In press, *British Journal of Psychiatry*_Sept 2007

Efficacy of three screening instruments in the identification of autism spectrum disorder

Tony Charman*, Gillian Baird, Emily Simonoff, Tom Loucas, Susie Chandler,

David Meldrum, Andrew Pickles

Correspondence to:

*Tony Charman, Behavioural & Brain Sciences Unit, UCL Institute of Child Health, 30 Guilford Street, London, WC1N 1EH, UK. t.charman@ich.ucl.ac.uk

Running head: Efficacy of screening instruments for ASD

Word count text (excluding References, Tables and Figures): 4281

Efficacy of three screening instruments in the identification of autism spectrum disorder

ABSTRACT

Background: Screening instruments to identify cases of autism spectrum disorders (ASD) have not previously been compared in the same sample.

Aims: To compare the Social Communication Questionnaire (SCQ), the Social Responsiveness Scale (SRS) and the Children's Communication Checklist (CCC).

Method: Screen and diagnostic assessments on 119 nine to thirteen-year-old children with special educational needs (SEN) with and without ASD were weighted to estimate screen characteristics for a realistic target population.

Results: The SCQ performed best (AUC 0.90; Sensitivity=0.86; Specificity=0.78). The SRS had a lower AUC (0.77) with high Se (0.78) and moderate Sp (0.67). The CCC had a high sensitivity but lower specificity (AUC=0.79; Se=0.93; Sp=0.46). The AUC of the SRS and CCC was lower for children with IQ<70. Behaviour problems reduced specificity for all screens.

Conclusions: The SCQ, the SRS and the CCC showed strong to moderate ability to identify ASD cases in this 'at risk' sample of school-age children with SEN.

Keywords: Autism, Autism Spectrum Disorders, Pervasive Developmental Disorders, Screening, Identification, SNAP cohort

Words: 154

INTRODUCTION

There is considerable interest in screening instruments to identify children with possible autism spectrum disorders (ASD) for a more in-depth diagnostic assessment. Recently developed screening instruments that have demonstrated promising instrument properties in initial validation studies include the Social Communication Questionnaire (SCQ: Sensitivity (Se)=0.85; specificity (Sp)=0.75; Berument *et al*., 1999) and the Social Responsiveness Scale (SRS: Se=0.85; Sp=0.75; Constantino & Gruber, 2005). The Children's Communication Checklist (CCC; Bishop, 1998) has a Pragmatic Composite (PC) subscale that has been shown to discriminate well between autism and non-autism cases (Bishop & Baird, 2001). For clinicians and researchers a key consideration is *which* screen is most appropriate to our service or our study? In the present study we directly compare the instrument properties of the SCQ, SRS and CCC to identify ASD cases in the Special Needs and Autism Project (SNAP; Baird *et al*., 2006) cohort of 9 to 14-year-old children with special educational needs (SEN) with and without ASD.

METHOD

The study was approved by the South East Multicentre Research Ethics Committee (REC) (00/01/50).

Screening instruments

Social Communication Questionnaire (SCQ; Rutter et al., 2003)

The SCQ is a 40-item parent-report questionnaire asking about characteristic autistic behaviour. Each item is scored 0 or 1 with 1 being the score for endorsement of each autism symptom. Total scores can range from 0 to 39 (the first item is a language screening question that is not included in the total score). It is based on the Autism Diagnostic Interview-Revised (ADI-R; Lord *et al*., 1994) and has established validity with a diagnosis of autism (Berument

et al., 1999). Nineteen items rate current behaviour and 20 rate behaviour when the child was aged 4-5 years of age. The recommended cut-off for ASD/PDD is =>15.

Social Responsiveness Scale (SRS; Constantino & Gruber, 2005)

The SRS is a 65-item rating scale asking about characteristic autistic behaviour. Each item is scored from 0 ('never true') to 3 ('almost always true') as best describes the child's behaviour in the past 6 months. Total scores can range from 0 to 195. For the present analysis the cut-point that best discriminated children with and without ASD (>=75) was chosen (Constantino & Gruber, 2005, p.38). SRS scores discriminate between children with and without ASD and are strongly correlated with ADI-R domain scores (r=0.65 to 0.77; Constantino *et al*., 2003).

