Original Article

Evaluating Reliability and Predictive Validity of the Persian Translation of Quantitative Checklist for Autism in Toddlers (Q-CHAT)

Mehrdad Mohammadian, MD[1]
Hadi Zarafshan, MA[2]
Mohammad Reza Mohammadi, MD[2]
Issa Karimi, BA[1]

1 Iran University of Medical Sciences, Tehran, Iran .
2Tehran University of Medical Sciences, Psychiatry & Psychology Research Center, Tehran, Iran.
.

Corresponding author:
Hadi Zarafshan ,
Tehran University of Medical Science, Psychiatry & Psychology Research Center ,
Tehran, Iran .
Tel: +982155421959
E-mail :
h-zarafshan@razi.tums.ac.ir

Objective: Early screening of autism increases the chance of receiving timely intervention. Using the Parent Report Questionnaires is effective in screening autism. The Q-CHAT is a new instrument that has shown several advantages than other screening tools. Because there is no adequate tool for the early screening of autistic traits in Iranian children, we aimed to investigate the adequacy of the Persian translation of Q-CHAT.
Method: At first, we prepared the Persian translation of the Quantitative Checklist for Autism in Toddlers (Q-CHAT). After that, an appropriate sample was selected and the check list was administered. Our sample included 100 children in two groups (typically developing and autistic children) who had been selected conveniently .Pearson's r was used to determine test-retest reliability, and Cronbach's alpha coefficient was used to explore the internal consistency of Q-CHAT. We used the receiver operating characteristics curve (ROC) to investigate whether Q-CHAT can adequately discriminate between typically developing and ASD children or not. Data analysis was carried out by SPSS 19.
Result: The typically developing group consisted of 50 children with the mean age of 27.14 months, and the ASD group included50 children with the mean age of 29.62 months. The mean of the total score for the typically developing group was 22.4 (SD=6.26) on Q-CHAT and it was 50.94 (SD=12.35) for the ASD group, which was significantly different (p=0.00).The Cronbach's alpha coefficient of the checklist was 0.886, and test-retest reliability was calculated as 0.997 (p<0.01). The estimated area under the curve (AUC) was 0.971. It seems that the total score equal to 30 can be a good cut point to identify toddlers who are at risk of autism (sensitivity= 0.96 and specificity= 0.90).
Conclusion: The Persian translation of Q-CHAT has good reliability and predictive validity and can be used as a screening tool todetect18 to 24 months old children who are at risk of autism.

Keywords: *Autism, Early Screening, Q-CHAT, Reliability, Predictive Validity*

Iran J Psychiatry 2015; 10:1: 64-70

Autism spectrum disorder (ASD) is a developmental disorder recognized with impairments in social communication and social interaction, and restricted and repetitive patterns of behavior, interest and activities (1).
Numerous studies have shown that early intensive intervention is crucial for improvement, and it decreases symptoms of autism in children (2-4). To establish early intervention, the age on which detection is done should be decreased (5, 6).
In recent years, many researchers have tried to identify early signs of autism to screen children who are at risk of autism in early stages. Most experts have emphasized that autism can be diagnosed at 2 years of age and its signs can be detected before the age of 24 months (7-9).
Since there is no biomarker recognized for autism, the diagnosis should be done based on a behavioral profile using screening tools (3, 10). There are two general methods for screening children who are at risk for autism: The first method is based on an evaluation that is made by a clinician and the second is based on using the Parent Report Questionnaire (11)
Using the Parent Report Questionnaire to screen autism has several benefits such as simplicity in implementation, cost effectiveness and time saving(12). Moreover, it is shown that screening based on the Parent Report Questionnaire is as

accurate as screening based on clinicians' examination (13).

In recent years, several instruments have been developed to screen and diagnose autism in children under three years of age; namely, Checklist for Autism in Toddlers (CHAT)(14), Modified Checklist for Autism in Toddlers (M-CHAT)(15), the Early Screening of Autistic Traits Questionnaire (ESAT)(16), The First Year Inventory (FYI)(17). Allison et al. argue that there are several limitations in the current instruments that make us use new instruments; they have mentioned important limitations such as low level of sensitivity, inadequacy to use in general population, focusing on specific ages and using binary scoring system that lead to missing lower degrees of symptoms(18).They introduced a new instrument called Quantitative Checklist for Autism in Toddlers (Q-CHAT) which is the new version of the Checklist for Autism in Toddlers (CHAT)(18).This instrument is called the Parent Report Questionnaire with 25 items scored on a five point Liker scale. In a preliminary study, Allison et al. investigated scores distribution, test retest reliability and internal consistency of Q-CHAT. Their study has shown good reliability (0.82) and internal consistency for this questionnaire(0.67 and 0.81 in general population and ASD group, respectively). It is also shown that scores of Q-CHAT in the general population is approximately normally distributed(18).In another study, Allison et al. examined the predictive validity of the full length and the short version of Q-CHAT, using area under the curve (AUC) measure(19). The area under the curve (AUC) was 0.965 and 0.920 for short and full version of Q-CHAT, respectively (19)

Because no acceptable tool was available for the early screening of autistic traits in Iranian children, we aimed to investigate the adequacy of the Persian translation of Q-CHAT to distinguish the typically developing children fom those children who are at risk of autism.

