

# Sensitivity and specificity of early screening for autism

Pål Surén, Alexandra Saasen-Havdahl, Michaeline Bresnahan, Deborah Hirtz, Mady Hornig, Catherine Lord, Ted Reichborn-Kjennerud, Synnve Schjølberg, Anne-Siri Øyen, Per Magnus, Ezra Susser, W. Ian Lipkin and Camilla Stoltenberg

## Background

Early identification and diagnosis is beneficial for children with autism spectrum disorder (ASD). Universal early screening is recommended by many experts, but disputed because evidence is limited, and sensitivity and specificity in general populations are largely unknown.

## Aims

To estimate the sensitivity and specificity of early population-based screening for ASDs.

## Method

The study was based on the Norwegian Mother and Child Cohort Study. The 36-month cohort questionnaire included the Social Communication Questionnaire (SCQ), a 40-item screening instrument for ASD.

## Results

A total of 58 520 mothers (58%) responded to the questionnaire. By the end of follow-up on 31 December 2015, 385 (0.7%) individuals with ASD had been identified among the responders' children. The distributions of SCQ scores in those with ASD and other children had large degrees of overlap. With the cut-off of 15 recommended in the SCQ manual, screening sensitivity was 20% (95% CI 16–24) for ASD overall. For children with ASD who had not developed phrase speech at 36 months, sensitivity was 46% (95% CI 35–57%), whereas it was 13% (95% CI 9–17) for children with ASD with phrase speech. Screening specificity was 99% (95% CI 99–99). With the currently recommended cut-off of 11, sensitivity increased to 42% for ASD overall (95% CI 37–47), 69%

(95% CI 58–79) for ASD without phrase speech and 34% (95% CI 29–40) for ASD with phrase speech. Specificity was then reduced to 89% (95% CI 89–90).

## Conclusions

Early ASD screening with a parent checklist had low sensitivity. It identified mainly individuals with ASD with significant developmental delay and captured very few children with ASD with cognitive skills in the normal range. Increasing sensitivity was not possible without severely compromising specificity.

## Declaration of interest

C.L. receives royalty for the Social Communication Questionnaire, which she has co-authored.

## Keywords

Autistic spectrum disorders; developmental disorders; epidemiology.

## Copyright and usage

© The Royal College of Psychiatrists 2019. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work.

Autism spectrum disorder (ASD) are neurodevelopmental disorders characterised by persistent impairment in reciprocal social communication and social interaction, and restricted, repetitive patterns of behaviour, interests and activities.<sup>1</sup> Symptoms usually present at an early age, typically during the second year of life.<sup>1</sup> Early identification and diagnosis is assumed to be beneficial, because it facilitates access to interventions that may improve the child's long-term developmental outcomes.<sup>2</sup> It may also be helpful in obtaining immediate assistance to children and their families.<sup>2</sup> However, universal ASD screening is controversial and the evidence supporting such screening is weak. The current guidelines from the American Academy of Pediatrics state that ASD screening should be included in general developmental check-ups at ages 18 and 24 months.<sup>2</sup> The British guidelines, on the other hand, do not recommend universal ASD screening.<sup>3</sup>

## Screening tools for ASD

A number of screening tools for ASD are available. For children under 30 months of age, the most widely used is the Modified Checklist for Autism in Toddlers, Revised with Follow-Up (M-CHAT-R/F).<sup>4</sup> The M-CHAT-R/F includes a standardised questionnaire and an additional follow-up interview for those who

score above a defined threshold on the questionnaire. In a population-based study of 15 612 children aged 16–30 months in the USA, the M-CHAT-R/F had a positive predictive value of 47.5% for ASD and 94.6% for all types of developmental concerns and delays.<sup>4</sup>

A recent review by the US Preventive Services Task Force of published studies of universal ASD screening in children below 36 months of age noted the lack of follow-up to identify false negatives.<sup>5</sup> Without this information, it has not been possible to estimate sensitivity and specificity of screening instruments, and therefore not possible to fully assess the effectiveness of ASD screening. This was also a limitation of the large study of the M-CHAT-R/F, which did not include complete follow-up of the screen-negative children.<sup>4</sup> Furthermore, it is not well-known whether screening performance is modified by other characteristics of the child and the family, and whether there are disadvantages in implementing universal screening for ASD.

Consequently, we do not know whether existing screening tools for ASD comply with the screening criteria recommended by the World Health Organization.<sup>6</sup> According to these criteria, health screening should not be implemented unless it has been demonstrated that the screening test is effective and that the benefits of screening outweigh the harm and the costs.

## Aims

In this study, we have estimated the sensitivity and specificity of the Social Communication Questionnaire (SCQ) for ASD screening at 36 months of age.<sup>7</sup> The SCQ is a 40-item questionnaire that was initially designed for ASD screening in individuals aged 4 years and older.<sup>7</sup> Our study was conducted in the Norwegian Mother and Child Cohort (MoBa).<sup>8</sup> The cohort participants have been followed for a number of years after screening to capture all individuals diagnosed with ASD.

## Method

### Study population

MoBa is a nationwide, prospective, population-based cohort that includes 114 552 children born from 1999 to 2009.<sup>8</sup> Mothers were recruited at the ultrasound examination provided free of charge to all pregnant women around gestational week 18. Of the women invited to participate, 41% consented. Fathers were also invited to participate if the mothers had consented first. The parents completed questionnaires during pregnancy, and have continued to receive questionnaires as the children grow older. Children with ASD diagnoses have been identified and ascertained through the Autism Birth Cohort (ABC) Study, a study of ASD nested within MoBa.<sup>9</sup>

The 36-month questionnaire was distributed to MoBa participants born in 2001 and later, a total of 100 364 children. Of these, there were 58 520 (58%) whose parents returned the questionnaire. ASD case ascertainment has been conducted until 31 December 2015. By that time, the age range of the children whose parents completed the 36-month questionnaire was 6.5–14.7 years (mean 10.2 years).