Children's Communication Checklist (CCC; Bishop, 1998)

Although not developed as a screen for ASD, the CCC is a 70-item rating scale asking about language and communication impairments, divided into 9 subscales. Each item is scored 0 ('does not apply'), 1 ('applies somewhat'), 2 ('definitely applies') to missing value ('unable to judge'). Some items ask about language and communication impairments and others about competencies. Two subscales assess aspects of language structure (syntax and speech); two assess aspects of autistic behaviour (social relationships and interests); five assess aspects of pragmatic communication (inappropriate initiation, coherence, stereotyped conversation, use of context, and rapport). These latter 5 scales can be combined into a Pragmatic Composite (PC). Bishop (1998) found that a CCC PC score of <=132 best identified children with pragmatic language impairment (PLI) and this also discriminated well between children with and without autism in a clinical sample, though less well between cases with Asperger syndrome/PDD-NOS and cases with ADHD (Bishop & Baird, 2001). The present study was started before the publication of the Children's Communication Checklist-Version 2 (CCC-2; Bishop, 2003).

SNAP cohort

SCQ, SRS and CCC data were collected on a subgroup of the Special Needs and Autism Project (SNAP) sample: a cohort of children 'at risk' for ASD due either to having received a local clinical ASD diagnosis or by having a Statement of Special Educational Needs (SEN) (Baird *et al*., 2006). A Statement of Special Educational Needs (SEN) is a legal document issued by the UK local educational authority when children require significant additional support in school due to any learning and/or behavioural problems. In the larger study, N=255 children from the cohort received a comprehensive diagnostic assessment (described below). We have previously reported on the screening properties of the SCQ in the total sample finding similar discrimination between ASD and non-ASD cases as in the original validation sample (Se=0.88; Sp=0.72; Chandler *et al*., in press). Parents of a sub-sample of children also completed both the SRS and the CCC in addition to the SCQ (N=119). This afforded us the opportunity to directly compare the instrument properties of the 3 screens in the same sample.

As part of a prevalence study of ASD, within a total population cohort of 56,946 children born between July 1st 1990 and December 31st 1991 (Baird *et al*., 2006) all those with a current clinical diagnosis of PDD (n=255) or considered 'at risk' for being an undetected case by virtue of having a Statement of SEN (n=1,515) were screened using the SCQ. A total of 1,066 SCQs were returned completed (return rate 60.2%), 31 families declined further participation, leaving 1,035 (return rate 58.5%) who returned the SCQ and opted in for further assessments. Mean (SD) age at SCQ screening in the whole SNAP sample was 10.3 (0.4) years.

A stratified subsample (by coincidence, also n=255) received a comprehensive diagnostic assessment including standardized clinical observation (Autism Diagnostic Observation Schedule – Generic (ADOS-G); Lord *et al*., 2000) and parent interview

assessments of autistic symptoms (ADI-R; Lord, *et al*., 1994), language and IQ, psychiatric comorbidities and a medical examination. The team used ICD-10 research criteria to derive a clinical consensus diagnosis of childhood autism, other ASDs or non-ASD (see Baird *et al*., 2006; for details). For 36 randomly selected cases project consensus diagnoses were compared to those of 8 internationally recognised experts using ICD-10 criteria (2 experts independently rated ADI, ADOS, psychometric findings and a clinical vignette for each case). Agreement between project consensus and expert autism/ASD/no-ASD diagnostic categories was 93% with (weighted) kappa 0.77 (see Baird *et al*. 2006; Figure 1, for details).

The following child characteristics were also collected: IQ, severity of autism symptoms measured by ADI-R and ADOS-G algorithm total scores; a total count of ICD-10 symptoms (0-12) was systematically completed as part of the diagnostic review process of every case; parent and teacher report of emotional and behavioural problems; adaptive behaviour was assessed using the Vineland Adaptive Behavior Scales (VABS; Sparrow *et al*, 1984). IQ was measured using the Wechsler Intelligence Scale for Children (WISC-III-UK; Wechsler, 1992) (N=118). The (weighted) mean (SE) Full Scale IQ of the sample was 73.4 (1.6) and the range was 40 to 136. 56% (weighted) of children had IQ<70. One child could not complete the WISC and their IQ was derived using Raven's Standard Progressive Matrices (SPM; Raven et al, 1990).

Parents and teachers completed the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997, 2001). Each subscale has 5 questions that are rated 'not true', 'somewhat true' and 'certainly true' and score 0-2 with higher scores indicating higher pathology. The 4 subscales *emotional problems, peer problems, conduct problems* and *hyperactivity* are summed to create a total problem score (range 0-40). Compared to SDQ UK norms (Meltzer et al, 2000), children whose teacher (>=16) and/or parent (>=17) 'total problem score' fell

about the 10% percentile were considered to have a high rate of behaviour problems in the present analysis.

