Journal of Child Psychology and Psychiatry 48:9 (2007), pp 932–940 doi:10.1111/j.1469-7610.2007.01762.x

Between a ROC and a hard place: decision making and making decisions about using the SCQ

Christina Corsello,[1] Vanessa Hus,[2] Andrew Pickles,[3] Susan Risi,[4] Edwin H. Cook, Jr.,[5] Bennett L. Leventhal,[5] and Catherine Lord[4]

[1]University of California, San Diego, USA; [2]Oxford University, UK; [3]University of Manchester, UK; [4]University of Michigan Autism and Communication Disorders Center, Ann Arbor, USA; [5]Department of Psychiatry, University of Illinois at Chicago, USA

Background: The Social Communication Questionnaire (SCQ), formerly the Autism Screening Questionnaire (ASQ), is based on a well-validated parent interview, the Autism Diagnostic Interview (ADI). It has shown promise as a screening measure for autism spectrum disorders (ASDs) in a research-referred older sample, though recent studies with younger children reported lower sensitivities when using the suggested cutoff of ≥15 to differentiate ASDs from children with nonspectrum disorders (NS). **Methods:** Diagnostic discrimination of the SCQ was evaluated alone and in combination with the ADOS (Autism Diagnostic Observation Schedule) in a clinical and research-referred sample of 590 children and adolescents (2 to 16 years), with best estimate consensus diagnoses of autism, pervasive developmental disorder, not otherwise specified (PDD-NOS) and non-ASD disorders. The SCQ was completed before the evaluation in most cases. Performance of the SCQ was also compared with the Autism Diagnostic Interview – Revised (ADI-R). **Results:** Absolute scores and sensitivity in the younger children and specificity for all groups were lower than reported in the original study. Using receiver operating curves (ROC) to examine the area under the curve (AUC), the SCQ was more similar to the ADI-R total score in differentiating ASD from NS disorders in the older (8–10, >11) than younger age groups (<5, 5–7). Lowering the cutoff score in the 2 younger groups improved sensitivity, with specificity remaining relatively low in all groups. Using the SCQ in combination with the ADOS resulted in improved specificity. Diagnostic discrimination was best using the ADI-R and ADOS in combination. **Conclusions:** Those interested in using the SCQ should consider adjusting cutoff scores according to age and purpose, and using it in combination with another measure. Sensitivity or specificity may be prioritized for research or screening depending on goals. **Keywords:** Autistic disorder, diagnosis, screening, Social Communication Questionnaire. **Abbreviations:** ADOS: Autism Diagnostic Observation Schedule; ADI-R: Autism Diagnostic Interview-Revised; ASD: autism spectrum disorders; PPD-NOS: pervasive developmental disorder, not otherwise specified; SCQ: Social Communication Questionnaire; AUC: area under the curve.

Over the past two decades, there has been significant progress in the development of reliable and valid standardized instruments for diagnosing autism and autism spectrum disorders (ASDs) (Lord, Rutter, DiLavore, & Risi, 1999; Rutter, LeCouteur, & Lord, 2003), which has led to more consistency in classification of ASDs and the ability to diagnose very young children with autism (Lord, 1995; Charman & Baird, 2002). This has resulted in an increasing need for reliable and valid screening instruments to determine who should receive a formal diagnostic assessment. Several screening instruments have been developed; most have limitations, including high false negative rates (Baird et al., 2000; Baron-Cohen, Allen, & Gillberg, 1992). The Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003), formerly the Autism Screening Questionnaire (ASQ; Berument, Rutter, Lord, Pickles, & Bailey, 1999), was developed based on the Autism Diagnostic Interview (ADI; LeCouteur et al., 1989), an earlier version of the Autism Diagnostic Interview – Revised (ADI-R; Rutter et al., 2003), a parent interview that demonstrates good validity and reliability. The SCQ is now widely available as a screening tool. However, there are few studies investigating the validity of the instrument, particularly with younger children, and with informants who are not already familiar with the traits and behaviors associated with autism.

The first published study on the SCQ was encouraging, reporting good sensitivity and specificity in a primarily British sample of 200 individuals (160 with ASDs and 40 language impaired, developmentally delayed, and conduct disorder controls), setting ≥ 15 as the cutoff to differentiate ASDs from children with non-spectrum diagnoses (NS) and ≥22 to differentiate children with autism from those without autism, using clinical diagnosis as the gold standard. However, the groups were older than those typically screened, with a mean age of 23 years in the autism sample. In addition, all the parents were

Conflict of interest statement: Authors S.R., C.L., C.C., and A.P. receive royalties for the SCQ, ADI-R, and/or ADOS; profits accrued from this study were donated to charity.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

participating in research studies on ASDs and had been interviewed using either the ADI or ADI-R and diagnosed with or without ASDs prior to receiving the SCQ (Berument et al., 1999). A more recent study with a younger, American sample was less encouraging, reporting a sensitivity of .74 and a specificity of .54 for differentiating ASD from NS (Eaves, Wingert, Ho, & Mickelson, 2006). The authors of this study suggested prorating scores rather than lowering the cutoff score to account for potential lower scores in younger nonverbal children, even though raw scores were not much different between the verbal and nonverbal groups, which resulted in improved sensitivity, but lower specificity.