### The SCQ

The SCQ items are modified questions from the standardised diagnostic interview for ASD, the Autism Diagnostic Interview – Revised (ADI-R).<sup>10</sup> The items were selected to detect deficits in social interaction and communication, and repetitive and stereotyped behaviours.<sup>7</sup> The first SCQ item is used to determine if the child has phrase speech. If so, the other 39 items are scored. A total score of  $\geq 15$  is considered screen-positive for ASD. If the child does not have phrase speech, only the 33 non-verbal items are scored.

Previous studies of the SCQ have utilised clinic-based study samples where children have already been diagnosed with ASD or referred for evaluations of ASD. The manual cut-off level of 15 is based on the initial validation study, which included participants from 4 to 32 years of age.<sup>7</sup> A subsequent American study showed that the SCQ had lower sensitivity in children under 8 years compared with children who were 8 years or older.<sup>11</sup> To achieve a sensitivity of 80% in children under 8 years of age, the cut-off had to be lowered to  $\geq 12$  for children aged 5–7 years and  $\geq 11$  for children under 5 years. More recent studies have also supported the use of 11 as the cut-off level in preschool and elementary-school children.<sup>12,13</sup>

### ASD screening in the MoBa cohort

The SCQ was embedded in the questionnaire distributed to MoBa participants at age 36 months. When the MoBa questionnaire was designed, there were no available screening instruments that had been tested in 3-year-olds. The SCQ was chosen because it captures all common symptoms of autism in children, and because the initial validation study indicated that it performed well at age 4 years and

older.<sup>7</sup> The investigators who initiated the screening procedures in 2004 decided not to use the seven verbal SCQ items for screening, because they assumed that a large proportion of children with ASD would not have developed spoken language by 3 years of age. In order to ensure high sensitivity, it was decided to use a more lenient cut-off level than 15, and screen-positivity for ASD was defined by a score of  $\geq 12$  on the 33 non-verbal SCQ items. Screen-positive children were invited to participate in a 1-day diagnostic assessment. A randomly selected group of age-matched controls was also invited. Controls were selected every 2 weeks among MoBa participants who had passed the age of 37.5 months during the 2 preceding weeks.

Even though the seven verbal SCQ items were not used for screening by the ABC Study, the data are still available. In this article, we have included all the SCQ items in the analyses. We have also estimated the sensitivity and specificity of the SCQ at different cut-off levels:

- the cut-off recommended by the SCQ manual:  $\geq 15$  for the 39 scored items.
- the cut-off recommended by more recent studies:  $\geq 11$  for the 39 scored items.
- the cut-off used in the ABC Study:  $\geq 12$  for the 33 non-verbal items.

### Other methods of identifying potential individuals with ASD

In addition to screening, the ABC Study used several methods to detect individuals with ASD in the cohort. The following participants were considered to potentially have ASD and invited to the clinical assessments.

- Children referred from their physicians or parents, if there was reason to suspect ASD. Referrals were elicited through annual newsletters to MoBa participants and information on the website of the Norwegian Institute of Public Health.
- Children whose parents reported autism, Asperger syndrome or autistic traits in the MoBa questionnaires at ages 5 and 7 years.
- Children diagnosed with ASD by specialist health services. Data from all hospitals and out-patient clinics are reported to the Norwegian Patient Register (NPR), beginning in 2008.<sup>14</sup> Reporting is linked to the government reimbursement system. We conducted annual linkages between MoBa and the NPR from 2008 to 2015 to identify MoBa participants with ASD diagnoses. Until all MoBa children reached 36 months of age in 2012, children with ASD diagnoses in the NPR were invited to the clinical assessments. After that, the diagnoses have been confirmed through medical record reviews, as described below.

### ASD case ascertainment

The clinical assessment of the ABC Study included the ADI-R and the Autism Diagnostic Observation Schedule.<sup>15</sup> Diagnoses were based on the DSM-IV-TR.<sup>16</sup> The ASD case definition included codes 299.00 (autistic disorder), 299.80 (Asperger disorder) and 299.80 (pervasive developmental disorder not otherwise specified).

After 2012, case ascertainment has been conducted through medical record reviews at the clinics where the diagnoses were recorded. In Norwegian specialist health services, diagnoses are coded according to the ICD-10.<sup>17</sup> The ASD case definition of our reviews included codes F84.0 (childhood autism), F84.1 (atypical autism), F84.5 (Asperger syndrome), F84.8 (other pervasive developmental disorder) and F84.9 (pervasive developmental disorder, unspecified).