Order of completion of assessments

In the larger study the SCQ was used as the initial screening instrument to identify cases for in-depth diagnostic assessment so for all cases the SCQ was completed before the diagnostic assessments. Mean (SE) age at SCQ screening in the N=119 subsample with data on all 3 screens was 10.2 (0.4) years; range 9.5 to 11 years. The CCC was completed by parents immediately prior to the diagnostic assessment (Mean (SE) age = 12.0 (0.1) years; range 9.8 to 13.9 years). The SRS was completed at (Mean (SE) age = 12.6 (0.4) years; range 11.8 to 13.2 years). In 50 cases this was in advance of, and in 69 cases this was following, the diagnostic assessment and completion of the CCC. As part of the consensus clinical diagnostic process scores on the 3 screening instruments were not consulted (see Baird *et al*., 2006, for details). For the purposes of the current paper only children whose parents had completed all 3 questionnaires were included in the analysis. The children for whom all 3 screens were completed (N=119) differed from the remainder of the cohort (N=136) in terms of IQ (mean (SE) = 78.5 (1.8) vs. 67.4 (2.2); ANOVA $F(1,251)=15.0$, $p<.001$) but not parental education or child symptom severity.

Statistical Analysis

Stratification of the ASD/SEN sample was based on whether or not a child had a locally recorded ASD diagnosis (yes/no) and 4 levels of SCQ score (low (<8), moderately low (8-14), moderately high (15-21), high (≥22); see Baird *et al*., 2006; Figure 1 for details). Use of weights allowed all statistics such as means, group differences and screen performance measures to be presented as target population estimates, taking account not only of the differences in sampling proportions according to SCQ score and local ASD diagnosis, but also the differential response to the SCQ associated with a prior local ASD diagnosis, health

district and child's sex. Wald test statistics (adjusted t and F-tests) and p-values were calculated using the linearisation version of the robust parameter covariance matrix as implemented by the svy procedures of Stata 9 (Stata, 2005). A receiver-operator-characteristic (ROC) Area-Under-the-Curve (AUC) analysis was performed to assess and compare the discriminant power of the screening instruments in distinguishing ASDs (including autism) cases from non-ASD cases (Dunn, 2000; Hanley & McNeil, 1982). Application of the weights ensured that this comparison was fair in spite of the SCQ-stratified sample design. Confidence intervals for weighted AUC estimates, and tests were obtained by bootstrap resampling ROC procedures of Stata 9 (Stata, 2005), reverse coded in the case of the CCC.

RESULTS

Thirty-three cases received a clinical consensus diagnosis of childhood autism; 37 a clinical consensus diagnosis of other ASDs and 49 children did not meet clinical consensus diagnosis for autism or other ASD. Of the 37 children with an 'other ASD' diagnosis: 2 met ICD-10 criteria for 'atypical autism' due to late onset; 2 met ICD-10 criteria for 'atypical autism' due to an insufficient number of areas of abnormality; 29 met ICD-10 criteria for 'PDD other' due to sub-threshold symptomatology; 3 met ICD-10 criteria for 'PDD unspecified' due to lack of information (incomplete assessment, adopted children for whom early history was not available); and 1 met ICD-10 criteria for overactive disorder associated with mental retardation and stereotyped movements. Non-ASD diagnoses included intellectual disability (DSM-IV-TR 'mental retardation') and learning disabilities (N=27), language delay/disorder (N=7), hyperkinetic and/or conduct disorder (N=6) and a variety of other medical, sensory and developmental diagnoses (N=9).

Table 1 shows the weighted mean (SE) score of the sample on the three screens by consensus diagnostic group. As would be expected, the childhood autism cases scored higher than the other ASD and non-ASD cases on the SCQ and SRS and lower on the (reverse

scored) CCC PC. Similarly, the other ASD cases scored higher than the non-ASD cases on the SCQ and SRS and lower on the CCC PC. For the SCQ all three group-by-group comparisons were significant (other ASD vs. non-ASD: $F(1,118)=39.8$, $p<.001$; Childhood autism vs. non-ASD: $F(1,118)=186.4$, $p<.001$; Childhood autism vs. other ASD: $F(1,118)=28.2$, $p<.001$). For the SRS and CCC all three comparisons approached or reached significance, except for the Childhood autism vs. other ASD comparison on the SRS (SRS: other ASD vs. non-ASD: $F(1,118)$ 6.2, $p<.05$; Childhood autism vs. non-ASD: $F(1,118)=38.7$, $p<.001$; Childhood autism vs. other ASD: $F(1,118)=2.7$, $p=.10$; CCC: other ASD vs. non-ASD: $F(1,118)=13.2$, $p<.001$; Childhood autism vs. non-ASD: $F(1,118)=29.5$, $p<.001$; Childhood autism vs. other ASD: $F(1,118)=5.3$, $p<.05$).