Agreement between the SCQ and other instruments for diagnostic classification has varied depending on how diagnosis was defined. Recent studies have demonstrated that diagnostic accuracy is improved when clinical judgment is used in conjunction with standardized observational and parent measures (Lord et al., 2006; Risi et al., 2006). The most commonly used standardized parent measure in research studies is the ADI-R, which takes between 90 minutes and 3 hours to administer and may not be practical in many clinical and research settings. Other standardized parent measures are questionnaires, which are quicker and less expensive alternatives to a parent interview, such as the Gilliam Autism Rating Scales (GARS-II: Gilliam, 2006), the Social Responsiveness Scale (SRS: Constantino & Gruber, 2005), and the Autism Behavior Checklist (ABC: Krug, Arick, & Almond, 1979). The SCQ is also a questionnaire and, although designed as a screener, may be worth investigating as a contributor to diagnosis because it is based on a well-established and -studied diagnostic parent interview. Several investigators have been interested in using the SCQ in place of the ADI-R for research screening of children previously diagnosed in the community with an ASD. For this reason, they have compared the SCQ to the ADI-R. Bishop and Norbury (2002) found good categorical agreement between the SCQ and ADI-R, using a chi square analysis. Another recent British study with an older sample of individuals with Cohen syndrome reported good agreement between SCQ and ADI-R total scores ($r = .85$; Howlin & Karpf, 2004). Comparing the SCQ to an observational measure, the ADOS (Autism Diagnostic Observation Schedule; Lord, Rutter, DiLavore, & Risi, 1999) resulted in good agreement in one study (Howlin & Karpf, 2004) but not in another (Bishop & Norbury, 2002). When comparing the SCQ to the ADI-R and ADOS in combination, sensitivity for the SCQ remained high, but specificity dropped considerably (Howlin & Karpf, 2004).

The present multi-site study was designed to investigate how well the SCQ functions as a clinical screening instrument in a larger, younger American sample of children with ASD or non-spectrum disorders. All parents completed the SCQ prior to being interviewed with the ADI-R. Because of interest from clinicians and researchers in using the shorter, parent-completed SCQ in lieu of the longer, investigator-based interview format ADI-R, the SCQ is compared to the ADI-R with particular emphasis on its discriminative validity both alone and in combination with the ADOS.

Method

Subjects

The sample consisted of 590 children between 2 and 16 years who were consecutive referrals to two university-based clinics specializing in children with possible ASDs and/or were participants in research within the autism centers. Forty-three of the initial 633 participants were excluded because scores from missing items could have changed their SCQ classification. Parents signed an IRB-approved informed consent prior to participation. Consensus Best Estimate DSM IV (APA, 1994) diagnoses were made by two examiners (e.g., a child psychiatrist, clinical psychologist) who saw the child for 1–3 one- to three-hour sessions and had access to all assessment results (see below), as well as unstructured telephone teacher interviews. As shown in Table 1, nonverbal and verbal IQs ranged from profound mental retardation to superior intelligence.

The entire sample included 282 children with autism, 157 children with other ASDs, which included pervasive developmental disorder – not otherwise specified (PDD-NOS) and Asperger's disorder (AD), and 151 non-spectrum (NS) children, each of whom had been referred for possible autism or had been part of a control group in a research project and had at least one non-spectrum diagnosis (communication disorder = 36, ADHD = 30,

Table 1 Demographics and psychometrics by clinical diagnosis

	Clinical diagnosis		
	Autism ($n = 282$)	PDD-NOS ($n = 157$)	Non-spectrum ($n = 151$)
Age (months)*			
M	84.34	96.09	93.09
SD	43.58	47.92	45.53
Range	25–184	24–188	25–192
Nonverbal IQ**			
M	68.92	91.26	78.44
SD	27.58	25.28	26.75
Range	2–150	15–153	13–126
Verbal IQ**			
M	52.02	90.01	78.51
SD	30.10	27.73	27.50
Range	3–129	15–153	14–139
ADI-R Total**			
M	37.98	26.15	17.87
SD	9.32	11.61	10.62
Range	5–54	1–50	1–46
ADOS Total**			
M	17.19	9.62	5.49
SD	3.75	4.03	3.71
Range	1–24	1–21	0–20

*$p < .05$; **$p < .001$.

mental retardation = 26, Down sydrome = 18, fetal alcohol syndrome = 18, mood/anxiety disorder = 12, other developmental/psychiatric disorders = 11). Because the DSM-IV priority rule of diagnosing autism first was followed, only 3 children received a diagnosis of AD. Consequently, they were merged with PDD-NOS. The majority of the children were male ($n = 462$) and Caucasian ($n = 495$), with significantly fewer African Americans ($n = 43$), and other ethnicities ($n = 48$, 4 missing). The majority of the parents of the sample had some college or a higher level of education ($n = 451$, 38 missing).