**Table 1** Likelihood ratios of individual Social Communication Questionnaire items<sup>a</sup>

Item	Domain	Item content	No ASD group (n = 58 135)	Children with ASD without phrase speech (n = 86)		Children with ASD with phrase speech (n = 294)	
			Positive, %	Positive, %	Likelihood ratio	Positive, %	Likelihood ratio
1	C	No phrase speech	0.9	100.0	113.6	0.0	0.0
2	C	No conversation to and from	2.9	90.6	31.4	18.6	6.4
3	C	Echolalia	61.7	25.3	0.4	69.2	1.1
4	C	Inappropriate questions	38.2	1.2	0.0	27.1	0.7
5	C	Pronoun reversal	37.6	18.1	0.5	45.5	1.2
6	C	Neologisms	29.3	19.0	0.7	31.6	1.1
7	R	Verbal rituals	35.1	32.1	0.9	46.2	1.3
8	R	Compulsions and rituals	34.3	23.5	0.7	40.1	1.2
9	S	Inappropriate facial expressions	0.3	10.6	33.9	3.1	9.8
10	S	Use of other's body to communicate	39.5	72.9	1.8	72.9	1.8
11	R	Unusual preoccupations	17.4	54.1	3.1	54.1	3.1
12	R	Repetitive use of objects	34.1	68.2	2.0	51.2	1.5
13	R	Circumscribed interests	21.6	47.1	2.2	38.8	1.8
14	R	Unusual sensory interests	8.2	27.4	3.3	27.4	3.3
15	R	Hand and finger mannerisms	3.8	27.1	7.1	9.6	2.5
16	R	Complex body mannerisms	76.1	47.7	0.6	69.2	0.9
17	—	Self-injury	7.3	25.6	3.5	14.6	2.0
18	—	Unusual attachment to objects	25.5	30.6	1.2	31.4	1.2
19	S	No particular friends/best friend	38.7	83.3	2.2	55.2	1.4
20	C	No social chat	3.1	78.8	25.1	12.8	4.1
21	C	No imitation	5.3	40.0	7.6	7.5	1.4
22	C	No pointing to express interest	3.5	42.4	12.0	6.2	1.8
23	C	No gestures to request objects	27.1	43.5	1.6	38.7	1.4
24	C	No nodding to indicate 'yes'	7.8	67.1	8.5	24.3	3.1
25	C	No head shaking to indicate 'no'	11.9	49.4	4.1	25.3	2.1
26	S	Usually no eye gaze	1.7	26.5	15.3	6.6	3.8
27	S	No social smiling	1.0	10.6	11.1	1.4	1.5
28	S	No showing and directing attention	1.1	20.0	18.0	2.4	2.2
29	S	No offering to share	2.7	39.5	14.9	9.8	3.7
30	S	No seeking to share enjoyment	0.9	19.8	23.3	3.8	4.4
31	S	No offering of comfort	2.8	42.4	15.1	10.7	3.8
32	S	Lack of normal social overtures	3.6	12.9	3.6	3.8	1.0
33	S	Lack of normal range of facial expressions	0.1	16.5	133.8	2.1	16.9
34	C	No imaginative social play	4.3	37.2	8.6	13.0	3.0
35	C	No imaginative play	4.7	61.6	13.1	20.5	4.4
36	S	No interest in other children	2.5	34.1	13.7	11.1	4.4
37	S	No positive response to other children	2.5	16.7	6.7	10.4	4.2
38	—	Usually no attention to voice	2.7	34.1	12.5	8.6	3.1
39	S	No imaginative play with peers	15.7	89.5	5.7	45.9	2.9
40	S	No group play	8.1	69.8	8.6	27.0	3.3

ASD, autism spectrum disorder; R, repetitive and restricted behaviour; C, communication; S, social interaction.

a. Includes all responders to the Norwegian Mother and Child Cohort 36-month questionnaire (n = 58 520), with 385 individuals with ASD: 86 without phrase speech at 36 months, 294 with phrase speech at 36 months and 5 with missing information about phrase speech.

### Other factors affecting sensitivity, specificity and screening participation

To examine whether sensitivity and specificity were affected by other types of developmental delay, we stratified the individuals with ASD by the presence or absence of phrase speech (reported on the first SCQ item) and cognitive delay (IQ below 70 at cognitive testing). We examined whether screening participation and screening performance were associated with sociodemographic characteristics of the families: parental education, parental age, maternal parity and maternal living status. We also examined whether these sociodemographic characteristics were associated with children's risk of ASD.

### Statistical analyses

First, we evaluated how well the individual SCQ items differentiated between individuals with and without ASD. These abilities were evaluated by likelihood ratios (LRs).<sup>18</sup> An LR is the ratio between the probability of a positive score in the presence of a condition (true positivity) and the probability of a positive score in the absence of the condition (false positivity). A high LR value indicates

a high precision of a screening item, whereas an LR value close to one indicates that a screening item does not discriminate between individuals with and without the condition. The ability of a test to distinguish between 'cases' and 'non-cases' is generally considered to be poor when LR values are lower than 2.<sup>19</sup>

We then determined sensitivity, specificity and positive and negative predictive values of the SCQ at the different cut-off levels described previously. We also explored the screening performance of the SCQ through receiver operating characteristic (ROC) curves. ROC curves were created for the SCQ total (manual) score and for each of the three SCQ domains, i.e. communication, social interaction and repetitive behaviours. In an ROC curve, sensitivity (the true-positive proportion) is plotted on the y-axis, whereas 1 – specificity (the false-positive proportion) is plotted on the x-axis.<sup>18</sup> The area under the ROC curve (AUC) is a measure of the overall ability to differentiate between cases and non-cases, and can be interpreted as the probability that a randomly chosen pair of one case and one non-case will be correctly categorised.<sup>18</sup> An AUC value of 0.9–1 is considered excellent, while 0.80–0.89 is good and 0.70–0.79 is fair.<sup>20</sup> Values below 0.70 are poor, and tests with values close to 0.50 are of no value.<sup>20</sup> Associations between

**Table 2** Sensitivity and specificity of the Social Communication Questionnaire (SCQ) at different cut-off levels (total  $n = 58\,520$ )<sup>a</sup>

	All children with ASD	Children with ASD without phrase speech	Children with ASD with phrase speech
<i>SCQ total (manual) score <math>\geq 15</math><sup>b</sup></i>			
Sensitivity (95% CI)	20 (16–24)	46 (35–57)	13 (9–17)
<i>n/N</i>	76/379	39/85	37/294
Specificity (95% CI)	99 (99–99)	99 (99–99)	99 (99–99)
<i>n/N</i>	56 977/57 738	56 977/57 738	56 977/57 738
Positive predictive value (95% CI)	9 (7–11)	5 (3–7)	5 (3–6)
<i>n/N</i>	76/837	39/800	37/798
Negative predictive value (95% CI)	99 (99–100)	100 (100–100)	100 (99–100)
<i>n/N</i>	56 977/57 280	56 977/57 023	56 977/57 234
<i>SCQ total (manual) score <math>\geq 11</math><sup>b</sup></i>			
Sensitivity (95% CI)	42 (37–47)	69 (58–79)	34 (29–40)
<i>n/N</i>	160/379	59/85	101/294
Specificity (95% CI)	89 (89–90)	89 (89–90)	89 (89–90)
<i>n/N</i>	51 637/57 738	51 637/57 738	51 637/57 738
Positive predictive value (95% CI)	3 (2–3)	1 (1–1)	2 (1–2)
<i>n/N</i>	160/6261	59/6160	101/6202
Negative predictive value (95% CI)	100 (100–100)	100 (100–100)	100 (100–100)
<i>n/N</i>	51 637/51 856	51 637/51 663	51 637/51 830
<i>SCQ-33 score <math>\geq 12</math><sup>c</sup></i>			
Sensitivity (95% CI)	25 (20–29)	64 (52–74)	13 (9–17)
<i>n/N</i>	93/379	54/85	37/294
Specificity (95% CI)	99 (99–99)	99 (99–99)	99 (99–99)
<i>n/N</i>	57 346/57 852	57 346/57 852	57 346/57 852
Positive predictive value (95% CI)	16 (13–19)	10 (7–12)	7 (5–9)
<i>n/N</i>	93/599	54/560	37/543
Negative predictive value (95% CI)	100 (99–100)	100 (100–100)	100 (99–100)
<i>n/N</i>	57 346/57 632	57 346/57 377	57 346/57 603