-------Table 1 about here------

Total scores on the three screens were highly and significantly correlated (weighted correlation coefficients SCQ–SRS=0.68; SCQ–CCC=-0.66; SRS–CCC=-0.75, all $p<.001$). Table 2 shows the correlations between total scores on the three screens and scores on the ADI-R, ADOS-G, ICD-10 symptom count, IQ, BPVS, Vineland ABC and parent and teacher SDQ. All three screens were more highly correlated with ADI-R total score and ICD-10 symptom count than ADOS-G total score. SCQ and SRS scores were unrelated to IQ and scores on the CCC were only weakly associated, with lower IQ being associated with poor pragmatic ability ($r=0.20$, $p<.05$). All three screens were unrelated to language ability as measured by the BPVS. Scores on all three screens was also significantly associated with the adaptive behaviour composite of the VABS and with parent, but less so with teacher, completed SDQs.

-------Table 2 about here------

The area under the curve (AUC), Sensitivity (Se), Specificity (Sp), Positive predictive value (PPV) and Negative predictive value (NPV) of the three screens in predicting ASD

versus non-ASD status are shown in Table 3 and the ROC curves are shown in Figure 1. The SCQ had a higher AUC (0.90) than the SRS (0.77; p=.05) and the CCC (0.79, p=.05), reflecting both its high sensitivity (Se=0.86) and specificity (Sp=0.78). The AUC of the SRS and CCC did not differ from one another (p=0.84). The SRS had high Se (0.78) but only moderate Sp (0.67); whilst the CCC had a high sensitivity but a low specificity (Se=0.93; Sp=0.46).

-------Table 3 and Figure 1 about here------

In order to examine whether the three screens performed differently in subsamples of children (children with low (<70) vs. high (>=70) IQ; children with vs. without parent and/or teacher rated borderline behavioural problems on the SDQ) AUC analyses were repeated for these subgroups. Note that these analyses should be treated with caution as the confidence intervals for some of the parameters are wide, reflecting smaller subsamples. However, while the SCQ and CCC performed similarly in the low IQ subsample as they did for the whole sample, the SRS had a lower AUC (0.67) reflecting its lower specificity (0.57). In the high IQ subsample the AUC was similar for all 3 screens (SCQ=0.90; SRS=0.87; CCC=0.88). All three screens showed lowered specificity in the subsample with elevated behavioural problems (SCQ=0.57; SRS=0.41; CCC=0.30).

DISCUSSION

Summary of the screening properties

The SCQ performed similarly to the initial validation study (Se=0.85; Sp=0.75; Berument *et al*., 1999; Rutter *et al*., 2003) and somewhat better than in several recent studies that have included younger children that have reported reduced sensitivity (0.71; Corsello *et al*., in press; 0.71; Eaves *et al*., 2006a; 0.67; Lee *et al*., 2007) or reduced specificity (0.58; Allen *et al*., in press; 0.54; 0.71; Corsello *et al*., in press; Eaves *et al*., 2006b). The present study included a restricted age range only but in a large sample Corsello *et al*. (in press) found

that the sensitivity of the SCQ increased with age (Table 3; Corsello *et al*., in press), perhaps reflecting the emergence of the full range of autistic symptoms over time. This is supported by comparison of the mean SCQ scores for children with a diagnosis of childhood autism in the present (25.8) and previous studies. Whilst our figure is very similar to that obtained in the Berument *et al*. (1999) study (25.2) it is higher than that obtained in the Corsello *et al*. (in press) study that included children aged 2 to 16 years (20.3) and also the Eaves *et al*. (2006b) that included children aged 2 to 6 years (19.2).

The SRS had a lower sensitivity in the current sample than in the original validation study (Se=0.85; Sp=0.75; Constantino & Gruber, 2005) and both the SRS and CCC had reduced specificity (0.57 and 0.41, respectively) in the low IQ subsample. The increased specificity of the SCQ compared to the SRS and CCC might be due to its fuller coverage of the third autism symptom domain of restricted and repetitive behaviors and interests. All 3 screens showed high discrimination between ASD and non-ASD children with IQ>=70 with AUC between 0.87 and 0.90. The CCC had a high sensitivity but low specificity, reflecting its broader cut-point for 'pragmatic impairment', rather than ASD per se. For some purposes, for example 'screening out' potential ASD cases when characterising a non-ASD comparison group in a research study, high sensitivity is prioritised and lower specificity does not bring costs and on the basis of the present data the CCC could be used for such a purpose. A recent report found that the CCC discriminated well between children with autism, children with attention deficit hyperactivity disorder (ADHD) and typical controls (discriminant function classification 78% in Study 1; 77% in Study 2; Geurts *et al*., 2004).