Procedure

Parents completed the SCQ for their child prior to the diagnostic assessment as part of a pre-evaluation packet. Clinicians who made the best estimate diagnosis were not aware of SCQ scores but were aware of ADI-R and ADOS scores and classifications.

Measures

The SCQ is a 40-item, parent-completed, screening questionnaire, based on the initial mandatory probes from the original ADI (LeCouteur et al., 1989) covering the areas of communication, reciprocal social interactions, and restricted and repetitive behaviors and interests (Rutter, Bailey et al., 2003). Each item is checked as 'yes' or 'no', and assigned a point rating of '1' (presence of abnormal behavior) or '0' (absence of abnormal behavior). The first item is not included in the scoring, as it indicates if the child has sufficient verbal skills for language items to be scored. If the child is not scored as verbal, the six language items are skipped. The points are summed and result in a total possible score of 0–33 for nonverbal children and 0–39 for verbal children. Totals are compared to a cut off of ≥15 for ASD and ≥22 for autism. There are two different versions of the SCQ: 1) a 'current' version designed for children under the age of 5 years and 2) a 'lifetime' version designed for children 5 years of age or older, with all questions based on lifetime or past behavior.

A trained examiner who had achieved research reliability (see Lord et al., 1994) administered the ADI-R (Rutter, LeCouteur et al., 2003) and ADOS (Lord et al., 1999) to all cases. Both measures contain questions about the three areas of behaviors associated with an ASD. Most items are scored on a '0–3' scale, with '0' indicating the absence of the specific abnormality and '3' indicating an abnormality that interferes with daily life. A code of '8' indicates an item is not applicable. Both measures have algorithms, with separate cutoff scores for each domain. On the ADI-R, fewer items are administered to nonverbal children and children under the age of 4, resulting in lower total algorithm scores for these children. Most ADI-R algorithm items are coded based on current or past behavior for children under 5 years of age and past behavior for children over the age of 5 years.

The ADOS (Lord et al., 1999) is a standardized observational scale consisting of about 10 tasks and 30 codes. It is organized into 4 separate modules, based on the age and expressive language level of the child, ranging from pre-verbal toddlers to verbally fluent adults. A child meets criteria for a classification of autism if the scores in the social and communication domains and the total on the algorithm meet or exceed cutoff scores. The total possible score for modules 1 and 2 is 24 and for Modules 3 and 4 is 22. Total scores were prorated to account for this difference in Table 1. In the present study, all examiners had achieved inter-rater reliability meeting standard criteria for each instrument prior to participating. More than one-third of administrations were double coded and reliability maintained at greater than 80% exact agreement on the ADOS and 90% for the ADI-R.

A variety of cognitive and developmental measures were administered depending on the age and developmental level of the child or adolescent in order to ensure that each individual reached a basal and had a ceiling on nonverbal and verbal tests: Differential Ability Scales (Elliott, 1990; $n = 289$), Mullen Scales of Early Learning (Mullen, 1995; $n = 192$), Wechsler Scales of Intelligence (Wechsler, 1991; $n = 29$), or Other = 75. Each test resulted in a standard score, with a mean of 100 and a standard deviation of 15. If a child was severely delayed and tested using a measure that did not have normative scores for his/her chronological age, ratio IQ scores were calculated by dividing the child's age equivalent by his/her chronological age and multiplying by 100.

Design

In order to evaluate the validity of the SCQ in this sample, we first attempted to replicate the results from the original study (Berument et al., 1999). We assessed discriminative validity by examining the area under the receiver operating curve (AUC), which is a plot of true positive versus false positive results to assess the ability of a test to detect the presence of a diagnosis. We then ran a confirmatory factor analysis and examined correlations between the SCQ and ADI-R totals and domain scores. Next, we ran a linear regression to determine which variables affected SCQ scores. Multivariate and univariate analysis of variance were used to assess the relationships between age, ADI-R scores, and SCQ scores. Finally, the discriminative validity of the SCQ and ADI-R were compared alone and in combination with the ADOS by examining the sensitivity (the ability to accurately classify children with an autism or ASD diagnosis: TP/ (TP + FN)), specificity (the ability to accurately exclude children without an autism or ASD diagnosis: TN/ (TN + FP)), and predictive power (how likely a child classified with ASD or autism on the measure is to have that diagnosis) of each measure and each combination of measures. Comparisons were made between children with autism (AUT) and those without autism, including PDD-NOS and NS disorders (Not AUT), as well as between children with ASD (including autism and PDD-NOS) and NS (non-spectrum disorders). Throughout the analyses, the groups will be referred to in these ways.

