appendix B: The Evidence Update Advisory Group and Evidence Update project team

## Evidence Update Advisory Group

The Evidence Update Advisory Group is a group of topic experts who review the prioritised evidence obtained from the literature search and provide the commentary for the Evidence Update.

**Professor Gillian Baird – Chair**

Consultant Paediatrician, Guy's and St Thomas' NHS Foundation Trust, London

**Dr Carole Buckley**

General Practitioner, Bristol

**Professor Tony Charman**

Chair in Clinical Child Psychology, Institute of Psychiatry, King's College London

**Professor Jonathan Green**

Professor of Child and Adolescent Psychiatry, University of Manchester

**Dr Diana Howlett**

Consultant Community Paediatrician, North Bristol NHS Trust

**Professor Ann Le Couteur**

Professor of Child and Adolescent Psychiatry, Institute of Health and Society, Newcastle University

**Professor Emily Simonoff**

Professor of Child & Adolescent Psychiatry, Institute of Psychiatry, King's College London

**Penny Williams**

Lead Clinical Specialist Speech and Language Therapist for Autism, Guys and St Thomas NHS Foundation Trust Community Services

## Evidence Update project team

**Marion Spring**

Associate Director

**Chris Weiner**

Consultant Clinical and Public Health Adviser

**Cath White**

Programme Manager

**Lynne Kincaid**

Medical Writer

**Bazian**

Information specialist support