ASD, Autism spectrum disorder; SCQ-33, the 33 SCQ items that do not require phrase speech.

a. Includes all responders to the Norwegian Mother and Child Cohort 36-month questionnaire ( $n = 58\,520$  including 385 individuals with ASD).

b. The SCQ manual score was considered valid if there were  $\leq 19$  missing values on the SCQ items (available for 379 children with ASD).

c. The SCQ-33 score was considered valid if there were  $\leq 16$  missing values on the 33 non-verbal SCQ items (available for 379 children with ASD).

screening participation, sociodemographic characteristics and ASD diagnoses were assessed by Pearson's  $\chi^2$  tests.

## Ethics

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. MoBa is regulated under the Norwegian Health Registry Act. Written informed consent was obtained from mothers and fathers who participate. The consent covers linkages to health registries and reviews of medical records. The ABC Study has approval from the Regional Committee for Medical and Health Research Ethics for South-East Norway. Participation in the clinical assessments of the ABC Study was based on an additional written informed consent.

## Results

By the end of follow-up, a total of 385 individuals with ASD (0.7%) had been detected among the 58 520 children with questionnaire data at 36 months. Of those with ASD, only 56 were detected through screening and clinical assessments. The others were detected through referrals ( $n = 20$ ), reports of autism or Asperger syndrome in the 5- and 7-year questionnaires ( $n = 18$ ) and registry linkages ( $n = 291$ ).

There were 294 (76%) children with ASD who had developed phrase speech at 36 months, while 86 (22%) had not. The information about phrase speech was derived from item 1 of the SCQ. This information was missing for 5 children with ASD (1%). Absence of phrase speech indicates severe language delay, and was largely overlapping with cognitive delay ( $IQ < 70$ ) among the children with ASD.

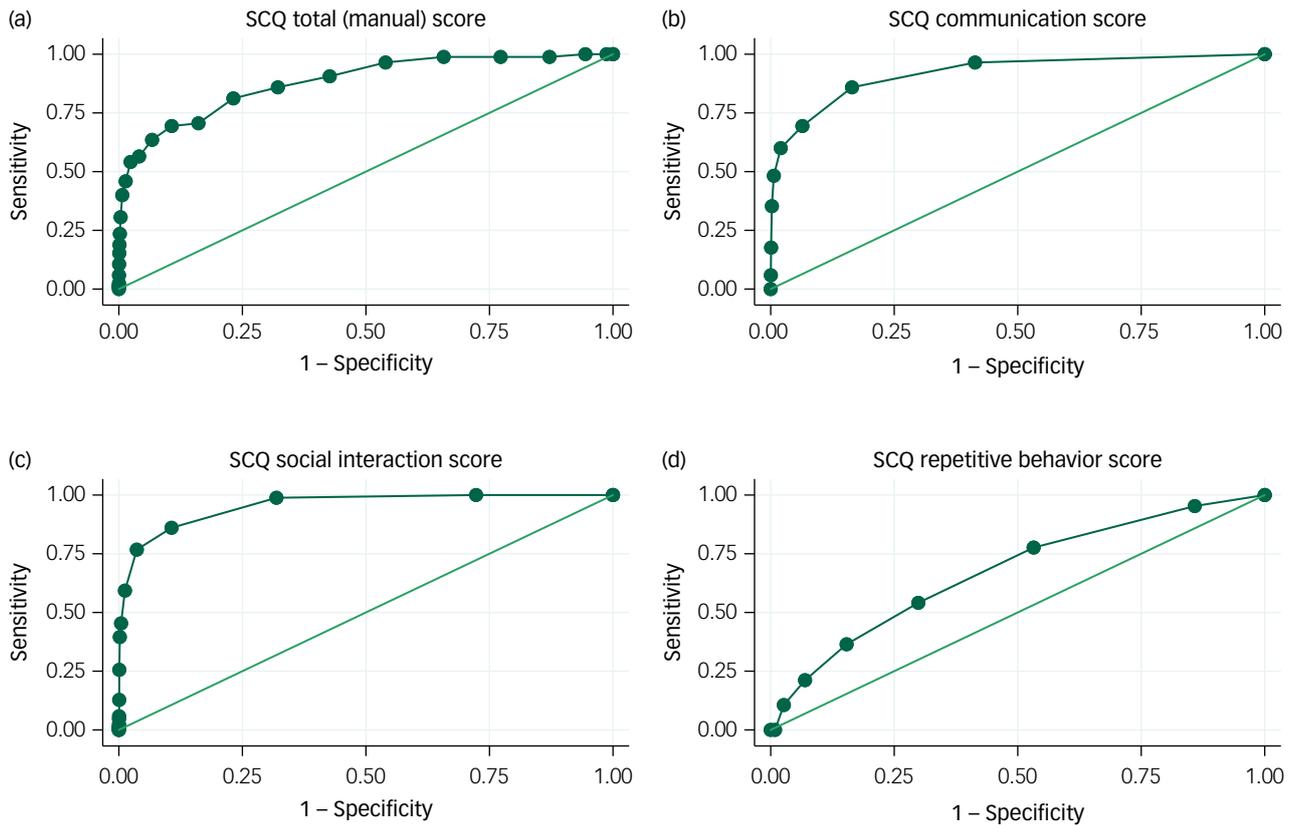
Consequently, we only present results of analyses stratified by the presence or absence of phrase speech.