Results

Background results

As shown in Table 1, there were significant differences in verbal IQ, $F(2,584) = 98.01$, $p < .001$, and nonverbal IQ scores, $F(2, 572) = 34.21$, $p < .001$,

among the three diagnostic groups, with significantly lower verbal and nonverbal IQ scores in the AUT group than the PDD-NOS and the NS groups, and significantly lower nonverbal and verbal IQ scores in the NS group than the PDD-NOS group. There was a significant difference in age, F (2, 589) = 3.96, $p < .05$, with the AUT group significantly younger than the PDD-NOS group. There was no significant difference between the groups in ethnicity, χ^2 (4, N = 586) = .44. There was a significant difference in gender, χ^2 (2, N = 589) = 16.03, $p < .001$, with more males than females in the AUT and PDD-NOS groups than the NS group. There was also a significant difference in maternal education, χ^2 (2, N = 552) = 4.29, $p < .05$, with fewer mothers of the NS group achieving a college degree or higher level of education than the other two groups.

Replication of previous findings

The discriminative validity of the SCQ for differentiating participants was assessed using the area under the receiver operating curve (AUC). In this sample, the AUC was .77 applying the suggested cutoff score of ≥15 for a classification of an ASD versus NS, with a sensitivity of .71 and specificity of .71, both of which were lower than reported in the original study. Differentiating AUT versus NS, the AUC was .81, and again applying the cutoff of ≥15, resulted in a sensitivity of .78 and a specificity of .71. When differentiating AUT from Not AUT the AUC was .74, and using the cutoff of ≥22, resulted in low sensitivity (.45) and relatively good specificity (.84).

As in the original study, results of independent group *t*-tests indicated that SCQ scores were significantly higher in the AUT and ASD groups than the NS group (AUT > NS, ASD > NS) for all comparisons of diagnostic and IQ groupings at $p < .001$ (see Table 2). Table 2 has been organized based on Tables 5 and 6 of the original study (Berument et al., 1999, p. 448) to allow for ease of comparison.

As previously reported (Berument et al., 1999), correlations were strong and significant between the SCQ and the ADI-R total scores ($r = .73$, $p < .001$), and the SCQ total score and the ADI-R social ($r = .70$, $p < .001$) and nonverbal communication ($r = .63$, $p < .001$) domains, and moderate and significant between the SCQ total score and the ADI-R restricted and repetitive ($r = .47$, $p < .001$) domain. Correlations between the ADI-R and SCQ social ($r = .69$, $p < .001$), nonverbal communication ($r = .65$, $p < .001$) and restricted and repetitive behavior domains ($r = .56$, $p < .001$) were also strong and significant.

Potential contributors to discriminative validity

The next step was determining possible reasons for lower SCQ scores and lower sensitivity and specificity in this younger American sample than in the original study. A hierarchical linear regression was run with SCQ score as the dependent variable. Preliminary analyses included the independent variables of ADI-R total score, which was entered first in order to control for differences in severity of autism, followed by age, verbal IQ, nonverbal IQ, and the categorical variables of SCQ version (current, lifetime), ethnicity (white, African American, other), gender (male, female), maternal education (college or higher, high school or lower), and birth order (first born or not). Preliminary analyses indicated that ADI-R total score and chronological age contributed

Table 2 Diagnostic discrimination using the cut-off of ≥15 for the total group and by IQ bands

	n	SCQ *M (SD)*	R.O.C.	Sensitivity	Specificity
Diagnostic groupings					
AUT*	281	20.26 (6.82)	AUT vs. NS = .81	.78	.71
ASD**	438	18.66 (7.14)	ASD vs. NS = .77	.71	.71
NS***	151	11.56 (6.81)			
Full scale IQ groupings					
IQ ≥ 70					
AUT	101	19.27 (7.01)	AUT vs. NS = .78	.74	.73
ASD	223	17.06 (7.19)	ASD vs. NS = .71	.62	.73
NS	95	11.66 (7.01)			
IQ = 50–69					
AUT	68	19.66 (6.97)	AUT vs. NS = .83	.74	.68
ASD	87	19.21 (7.09)	ASD vs. NS = .81	.74	.68
NS	28	10.50 (6.63)			
IQ = 30–49					
AUT	67	20.91 (6.58)	AUT vs. NS = .89	.82	.83
ASD	76	20.59 (6.50)	ASD vs. NS = .89	.81	.83
NS	18	9.89 (5.26)			
IQ < 30					
AUT	38	22.71 (5.90)		.92	.33
ASD	40	22.70 (5.92)		.93	.33
NS	6	17.33 (5.05)			

*AUT = autistic disorder; **ASD = autism spectrum disorder (including autism); ***NS = non-spectrum.