All 3 screens had lower specificity in the subsample with elevated levels of behavioural problems as measured by the SDQ. It is likely that that in response to questions on the screens that are meant to be measuring autism symptoms, some parents might be endorsing items that reflect aspects of their child's emotional, hyperactivity or conduct

difficulties. One previous study has reported high scores on the SCQ for children with mood and anxiety disorders in whom a clinical diagnosis of PDD had been excluded (Towbin *et al*., 2005), although such disorders are unlikely to be common in our sample as they usually do not form a reason for SEN registration. In the Towbin *et al*. study significantly more children fell above the ASD cut-point on the SRS and the Social Interaction Deviance Composite on the CCC-2 than above the ASD cutpoint on the SCQ (Figure 1, p. 458; Towbin *et al*., 2005).

In addition to the prevalence of ASD cases; in any particular clinical setting or research study the characteristics (e.g. clinical diagnosis, IQ, age) of the ASD and non-ASD cases, family factors (e.g. parental education, parental knowledge about autism), and methodological factors including whether the screen was completed prior to or after a diagnostic assessment, will also affect how a screen performs. Another factor that affects screen performance in relation to a clinical diagnosis of ASD is the time period of the behaviour sampled and the three instruments used in this study differed in this respect. Whilst the CCC and SRS ask parents to rate current behaviour (for the SRS the last 6 months; unspecified for the CCC), approximately half the items on the SCQ ask about behaviour in the 4-5 year period when autism symptoms may be at their most prototypical. One further consideration that the current study cannot address is whether screens perform differently in different countries due to cultural interpretation of the behaviours enquired about.

Example scenarios comparing use of the three screens

While statistics such as AUC have a statistical meaning they can be hard to translate into everyday examples to guide clinicians and researchers. To illustrate the potential impact of the different screen parameters on a hypothetical research study or clinical service we will outline two scenarios that summarise the consequences of choice of screen. In both scenarios assume that among 250 children to be screened, 100 are ‘true cases’ of ASD and 150 are ‘true non-ASD cases’. Using estimates from the present analysis this translates into: the number of

the 100 true ASD cases that are screen positive ('true positives') (SCQ=86; SRS=78; CCC PC=92); the number of the 100 true ASD cases that are screen negative ('false negatives') (SCQ=14; SRS=22; CCC PC=8); the number of the 150 true non-ASD cases that are screen negative ('true negatives') (SCQ=117; SRS=100.5; CCC PC=69); and the number of the 150 true non-ASD cases that are screen positive ('false positives') (SCQ=33; SRS=49.5; CCC PC=81). The relative importance of these figures depends on the purpose of using the screen in a particular study/service.

In the *first scenario*, consider the screen is being used to estimate the number of children within a special school system who have an ASD for purely administrative reasons. That is, there is no consequence or cost (such as a follow-up assessment) attached to being screen positive. In this scenario, the estimated prevalence of ASD will vary by a factor of ~50% depending on whether one is using the SCQ (119 screen positives i.e. cases identified), the SRS (127.5 cases identified) or the CCC PC (173 cases identified), with the latter clearly over-estimating the 'true' prevalence.

In a *second scenario*, consider a clinical service screening speech and language therapy referrals to identify children who should be 'fast tracked' into an expensive (and for parents sometimes anxiety provoking) multi-disciplinary assessment. For this aim high sensitivity is required for the screening procedure to be clinically useful. To maximize case-finding efficiency one could consider the proportion of cases correctly identified by the screen compared to cases missed by the screen (SCQ=6.14; SRS=3.55; CCC PC=11.5). However, one would also want to minimize screen false positives in order not to use expensive assessment resource on children who do not have ASD and to avoid provoking unnecessary anxiety in parents. Here the most relevant figure is the number of false positives (SCQ=33; SRS=49.5; CCC PC=81). In this scenario the SRS was least efficient in terms of case-finding and the CCC least cost-effective, with the SCQ performing best. Other scenarios would

produce different outcomes and the choice of which screen to use and which criterion to adopt could depend on the relative costs of false positives and false negatives – although these costs fall on different parties (false positives tend to cost services; whereas false negatives tend to cost the child and parent). Clinicians and researchers need to estimate as best they can the implications for their service/study of which screen to use for any particular purpose.

Strengths and limitations of the present study

The strengths of the present study include: testing the ability of different screening instruments to identify cases of ASD in the same sample; the generalisability of the findings due to the population weighting procedure; the calculation of confidence intervals around the instrument parameter estimates; and the inclusion of both low and high IQ children. The comprehensive diagnostic assessment and use of a clinical consensus decision-making process that was corroborated by independent expert rating (see Baird *et al*., 2006) are also strengths. Furthermore, the screens were able to differentiate those with ASD from those without ASD but with intellectual disability and language delay despite the considerable degree of symptom overlap between these conditions, especially in the area of impairments in communication.