In non-verbal children with ASD, mean SCQ total score was 13.7 (s.d. = 3.3). In children who were verbal with ASD, the distribution of SCQ scores had large degrees of overlap with those of children without ASD, both for the total score and the different SCQ domains. Mean SCQ total score for children who were verbal with ASD was 9.3 (s.d. = 4.4), whereas the mean SCQ score for children without ASD was 6.2 (s.d. = 3.3).

In Table 1, we show the proportions of positive responses, with corresponding LR estimates, for each of the 40 SCQ items. The ability to differentiate between individuals with and without ASD varied widely. The items covering repetitive language (3–6) and behaviours (7–8, 11–16) had high levels of endorsement in those with ASD and the general population alike, with correspondingly low LR values. In non-verbal children with ASD, the communication items that do not require phrase speech (20–25, 34–35) performed better, with LR values above two for all items except one. The 15 social interaction items (9–10, 19, 26–33, 36–37, 39–40) all had LR values above two for non-verbal children with ASD, and a majority of these items had LR values above 10.

In verbal children with ASD, the repetitive language items had LR values close to one, which indicates a lack of ability to distinguish these children from children without ASD. LR values were somewhat higher for the other communication items, but three out of eight were below two. The performance of the social interaction items was generally better in verbal children with ASD, with 11 out of 15 LR values above 2.

The majority of children with ASD scored below the cut-off level of  $\geq 15$  recommended in the SCQ manual (Table 2). The sensitivity was 20% (95% CI 16–24) for ASD as a whole. For non-verbal ASD it was 46% (95% CI 35–57), whereas for verbal ASD it was 13% (95% CI 9–17). Lowering the cut-off to 11 increased sensitivity to



**Fig. 1** Receiver operating characteristic (ROC) curves for the Social Communication Questionnaire (SCQ): autism spectrum disorder (ASD) without phrase speech at 36 months.

(a) SCQ total (manual) score, area under ROC curve 0.88 (95% CI 0.84–0.92); (b) SCQ communication score, area under ROC curve 0.92 (95% CI 0.89–0.95); (c) SCQ social interaction score, area under ROC curve 0.95 (95% CI: 0.94–0.97); (d) SCQ repetitive behaviour score; area under ROC curve 0.67 (95% CI: 0.62–0.73). Includes all responders to the Norwegian Mother and Child Cohort 36-month questionnaire ( $n = 58\,520$ ). The sample includes 385 individuals with ASD: 86 without phrase speech at 36 months, 294 with phrase speech at 36 months, and 5 with missing information about phrase speech. For children with ASD without phrase speech, the communication score (b) includes the eight communication items that do not require phrase speech (items 20–25, 34–35). The social interaction score (c) includes the 15 social interaction items (items 9–10, 19, 26–33, 36–37, 39–40). The repetitive behaviour score (d) includes the seven repetitive behaviour items that do not require phrase speech (items 8, 11–16).

42% (95% CI 37–47) for ASD overall, and to 69% (95% CI 58–79) for non-verbal ASD and 34% (95% CI 29–40) for verbal ASD. Specificity was then substantially reduced, from 99% (95% CI 99–99) to 89% (95% CI 89–90). The positive predictive value was reduced from 9% (95% CI 7–11) to 3% (95% CI 2–3). With the cut-off level that was originally chosen for the ABC Study ( $\geq 12$  on the 33 non-verbal items), sensitivity was 64% (95% CI 52–74) for non-verbal ASD and 13% (95% CI 9–17) for verbal ASD.

In Fig. 1, we show ROC curves for non-verbal children with ASD. For the communication and social interaction domains, AUC estimates were above 0.9, indicating excellent performance. For the repetitive behaviour domain, on the other hand, the AUC value of 0.67 (95% CI 0.62–0.73) was poor. Still, the AUC for the total SCQ score was 0.88 (95% CI 0.84–0.92), which is considered good for a screening test.

ROC curves for verbal children with ASD are shown in Fig. 2. For the total SCQ score, the AUC was 0.71 (95% CI 0.68–0.74). This is just above the level of 0.7 that separates fair performance from poor performance. There were no apparent cut-off levels that would allow for both high sensitivity and high specificity. AUC estimates were somewhat higher for the social interaction domain, with an AUC of 0.71 (95% CI 0.68–0.74), than for the communication domain, with an AUC of 0.65 (95% CI 0.62–0.68), and the repetitive behaviour domain, with an AUC of 0.62 (95% CI 0.58–0.65).