Table 3 Diagnostic discrimination by age with suggested cutoff of ≥15

	n	SCQ *M (SD)*	*t*	ROC	Sensitivity	Specificity
<5 years						
ASD	157	17.27 (6.40)	6.16***	.77	.68	.74
NS	43	10.40 (6.82)				
5–7 years						
ASD	109	17.04 (6.72)	3.96***	.70	.63	.67
NS	42	12.29 (6.28)				
8–10 years						
ASD	42	19.07 (6.85)	4.77***	.81	.71	.82
NS	22	11.05 (5.40)				
>11 years						
ASD	130	21.50 (7.61)	6.89***	.80	.80	.66
NS	44	12.30 (7.83)				

*$p < .05$; **$p < .01$; ***$p < .001$.

significantly to the SCQ total score, together accounting for a significant amount of the variance, $R^2 = .73$, $F(9, 448) = 55.90$, $p < .001$. There was no effect of ethnicity, gender, maternal education, version, or birth order. Because ADI-R scores are related to IQ, making it possible that co-varying ADI-R scores precluded IQ effects on the SCQ, the regression was re-run, replacing ADI-R score with the categorical variable of diagnosis (ASD, NS). In this case, age remained a significant predictor, and verbal IQ, diagnosis, and SCQ version became significant predictors.

In order to determine how SCQ scores differed according to the children's age, 4 different age groups were created (<5, 5–7, 8–10, > 11), taking into account the ages at which 'current' or 'most abnormal 4–5' information is evaluated. *T*-tests were run comparing ASD and NS in each age group, as shown in Table 3. SCQ scores differed significantly by diagnosis within each age group. In addition, sensitivity increased with age. Not surprisingly, the >11 year age children, the group most similar in age to the sample in Berument et al. (1999), had the highest scores.

Verbal level (verbal, nonverbal) and age group (<8, ≥8) were also examined more closely to be sure that verbal level (which affected possible total number of SCQ items) did not account for lower SCQ scores in younger groups. Of the entire sample, 74% was verbal, with a higher percentage of verbal children in the older (87%) than younger group (66%). A univariate analysis of variance was run, with SCQ score as the dependent variable and diagnostic group (ASD, NS), language level (verbal, nonverbal), and age group (<8 years, ≥8 years) as independent variables, co-varying nonverbal IQ scores. As described earlier, there was a main effect for diagnostic group, $F(1, 569) = 51.44$, $p < .001$ and age group, $F(1, 569) = 8.28$, $p < .01$, with the older group scoring higher ($M = 18.40$, $SD = 8.40$) than the younger group ($M = 15.76$, $SD = 6.99$). There was not a significant difference in scores between the nonverbal ($M = 17.66$, $SD = 7.05$) and verbal groups ($M = 16.56$, $SD = 7.91$). Scores were actually slightly higher in the nonverbal group, even though their verbal items had been scored as zeroes.

Comparison of the SCQ and the ADI-R alone and in combination with the ADOS

The discriminative validity of the ADI-R and the SCQ was compared for each measure alone and in combination with the ADOS. On the SCQ, the cutoff of ≥15 was used because the cutoff of ≥22 for classifying autism resulted in very low sensitivity. Comparisons were made first using a classification of AUT versus Not AUT because the standard use of the ADI-R results in a classification of autism or not and then using a classification of ASD versus NS, because the SCQ is typically used to screen children for more broadly defined ASD. In the latter comparison, the ADI-R classification of ASD was defined in two ways: a) meeting or exceeding the cutoff score in either communication or reciprocal social interactions and falling within two points of the cutoff on the other domain, or b) within one point on both the communication and social domains (an algorithm adopted by several American research networks; see Risi et al., 2006).

When classifying AUT versus Not AUT and ASD versus NS, not surprisingly, the combination of the ADI-R and the ADOS resulted in the best balance of sensitivity and specificity, which is desirable when using the measures diagnostically (see Table 4). Of the 5 children inaccurately classified with autism on both measures, 3 were under 5 years of age and had severe mental retardation. Of the 10 children inaccurately classified as NS, all but 1 received the classification of autism on the ADOS but not on the ADI-R, suggesting that clinicians tended to rely on their observations when the measures disagreed with one another. No children diagnosed with autism received a NS classification on both measures. When the cutoff of ≥15 was used, the SCQ resulted in lower sensitivity than the ADI-R, which is not ideal for a screening measure when including as many children who may have an ASD as possible is the goal.