One limitation of the present study is that the age of the 'at risk' sample at the time of screening (9-13 years) is later than would be required for first-level screening of young children, although it is still an age at which many children get referred for possible ASD, in particular to child and adolescent mental health services (see Skuse et al., 2005). Second, the SCQ was derived from items on the ADI-R and information from the ADI-R was part of the information used to achieve a clinical consensus diagnosis. This might, in part, explain the higher prediction of the SCQ compared to the SRS and CCC found in the current study. Another limitation is the relatively modest sample size, in particular when the high/low IQ and high/low behavioural problem subsamples were examined, leading to relatively wide CIs.

However, the sample size compares well with the majority of published data available on the SCQ, SRS and CCC.

Clinical implications

A score on a screening instrument is not a diagnosis. For diagnostic assessment a full parental interview regarding current and past development and behaviour, and structured observation of the child, preferably including a peer-group setting, is essential. Corsello *et al*. (in press) found that the addition of information from the ADOS-G to information from the SCQ significantly improved specificity to detect ASD cases. They recommend that for some clinical or research purposes a multi-stage assessment beginning with the SCQ; followed by the ADOS-G; and then including information from the time intensive ADI-R assessment might be appropriate. This study demonstrates that for some clinical and research purposes the SCQ, and to a lesser extent the SRS and CCC, can all be efficient first-level screens to identify children with possible ASD for a more in-depth assessment. Child characteristics such as IQ and behavioural problems affect performance of the screens and need, in addition to considerations regarding consequences/costs of a screen positive or negative result, to be borne in mind when choosing which screen to employ for any particular clinical or research purpose.

Affiliations and Acknowledgements

Tony Charman, PhD, UCL Institute of Child Health, London; Gillian Baird, FRCPCH, Guy's & St Thomas' NHS Foundation Trust, London; Emily Simonoff, MD, Institute of Psychiatry, Kings College London; Tom Loucas, PhD, School of Psychology and Clinical Language Sciences, University of Reading; Susie Chandler, PhD, UCL Institute of Child Health, London; David Meldrum, FACP, Chatswood Assessment Centre, Sydney; Andrew Pickles, PhD, Division of Epidemiology & Health Sciences, University of Manchester. We are grateful to the children and families who participated in the study and to the clinical teams in South Thames, whose collaboration made the study possible. We thank the expert group, Patrick Bolton, Antony Cox, Anne Gilchrist, Rebecca Landa, Ann Le Couteur, Catherine Lord, Lennart Pedersen and Michael Rutter. Thanks also to Iris Carcani-Rathwell, Greg Pasco, Samantha Ross, Emma Rowley, Vicky Slonims and Martha Turner for their help with assessments; and to Dorothy Bishop for comments on an earlier version of this manuscript.

Abbreviations: Area Under the Curve (AUC); Autism Spectrum Disorders (ASD); Children's Communication Checklist (CCC); Negative Predictive Value (NPV); Positive Predictive Value (PPV); Pragmatic Composite (PC); Sensitivity (Se); Social Communication Questionnaire (SCQ); Social Responsiveness Scale (SRS); Special Needs and Autism Project (SNAP); Special Educational Needs (SEN); Specificity (Sp)

Declaration of interest: The study was funded by the Wellcome Trust and the Department of Health, UK. AP receives royalties from the SCQ and ADOS-G. All other authors have no conflict of interest.

REFERENCES

Allen, C.W., Silove, N., Williams, K. *et al* (in press) Validity of the Social Communication Questionnaire in assessing risk of autism in preschool children with developmental problems. *Journal of Autism and Developmental Disorders.*

American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders 4th Edn. – Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association

Baird, G., Simonoff, E., Pickles, A. *et al* (2006) Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368*, 210-215

Berument, S. K., Rutter, M., Lord, C., *et al* (1999) Autism screening questionnaire: diagnostic validity. *British Journal of Psychiatry, 175*, 444-451

Bishop, D. V. M. (1998) Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*, 879-891.

Bishop, D.V.M. (2003) *Children's Communication Checklist-Version 2 (CCC-2)*. London: Psychological Corporation

Bishop, D.V.M. & Baird, G. (2001) Parent and teacher report of pragmatic aspects of communication: Use of the Development of the Children's Communication Checklist in a clinical setting. *Developmental Medicine and Child Neurology, 43*, 809-818.

Chandler, S., Charman, T., Baird, G. *et al* (in press) Validation of the Social Communication Questionnaire in a population cohort of children with autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*.