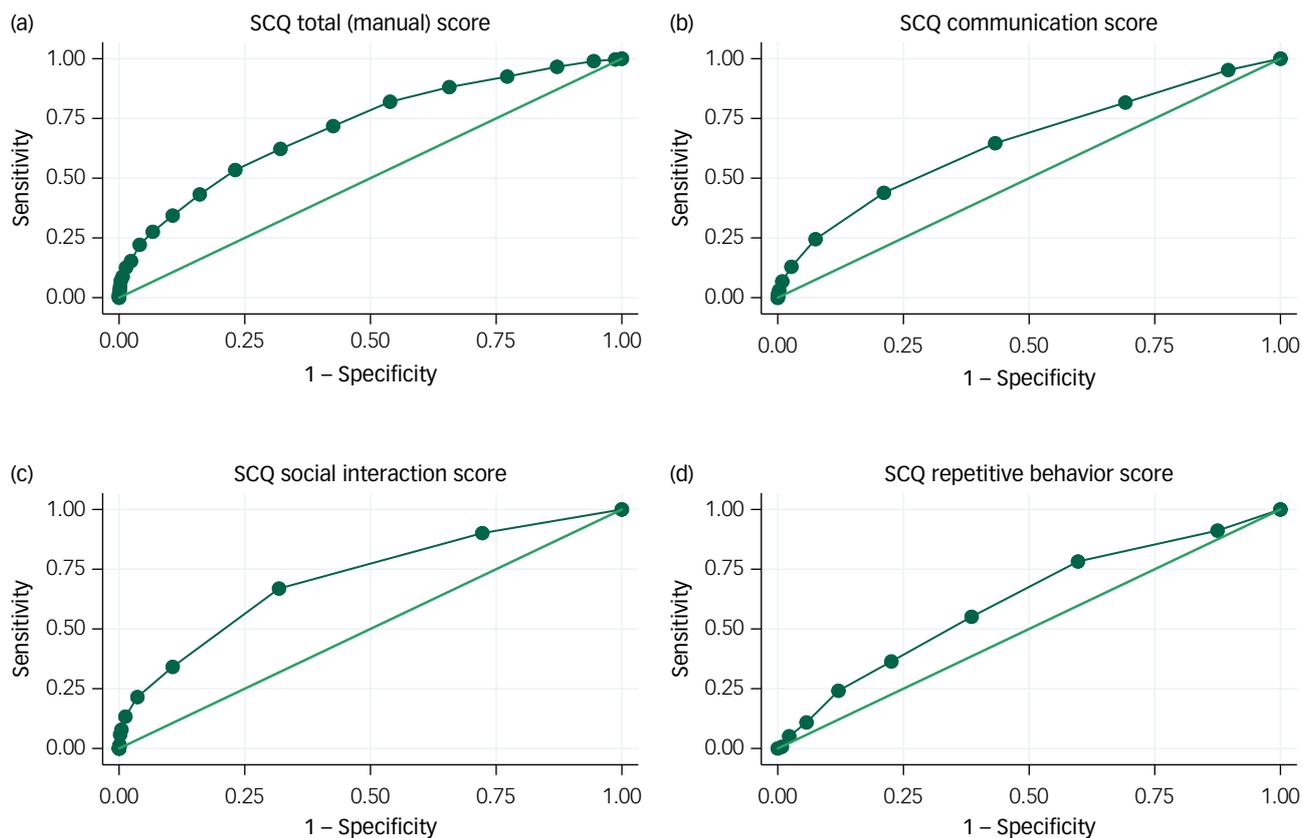
The 36-month questionnaire also included questions about whether the child had already been referred for evaluations by educational services, developmental disability services or child psychiatry services. Of the 86 non-verbal children with ASD, 71 (83%) had been referred for services by age 3 years. Of the 294 verbal children with ASD, only 52 (18%) had been referred at this age.

A large proportion of the cohort (42%) did not participate in the ASD screening, since the parents did not return the 36-month questionnaire. As shown in Supplementary Table 1 available at <https://doi.org/10.1192/bjo.2019.34>, the ones most likely to not respond to the questionnaire were parents without a college or university education, young parents and mothers who were single. Those characteristics were also associated with a higher risk of having children with ASD. Because the sensitivity was so low for ASD overall, we have not included analyses of how screening performance may vary across sociodemographic characteristics of the children and their families.

## Discussion

### Interpretation of the main findings

We investigated the sensitivity and specificity of population-based ASD screening with the SCQ at age 36 months. Most of the individuals with ASD had developed phrase speech at 36 months and did



**Fig. 2** Receiver operating characteristic (ROC) curves for the Social Communication Questionnaire (SCQ): autism spectrum disorder (ASD) with phrase speech at 36 months.

(a) SCQ total (manual score), area under ROC curve 0.71 (95% CI 0.68–0.74); (b) SCQ communication score, area under ROC curve 0.65 (95% CI 0.62–0.68); (c) SCQ social interaction score, area under ROC curve 0.71 (95% CI 0.68–0.74); (d) SCQ repetitive behaviour score, area under ROC curve 0.62 (95% CI 0.58–0.65). Includes all responders to the Norwegian Mother and Child Cohort (MoBa) 36-month questionnaire ( $n = 58\,520$ ). The sample includes 385 individuals with ASD: 86 without phrase speech at 36 months, 294 with phrase speech at 36 months and 5 with missing information about phrase speech. For children with ASD with phrase speech, the communication score (b) includes the 13 communication items (items 2–6, 20–25, 34–35). The social interaction score (c) includes the 15 social interaction items (items 9–10, 19, 26–33, 36–37, 39–40). The repetitive behaviour score (d) includes the eight repetitive behaviour items (items 7–8, 11–16).

not have significant cognitive delays. Their SCQ score distributions were largely overlapping with the distributions in those without ASD, and there was no obvious way of achieving higher levels of sensitivity without severely compromising specificity. Our findings suggest that it is hard to detect these children through parental report at this age. The ASD symptoms may be too subtle or not yet apparent,<sup>21</sup> or the symptoms may be underreported when parents are not concerned or actively seeking help.<sup>22</sup>

Sensitivity was higher for the children with ASD who had not developed phrase speech and were significantly delayed in their cognitive development. The majority of these children had already been referred to ordinary health or educational services before screening was conducted, suggesting that screening would not be necessary to elicit access to services for them. Screening may still have benefits, though, because it may accelerate the diagnostic process of differentiating ASD from other developmental disorders, thus leading to earlier initiation of adequate interventions. Another benefit, which has not been assessed in this study, is that screening may also detect children with other types of developmental delay.

The SCQ was originally validated for older children. It might have performed better at older ages in this cohort, but then the benefits of early identification would be reduced. The young age at screening may explain why the SCQ questions about repetitive language and behaviours were of little use in distinguishing individuals with ASD from those without. Repetitive language and behaviours are core features of ASD, but such behaviours are also common in

typically developing 3-year-olds.<sup>23</sup> Furthermore, the fact that the mothers were not aware that they were completing a screening instrument for autism could have made them more likely to interpret the ‘pathological’ response as favourable, i.e., as signs of the child being active and inquisitive, and practising new skills.

The SCQ questions about social interaction and non-verbal communication query deficits that are expected to be present at an early age in children with ASD, such as lack of eye gaze, social smiling or joint attention. The items of these domains had high levels of endorsement for children with ASD with significant developmental delay, but not for the majority of those with ASD with phrase speech. Limiting the use of SCQ to these two domains would not substantially improve sensitivity.