Table 4 Classification of autism or not autism and ASD or NS with the ADI-R and SCQ Alone and in combination with the ADOS

	n	Sensitivity	Specificity	PPV[1]	NPV[2]
AUT vs. Not AUT					
ADI-R	590	.90	.58	.66	.86
ADOS	571	.94	.76	.78	.93
SCQ ≥ 15	590	.78	.57	.62	.74
SCQ ≥ 12	590	.88	.43	.58	.80
ADI-R & ADOS	571	.85	.87	.86	.86
SCQ ≥ 15 & ADOS	571	.73	.85	.82	.78
SCQ ≥ 12 & ADOS	571	.83	.80	.79	.83
ASD vs. Not					
ADI-R	590	.90	.54	.85	.64
ADOS	571	.90	.77	.92	.72
SCQ ≥ 15	590	.71	.71	.88	.45
SCQ ≥ 12	590	.82	.56	.84	.51
ADI-R& ADOS	571	.83	.86	.95	.64
SCQ ≥ 15 & ADOS	571	.66	.92	.96	.48
SCQ ≥ 12 & ADOS	571	.76	.86	.94	.55

[1]Positive predictive value.
[2] Negative predictive value.

Specificity for both the ADI-R and the SCQ was relatively low. When the ADOS was used in combination with the SCQ ≥ 15, specificity increased considerably.

Using the less stringent SCQ cutoff of ≥12 improved sensitivity, but not surprisingly, lowered specificity. Specificity was lowest when differentiating AUT and Not AUT, with slightly more than half of false positives (63%) accounted for by PDD-NOS cases. Using the SCQ ≥ 12 in combination with the ADOS resulted in sensitivity comparable to the ADI-R and ADOS combination in discriminating AUT from Not AUT, but was somewhat lower than the ADI-R and ADOS combination in discriminating ASD from NS (see Table 4). Positive and negative predictive values are reported, but must be considered in light of the ASD-skewed distribution of this clinic-referred sample.

Given the effects of age, the sensitivity and specificity of the ADI-R and SCQ were compared for different age groups. The SCQ, as shown in Table 5, consistently accounts for less AUC than the ADI-R, when using the ADI-R total score, with relatively little variation across age. It was not possible to identify a single cutoff on the SCQ that worked equally well across age groups. The measure achieved reasonable sensitivity but relatively low specificity for AUT versus Not AUT comparisons at each age. However, because the SCQ is intended as a screener for ASD, the ASD versus NS comparisons may be of greater interest

Examining the AUC for each of the four age groups in differentiating ASD from NS for both the ADI-R and SCQ suggested that the SCQ was most effective in the 8–10 and >11 years age groups (see Table 5), with performance similar to that of the ADI-R. Sensitivity was low in the younger two age groups when using the suggested cutoff score of ≥15. Specificity was quite low in all groups with the exception of the 8–10-year-old group. In order to achieve sensitivity of 80%, cutoff scores would need to be lowered in

Table 5 Classification of AUT versus Not AUT and ASD versus NS for the ADI-R and SCQ with modified cutoffs for age groups

	Total Group ($N = 590$)	<5 years (201)	5–7 years (151)	8–10 years (64)	11+ years (174)
AUT vs. Not AUT					
ROC-AUC: SCQ	.74	.71	.76	.75	.80
ADI-R	.85	.81	.88	.81	.90
Autism cut (≥22)					
Sensitivity	.45	.32	.36	.44	.71
Specificity	.84	.89	.90	.85	.73
ASD cut (≥15)					
Sensitivity	.78	.72	.76	.80	.88
Specificity	.57	.59	.62	.64	.47
SCQ 80% sensitivity					
Cutoff number	14	13	14	15	19
Specificity	.52	.51	.55	.64	.65
SCQ 80% specificity					
Cutoff number	21	20	18	20	24
Sensitivity	.48	.43	.57	.68	.60
ASD vs. NS					
ROC-AUC: SCQ	.77	.77	.70	.81	.80
ADI-R	.84	.85	.80	.83	.85
ASD cut (≥15)					
Sensitivity	.71	.68	.63	.71	.80
Specificity	.71	.74	.67	.82	.66
SCQ 80% sensitivity					
Cutoff number	12	11	12	14	15
Specificity	.56	.60	.50	.77	.66
SCQ 80% specificity					
Cutoff number	18	17	18	15	19
Sensitivity	.56	.58	.42	.71	.68

order to distinguish ASD from NS for children under the age of 8 years. Specificity remained relatively low in all groups.