Constantino, J.N., Davis, A.A., Todd, R.D. *et al* (2003) Validation of a brief quantitative measure of autistic traits: Comparison of the Social Responsiveness Scale with the Autism Diagnostic Interview-Revised. *Journal of Autism and Developmental Disorders, 33,* 427-433

Constantino, J. N. & Gruber, C. P. (2005) *Social Responsiveness Scale (SRS)*. Los Angeles, LA: Western Psychological Services.

Corsello, C., Hus, V., Pickles, A. *et al* (in press) Between a ROC and a hard place: Decision making and making decisions about using the SCQ. *Journal of Child Psychology and Psychiatry.*

Dunn, G. (2000) *Statistics in Psychiatry*. London: Arnold.

Eaves, L.C, Wingert, H. & Ho, H.H. (2006a) Screening for autism - Agreement with diagnosis. *Autism*, 10, 229-242

Eaves, L.C, Wingert, H.D, Ho, H.H. *et al* (2006b) Screening for autism spectrum disorders with the social communication questionnaire. *Journal of Developmental and Behavioral Pediatrics*, 27, S95-S103

Geurts, H.M., Verté, S., Oosterlaan, J. *et al* (2004) Child the Children's Communication Checklist differentiate between children with autism, children with ADHD and normal controls? *Journal of Child Psychology and Psychiatry, 45*, 1437-1453.

Goodman, R. (1997) The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*, 581-586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1337-1345.

Hanley, J. A. & McNeil, B. J. (1982) The meaning and use of the area under a receiver operating characteristic curve. *Radiology*, 143, 29-36

Lee, L. C., David, A. B., Rusyniak, J. *et al* (2007) Performance of the Social Communication Questionnaire in children receiving special education services. *Research in Autism Spectrum Disorders, 1*, 126-138.

Lord, C., Risi, S., Lambrecht L. *et al* (2000) The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205-223

Lord, C., Rutter, M. & LeCouteur, A. (1994) Autism Diagnostic Interview-Revised - A Revised Version of A Diagnostic Interview for Caregivers of Individuals with Possible Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders, 24*, 659-685

Meltzer, H., Gatward, R., Goodman, R. *et al* (2000) *Mental health of children and adolescents in Great Britain*. London: The Stationery Office

Raven, J. C., Court, J. H. & Raven, J. (1990) *Standard Progressive Matrices.* Oxford: Oxford University Press

Rutter, M., Bailey, A. & Lord, C. (2003) *Social Communication Questionnaire (SCQ)*. Los Angeles, LA: Western Psychological Services

Skuse, D. H., Mandy, W. P. L. *et al* (2005) Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *British Journal of Psychiatry, 187*, 568-572

Sparrow, S., Balla, D. & Cichetti, D. (1984) *Vineland Adaptive Behaviour Scales.* Circle Pines, Minnesota: American Guidance Services

Stata (2005) *Stata Statistical Software Release 9.0: Survey Data Manual* College Station, TX: Stata Corporation

Towbin, K. E., Pradella, A., Gorrindo, T. *et al* (2005) Autism spectrum traits in children with mood and anxiety disorders. *Journal of Child and Adolescent Psychopharmacology, 15*, 452-464

Wechsler D (1992) *Wechsler Intelligence Scale for Children (III- UK Edition).* London: The Psychological Corporation

World Health Organisation (1993) *Mental Disorders: A Glossary and Guide to their Classification in Accordance with the 10th Revision of the International Classification of Diseases: Research Diagnostic Criteria (ICD-10)*. Geneva: WHO

Table 1 Scores on the three screens by diagnostic group.

	Non-ASD Mean (SE) N=49	**Other ASD** Mean (SE) N=37	**Childhood autism** Mean (SE) N=33
SCQ	9.5 (1.1)	19.2 (1.1)	25.8 (0.5)
SRS	68.0 (6.2)	97.8 (10.2)	116.1 (4.6)
CCC	131.9 (2.7)	120.3 (1.8)	114.5 (1.8)

Table 2 Correlations between total scores on the three screens and scores on the ADI-R, ADOS-G, ICD-10 symptom count, IQ, BPVS, Vineland ABC and parent and teacher SDQ.