In the current American guidelines, ASD screening is recommended as part of general well-child visits at ages 18 and 24 months.<sup>2</sup> The 23 items of the M-CHAT were included in the MoBa 18-month questionnaire. Previous analyses have shown that most children with ASD scored below the pre-defined cut-off levels on this instrument as well.<sup>24</sup> This suggests that screening would not have performed any better at earlier ages, or with another screening instrument.

### The challenges of autism screening

A critical challenge in autism screening is the lack of a clear boundary between the autism spectrum and the normal range of

development. Autistic traits are continuously distributed in the population.<sup>25</sup> The threshold for receiving a diagnosis depends on clinical judgement, which may vary between clinicians and change over time. In screening, the symptom scores of children with ASD may also overlap with those of other children. These features of the disorder suggest that it does not easily lend itself to detection by screening, at least not for children without general developmental delay. Many researchers and clinicians will still argue that efforts should be made to identify children with ASD as early as possible, because they are likely to have a lot to gain from early intervention.<sup>5</sup> If this is to succeed, it will require novel approaches that move beyond screening with a uniform single-informant checklists at a single time point.<sup>26</sup>

The harms and costs of screening must also be considered. The currently recommended cut-off level for the SCQ score is  $\geq 11$ .<sup>11</sup> When we tested this cut-off in our analyses, the sensitivity was 42% for ASD overall and the specificity was 89%. If we had used this cut-off in practice, 11% of all the children would have been defined as screen-positive and required further follow-up. Facilitating follow-up for such a large number of children would be very expensive within the ordinary healthcare system. Another problem is the psychological distress inflicted on so many parents by labelling their children as 'screen-positive for autism' when most of them do not have the disorder.

Another challenge of universal health screening is that those who need it the most may be the least likely to participate. This was illustrated by the participation in screening and clinical assessments in our study, which was considerably lower than expected. The proportions responding to the 36-month questionnaire were considerably lower among parents with lower socioeconomic status. The fact that the SCQ was embedded in a large questionnaire may have contributed to the attrition. On the other hand, this is a cohort where most participants are supposedly motivated to participate in research. Participation might have been even lower if the screening had been conducted in a population with a higher representation of disadvantaged parents and ethnic minorities. The problems with low participation might have been mitigated if ASD screening was integrated into routine clinical practice. Nonetheless, integrating the use of the SCQ into routine clinical practice would probably not have improved the sensitivity of screening.

## Strengths and limitations

The main strength of our study was its ability to capture false negatives and thereby produce reliable estimates of sensitivity and specificity of early ASD screening with the SCQ. This allowed us to remedy an acknowledged limitation of previous studies.<sup>5</sup> However, the study was not primarily designed as a screening study, and had no designated non-screened control group. This limits our ability to quantify the costs and benefits of ASD screening. Another limitation was the large proportion of non-responders to the screening questionnaire. These parents had a higher risk of having children with ASD, and the lack of data from these parents may have led to a slight underestimation of sensitivity. On the other hand, a number of individuals with ASD are still undetected among the younger children in the cohort. Most of these are likely to be verbal children with ASD with low SCQ scores at 36 months, which will probably lead to lower estimates of sensitivity when the cohort participants grow older.

In conclusion, early ASD screening with the SCQ had low sensitivity in this study. The screening captured mainly children with ASD with significant developmental delay, while missing those children with ASD who had cognitive skills in the normal range.

Currently available evidence does not support universal early screening programmes for ASD.

**Pål Surén**, MD, MPH, PhD, Physician and Researcher, Norwegian Institute of Public Health, Norway; **Alexandra Saasen-Havdahl**, PhD, Psychologist and Researcher, Norwegian Institute of Public Health; Nic Waals Institute, Lovisenberg Hospital, Norway; and MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, UK; **Michaeline Bresnahan**, PhD, Assistant Professor, Department of Epidemiology, Mailman School of Public Health, Columbia University; and New York State Psychiatric Institute, USA; **Deborah Hirtz**, MD, Physician and Researcher, National Institute of Neurological Disorders and Stroke, USA; **Mady Hornig**, MD, Associate Professor, Department of Epidemiology, Mailman School of Public Health, Columbia University, USA; **Catherine Lord**, PhD, Professor, Semel Institute for Neuroscience and Human Behavior, UCLA David Geffen School of Medicine, USA; **Ted Reichborn-Kjennerud**, MD, PhD, Department Director and Professor, Norwegian Institute of Public Health; and Institute of Clinical Medicine, University of Oslo, Norway; **Synnve Schjølberg**, MSc, Psychologist and Researcher, Norwegian Institute of Public Health, Norway; **Anne-Siri Øyen**, PhD, Psychologist and Researcher, Norwegian Institute of Public Health; and Nic Waals Institute, Lovisenberg Hospital, Norway; **Per Magnus**, MD, PhD, Center Director and Professor, Norwegian Institute of Public Health, Norway; **Ezra Susser**, MD, DrPH, Professor, Department of Epidemiology, Mailman School of Public Health, Columbia University; and New York State Psychiatric Institute, USA; **W. Ian Lipkin**, MD, Center Director and Professor, Center for Infection and Immunity, Mailman School of Public Health, Columbia University; and Vagelos College of Physicians and Surgeons, Columbia University, USA; **Camilla Stoltenberg**, MD, PhD, Director General and Professor, Norwegian Institute of Public Health; and Department of Global Public Health and Primary Care, University of Bergen, Norway

**Correspondence:** Pål Surén, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, N-0403 Oslo, Norway. Email: [pal.suren@fhi.no](mailto:pal.suren@fhi.no)

First received 13 Dec 2018, final revision 15 Apr 2019, accepted 24 Apr 2019

## Funding

P.S. is supported by the OAK Foundation (grant OCAY-12-326). A.S.-H. is supported by the South-East Norway Regional Health Authority (grant 2018059). She is also a member of the MRC Integrative Epidemiology Unit, which is supported by the University of Bristol and the UK Medical Research Council (MC\_UU\_12013/1). C. L. receives funding from R91HD081199, R01MH104423 and SFARI 345327. P.M. is supported by the Research Council of Norway through its Centers of Excellence Funding Scheme, project number 262700. MoBa receives funding from the Norwegian Ministry of Health and Care Services and the Ministry of Education and Research. The cohort has been extensively supported by the NIH/NIEHS (contract N01-ES-75558) and the NIH/NINDS (grant 1 U01 NS 047537-01 and grant 2 U01 NS 047537-06A1). The NIH/NINDS grants have also funded the Autism Birth Cohort Study.