Discussion

Several studies, including the present investigation, have found good agreement between the ADI-R and the SCQ, based on correlations between the measures. However, when comparing how well the measures classify individuals to clinical diagnosis, the ADI-R captured more children with autism or an ASD than the SCQ (≥15). In theory, a screener should be more inclusive than a diagnostic measure. Both instruments, used in isolation, included a relatively large number of children without autism or an ASD. Combining one of them with a standardized diagnostic observation resulted in the best diagnostic discrimination. The combination of the ADOS and ADI-R resulted in the most accurate diagnostic discrimination between AUT and Not AUT and between ASD and NS. Using the SCQ with a lowered cutoff of ≥12 in combination with the ADOS resulted in comparable *sensitivity* (ability to capture children with an ASD on the measure) to the combination of the ADI-R and the ADOS when differentiating AUT from Not AUT and comparable *specificity* (ability to accurately exclude NS children on the measure) to the ADI-R and ADOS when differentiating ASD from NS. It was not possible to improve both sensitivity and specificity of the SCQ and ADOS combination to a level comparable to that of the ADI-R and ADOS used together. Thus, clinicians and researchers wishing to take advantage of the ease of using the SCQ will need to determine their priority, capturing as many children with ASD as possible or excluding NS children.

Consistent with previous studies with younger American samples, younger children scored lower than older children on the SCQ, with the SCQ (≥15) missing a large number of younger children with ASDs. In the present sample, SCQ scores were higher for children beginning at about age 8. However, it is important to remember these are cross-sectional data and do not necessarily reflect increases in the scores of individuals. Biases resulting from age of referral or means of recruitment (research versus clinical) may have influenced these scores. When examining the AUC for total algorithm scores on the ADI-R and total SCQ scores in Table 5, the measures worked most similarly for the older two age groups when discriminating ASD from NS. Lowering the SCQ cutoff score to ≥11 or 12 resulted in better sensitivity for children <8 years of age, though at the cost of more NS children falsely classified as ASD. This result did not appear to be the effect of IQ. Although absolute SCQ scores were 3–5 points lower in this sample compared to the original British sample, IQ bands followed the same trends in both studies. In addition, no interactions emerged between age and IQ or age and verbal status. In fact, the nonverbal children scored slightly higher on the SCQ than the verbal children overall, even though they had 'missing data' on 6 verbal items. Consequently, lowering the cutoff score would be a more effective strategy than prorating scores in order to account for the skipped items for nonverbal children. Age differences in samples could partially explain the lower sensitivities in the US data compared to those of the British samples, as the participants were older in two of the British studies (Berument et al., 1999; Howlin & Karpf, 2004).

All studies have reported good agreement between the ADI-R and SCQ based on correlations. However, if a child has a high score on both measures, but one instrument's algorithm results in a classification of an ASD and the other does not, the high correlation is of little practical value when deciding how to characterize the child for research or clinical purposes. For this reason, it is important to report resulting classifications. Comparing a high SCQ score to a high score on *either* the ADI-R or ADOS results in the most agreements concerning classification of ASD, but may not be an accurate reflection of how well the SCQ alone works diagnostically, because of the low specificity of both the SCQ and ADI-R when used in isolation and of categorical agreement made on the basis of either the ADI-R or the ADOS (Risi et al., 2006). Comparing the SCQ to clinical judgment based on both the ADOS and ADI-R will result in the most conservative, and likely, the most accurate assessment of its diagnostic validity (Lord et al., 2006).

Limitations

The present study had a number of limitations that affect interpretation of the results. Data were not from a population sample so inferences about use of the SCQ as a general screening tool are not possible. Furthermore, the NS samples were idiosyncratically recruited groups of children who were either referrals to ASD clinics who were judged not to have an ASD or comparison groups in research projects with psychiatric or developmental disorders, other than ASD. The severity of these comparison groups' social and communication deficits as a result of having been suspected of having an ASD, even though not falling within ASD, likely contributed to findings of relatively poor specificity of all of the instruments, particularly for contrasts between ASD and NS. This select sample, which may be representative of a specialty clinic or research samples in autism research projects, may not be representative of children referred to general psychiatry or developmental clinics. The relatively small samples of children with NS (i.e., 18–151 per age cell) also limited comparisons within the age groups. Though one of the purposes of the paper was to address the use of the SCQ with younger participants, the sample size for

children <5 (180 ASD, 59 NS) was insufficient for more detailed study of age effects during the preschool period. Only 57 children were below the age of 3.

Because of the confounding of version of the SCQ with age (all children <5 received current) it was not possible to assess the effects of version independent of age, nor was it possible to assess the effects of verbal level independent of age and nonverbal IQ. More focused samples recruited specifically with such comparisons in mind will need to be studied to address these questions.