	SCQ	**SRS**	**CCC**
ADI (N=117)	.83***	.59***	-.58***
ADOS (N=119)	.45***	.48***	-.36***
ICD (N=119)	.72***	.59***	-.55***
IQ (N=118)	.01	-.09	.20*
BPVS (N=118)	.12	.09	.06
VABS (N=103)	-.38***	-.44***	.43***
pSDQ (N=119)	.57***	.73***	-.67***
tSDQ (N=108)	.34***	.36***	-.47***

** p<.01; *** p<.001

Table 3 SCQ, SRS and CCC instrument properties (weighted values)

SCQ ASD cut-point >=15	**SRS ASD cut-point >=75**	**CCC PC cut-point =<132**
Whole sample N=119	*Whole sample N=119*	*Whole sample N=119*
AUC .90 (95% CIs .81 to .96)	AUC .77 (95% CIs .61 to .90)	AUC .79 (95% CIs .64 to .91)
Se .86 (95% CIs .65 to .96)	Se .78 (95% CIs .57 to .92)	Se .93 (95% CIs .87 to .97)
Sp .78 (95% CIs .60 to .93)	Sp .67 (95% CIs .46 to .84)	Sp .46 (95% CIs .28 to .68)
PPV .74 (95% CIs .56 to .92)	PPV .63 (95% CIs .46 to .82)	PPV .56 (95% CIs .41 to .75)
NPV .88 (95% CIs .72 to .97)	NPV .81 (95% CIs .61 to .94)	NPV .90 (95% CIs .81 to .96)
Low IQ N=44	*Low IQ N=44*	*Low IQ N=44*
AUC .92 (95% CIs .74 to .99)	AUC .67 (95% CIs .38 to .93)	AUC .72 (95% CIs .47 to .92)
Se .97 (95% CIs .88 to 1.00)	Se .78 (95% CIs .46 to 1.00)	Se .99 (95% CIs .96 to 1.00)
Sp .73 (95% CIs .42 to .99)	Sp .57 (95% CIs .29 to .84)	Sp .41 (95% CIs .14 to .70)
PPV .68 (95% CIs .38 to .98)	PPV .52 (95% CIs .29 to .78)	PPV .50 (95% CIs .27 to .75)
NPV .98 (95% CIs .89 to 1.00)	NPV .81 (95% CIs .44 to 1.00)	NPV .99 (95% CIs .91 to 1.0)
High IQ N=75	*High IQ N=75*	*High IQ N=75*
AUC .90 (95% CIs .77 to .97)	AUC .87 (95% CIs .73 to .95)	AUC .88 (95% CIs .73 to .97)
Se .77 (95% CIs .51 to .94)	Se .78 (95% CIs .61 to .91)	Se .88 (95% CIs .76 to .96)
Sp .85 (95% CIs .60 to .98)	Sp .80 (95% CIs .58 to .94)	Sp .53 (95% CIs .30 to .86)
PPV .83 (95% CIs .58 to .97)	PPV .78 (95% CIs .53 to .94)	PPV .63 (95% CIs .42 to .89)
NPV .80 (95% CIs .57 to .96)	NPV .80 (95% CIs .60 to .91)	NPV .83 (95% CIs .68 to .94)
Low SDQ N=33	*Low SDQ N=33*	*Low SDQ N=33*
AUC 1.0 (95% CIs .99 to 1.00)	AUC .77 (95% CIs .42 to .94)	AUC .86 (95% CIs .67 to .98)
Se .87 (95% CIs .54 to 1.00)	Se .57 (95% CIs .15 to .93)	Se .94 (95% CIs .70 to 1.00)
Sp .99 (95% CIs .96 to 1.00)	Sp .93 (95% CIs .76 to 1.00)	Sp .60 (95% CIs .34 to .90)
PPV .91 (95% CIs .61 to 1.00)	PPV .47 (95% CIs .00 to 1.00)	PPV .21 (95% CIs .05 to .57)
NPV .99 (95% CIs .95 to 1.00)	NPV .95 (95% CIs .87 to .99)	NPV .99 (95% CIs .95 to 1.0)
High SDQ N=77	*High SDQ N=77*	*High SDQ N=77*
AUC .83 (95% CIs .66 to .92)	AUC .67 (95% CIs .44 to .87)	AUC .66 (95% CIs .42 to .87)
Se .86 (95% CIs .66 to .97)	Se .80 (95% CIs .55 to .95)	Se .95 (95% CIs .89 to .99)
Sp .57 (95% CIs .29 to .87)	Sp .41 (95% CIs .15 to .71)	Sp .30 (95% CIs .09 to .66)
PPV .73 (95% CIs .52 to .91)	PPV .64 (95% CIs .43 to .82)	PPV .64 (95% CIs .45 to .83)
NPV .76 (95% CIs .37 to .95)	NPV .61 (95% CIs .24 to .93)	NPV .81 (95% CIs .52 to .95)

Key: Area Under the Curve (AUC); Sensitivity (Se); Specificity (Sp); Positive Predictive Value (PPV); Negative Predictive Value (NPV);

Figure 1 ROC curves for the SCQ, SRS and CCC