## Acknowledgements

We are grateful to the families participating in MoBa and the ABC Study. We thank Therese Wardenær Bakke and Kaja Schau Knatten for their efforts in facilitating the data collection.

## Supplementary material

Supplementary material is available online at <https://doi.org/10.1192/bjo.2019.34>.

## References

- 1 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders – DSM-IV (4th edn)*. American Psychiatric Association, 2013.
- 2 Johnson CP, Myers SM. Identification and evaluation of children with autism spectrum disorders. *Pediatrics* 2007; **120**: 1183–215.
- 3 National Institute for Health and Care Excellence (NICE). *Autism Spectrum Disorder in Under 19s: Recognition, Referral and Diagnosis*. NICE, 2011 (<https://www.nice.org.uk/guidance/cg128>).
- 4 Robins DL, Casagrande K, Barton M, Chen CM, Dumont-Mathieu T, Fein D. Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics* 2014; **133**: 37–45.
- 5 McPheeters ML, Weitlauf A, Vohorn A, Taylor C, Sathe NA, Krishnaswami S, et al. *U.S. Preventive Services Task Force Evidence Syntheses, formerly Systematic Evidence Reviews. Screening for Autism Spectrum Disorder in Young Children: A Systematic Evidence Review for the US Preventive Services Task Force*. Agency for Healthcare Research and Quality, 2016.
- 6 Andermann A, Blancquaert I, Beauchamp S, Dery V. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. *Bull World Health Organ* 2008; **86**: 317–9.
- 7 Berument SK, Rutter M, Lord C, Pickles A, Bailey A. Autism screening questionnaire: diagnostic validity. *Br J Psychiatry* 1999; **175**: 444–51.

- 8 Magnus P, Birke C, Vejrup K, Haugan A, Alsaker E, Daltveit AK, et al. Cohort profile update: the Norwegian mother and child cohort study (MoBa). *Int J Epidemiol* 2016; **45**: 382–8.
- 9 Stoltenberg C, Schjolberg S, Bresnahan M, Hornig M, Hirtz D, Dahl C, et al. The Autism Birth Cohort: a paradigm for gene-environment-timing research. *Mol Psychiatry* 2010; **15**: 676–80.
- 10 Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 1994; **24**: 659–85.
- 11 Corsello C, Hus V, Pickles A, Risi S, Cook Jr EH, Leventhal BL, et al. Between a ROC and a hard place: decision making and making decisions about using the SCQ. *J Child Psychol Psychiatry* 2007; **48**: 932–40.
- 12 Moody EJ, Reyes N, Ledbetter C, Wiggins L, DiGiuseppi C, Alexander A, et al. Screening for autism with the SRS and SCQ: variations across demographic, developmental and behavioral factors in preschool children. *J Autism Dev Disord* 2017; **47**: 3550–61.
- 13 Barnard-Brak L, Brewer A, Chesnut S, Richman D, Schaeffer AM. The sensitivity and specificity of the social communication questionnaire for autism spectrum with respect to age. *Autism Res* 2016; **9**: 838–45.
- 14 Suren P, Bakken IJ, Aase H, Chin R, Gunnes N, Lie KK, et al. Autism spectrum disorder, ADHD, epilepsy, and cerebral palsy in Norwegian children. *Pediatrics* 2012; **130**: e152–8.
- 15 Lord C, Risi S, Lambrecht L, Cook Jr EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 2000; **30**: 205–23.
- 16 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (4th edn), Text Revision*. American Psychiatric Association, 2000.
- 17 World Health Organization. *International Statistical Classification of Diseases and Related Health Problems 10th Revision*. World Health Organization, 1990.
- 18 Hunink M, Glasziou P, Siegel J, Weeks J, Pliskin J, Elstein A, et al. *Decision Making in Health and Medicine*. Cambridge University Press, 2001.
- 19 Soreide K, Korner H, Soreide JA. Diagnostic accuracy and receiver-operating characteristics curve analysis in surgical research and decision making. *Ann Surg* 2011; **253**: 27–34.
- 20 Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves. *Surgery* 2016; **159**: 1638–45.
- 21 Lai MC, Lombardo MV, Chakrabarti B, Baron-Cohen S. Subgrouping the autism 'spectrum': reflections on DSM-5. *PLoS Biol* 2013; **11**: e1001544.
- 22 Havdahl KA, Bishop SL, Suren P, Oyen AS, Lord C, Pickles A, et al. The influence of parental concern on the utility of autism diagnostic instruments. *Autism Res* 2017; **10**: 1672–86.
- 23 Evans DW, Leckman JF, Carter A, Reznick JS, Henshaw D, King RA, et al. Ritual, habit, and perfectionism: the prevalence and development of compulsive-like behavior in normal young children. *Child Dev* 1997; **68**: 58–68.
- 24 Stenberg N, Bresnahan M, Gunnes N, Hirtz D, Hornig M, Lie KK, et al. Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatr Perinat Epidemiol* 2014; **28**: 255–62.
- 25 Constantino JN, Todd RD. Autistic traits in the general population: a twin study. *Arch Gen Psychiatry* 2003; **60**: 524–30.
- 26 Reiersen AM. Early identification of autism spectrum disorder: is diagnosis by age 3 a reasonable goal? *J Am Acad Child Adolesc Psychiatry* 2017; **56**: 284–5.