Summary and conclusions

It is clear from other research (Bishop & Norbury, 2002; Risi et al., 2006) that the various autism diagnostic or descriptive measures each measure something slightly different. Selection of instruments and cutoffs, therefore, needs to be made taking into consideration the characteristics of the children and their families. Thus, if an investigator conducts a study in which a relatively small, clearly diagnosed sample of young children with narrowly defined autism is the priority, he/she might use the SCQ with a cutoff of ≥12 to first screen individuals who have a community diagnosis of autism or a related disorder, then administer an ADOS to screened positive cases, which would increase specificity while maintaining sensitivity and then, only for ADOS positive cases, administer an ADI-R. On the other hand, if a researcher's priority was to identify individuals with broadly defined ASD as efficiently as possible, he/she might recruit cases with community diagnoses of any ASD or potentially overlapping disorder, administer an ADOS and an SCQ, and use a cutoff of ≥15 and ADOS ASD thresholds as the final criteria. A health management organization (HMO), in order to determine appropriate referrals to a specialized evaluation clinic for children with ASD, might screen cases where there is parental or professional concern, using the 80% sensitivity cutoffs for different ages on Table 5, followed by an ADOS. An ADI-R might be used only when more information about diagnosis was needed for clinical purposes. Awareness of the properties of the instruments gives clinicians and researchers a wider variety of options, which can be helpful as long as the consequent limitations are considered as well.

Acknowledgements

We gratefully acknowledge the families who participated in these research projects and the staff at the TEACCH centers in North Carolina, the Department of Psychiatry at the University of Chicago, and the University of Michigan Autism and Communication Disorders Center. This work was supported by grants NIMH R01 MH066496 and R01 MH46865 to Dr Lord, and was carried out as part of the NICHD/NIDCD Collaborative Programs for Excellence in Autism (CPEA).

Correspondence to

Christina M. Corsello, Child & Adolescent Services Research Center (CASRC), 3020 Children's Way MC 5033, San Diego, CA 92123, USA; Tel: (858) 966-7703 extension: 3638; Fax: (858) 966-7704; Email: ccorsello@ucsd.edu

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th edn). Washington, DC: American Psychiatric Association.

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 694–702.

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry*, *161*, 839–843.

Berument, S.K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, *175*, 444–451.

Bishop, D.V., & Norbury, C.F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: A study using standardized diagnostic instruments. *Journal of Child Psychology and Psychiatry*, *43*, 917–929.

Charman, T., & Baird, G. (2002). Practitioner review: Diagnosis of autism spectrum disorder in 2- and 3-year-old children. *Journal of Child Psychology and Psychiatry*, *43*, 289–305.

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

Eaves, L.C., Wingert, H.D., Ho, H.H., & Mickelson, E.C. (2006). Screening for autism spectrum disorders with the Social Communication Questionnaire. *Journal of Developmental and Behavioral Pediatrics*, *27*, 95–103.

Elliott, C. (1990). *Differential Ability Scales*. Austin, TX: Harcourt Assessments, Inc.

Gilliam, J. (2006). *Gilliam Autism Rating Scale*, 2nd Edition Manual. Minneapolis, MN: Pearson Assessments.

Howlin, P., & Karpf, J. (2004). Using the Social Communication Questionnaire to identify 'autistic spectrum' disorders associated with other genetic conditions – findings from a study of individuals with Cohen syndrome. *Autism*, *8*, 175–182.

Krug, D.A., Arick, J.R., & Almond, P.G. (1979). Autism screening instrument for educational planning background and development. In J. Gilliam (Ed.), *Autism:*

Diagnosis, instruction, management and research. Austin, TX: University of Texas at Austin Press.

LeCouteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holdgrafer, M., et al. (1989). Autism Diagnostic Interview: A standardized investigator-based instrument. *Journal of Autism and Developmental Disorders*, *19*, 363–387.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, *36*, 1365–1382.

Lord, C., Risi, S., Dilavore, P., Shulman, C., Thurm, A., & Pickles, A. (2006). Autism from 2 to 9 years of age. *Archives of General Psychiatry*, *63*, 694–701.

Lord, C., Rutter, M., DiLavore, P.C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule – WPS* (WPS edn). Los Angeles, CA: Western Psychological Services.

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 & Developmental Disorders*, *24*, 659–685.

Mullen, E. (1995). *Mullen Scales of Early Learning*. Minneapolis, MN: Pearson Assessments.

Risi, S., Lord, C., Corsello, C., Chrysler, C., Szatmari, P., Cook, E.H., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 1094–1103.

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

Rutter, M., LeCouteur, A., & Lord, C. (2003). *Autism Diagnostic Interview–Revised Manual*. Los Angeles, CA: Western Psychological Services.

Wechsler, D. (1991). *Wechsler Intelligence for Children – Third Edition*. San Antonio, TX: The Psychological Corporation.

Manuscript accepted 19 February 2007