Material and Methods

To investigate the properties of the Persian translation of the Quantitative Checklist for Autism in Toddlers (Q-CHAT), we went through several stages. At first, we prepared the Persian version of the Quantitative Checklist for Autism in Toddlers (Q-CHAT). Then, we selected an appropriate sample and administered the checklist.

We collected data and used several analyses to explore the characteristics of the Persian translation of the Checklist for Autism in Toddlers (Q-CHAT).

Participants

Our sample included 100 children in two groups (typically developing and autistic children) who were selected conveniently. The typically developing group was selected from a hospital nursery, and based on their parent reports they had no history of developmental delay. To confirm that they did not have any developmental delay, they were evaluated by one of the authors.

Autistic group were selected among clients of a psychiatric hospital and an autism center in Tehran. At first, children who were suspected to have autism due to their developmental delay (especially in communication) were evaluated by a child psychiatrist based on the diagnostic criteria of DSM-IV-TR, and if they were diagnosed with autism, then they were referred to a second child psychiatrist to confirm their diagnosis. Children who were diagnosed by both psychiatrists were selected as the ASD group.

Children who had severe physical disability or were visually or hearing impaired were excluded from the study.

After selecting the appropriate sample, the checklists were filled out by mothers of the selected children.

Measures

Quantitative Checklist for Autism in Toddlers (Q-CHAT) has 25 items and is answered by parents with the purpose of screening the early signs of autism in 18-24 months old toddlers. Items are scored in five level from zero to four (Half the items are reverse-scored).The total score of the checklist was calculated by summing up the scores of each item.

The original version of the Quantitative Checklist for Autism in Toddlers (Q-CHAT) has been retrieved from the Autism Research Center website (http://www.autismresearchcentre.com/arc_tests) and it was then translated in to Persian by one of the authors. To ensure the accuracy of the translation, we back-translated it to English and compared our English version with the original. We made some changes and finalized the Persian version of the Quantitative Checklist for Autism in Toddlers (Q-CHAT).We used the same method of the original version to score the items.

The gold-standard in this study was a psychiatric diagnosis based on DSM-IV-TR criteria.

Data Analysis

We used several statistical methods to analyze the data. Data were entered into SPSS-19 and then descriptive indexes (frequency, mean and standard deviation) and inferential statically tests (independent t-test, Kolmogorov-Smirnov test, Cronbach's alpha, Pearson's r, and the receiver operating characteristic (ROC) curve) were computed to addressed the aim of the study.

Pearson's r was used to determine test-retest reliability, and Cronbach's alpha coefficient was used to explore the internal consistency of Q-CHAT. We used receiver operating characteristics curve (ROC) to investigate whether Q-CHAT can adequately discriminate between the typically developing children and ASD children or not..

Results

Typically developing group consisted of 50 children with the mean age of 27.14 months (SD=7.68);of whom,37 (74%) were boys and 13 (26%) were girls. The ASD group consisted of 50 children with the mean age of 29.62 months (SD=9.43); of whom, 38 (76%) were boys and 12 (24%) were girls .Based on the independent t-test, the mean age of the two groups was not significantly different (t (98) =-1.441, p=0.153, equal variances assumed)

The mean of the total scores for the typically developing group on Q-CHAT was 22.4 (SD=6.26) and it was 50.94 (SD=12.35) for the ASD group. Based on the independent t-test, the means of the total scores of the two groups were significantly different (t (72.657) =-15.564, p=0.00, unequal variances assumed).The comparison of the score distribution of the two groups is demonstrated in Figure 1. The Kolmogorov-Smirnov test revealed that scores in the typically developing group was normally distributed (p=0.2),but it did not apply to the ASD group (p=0.019).

In the typically developing group, the mean of the total scores for boys (m=23.97, SD=6) was significantly higher than girls (m=17.92, SD=4.8) (t (43) =3.033, p=0.004, equal variances assumed), but there were not any significant differences between boys and girls in the ASD group(t (48) =0.193, p=0.848, equal variances assumed).

The reliability of the checklist was calculated with two methods. The test-retest reliability on a sample of 30 ASD children, with a month interval was calculated as 0.997 ($p<0.01$). The internal consistency (Cronbach's alpha) of the checklist was 0.886for all participants. We also calculated Alpha for the ASD and normal group separately (0.741 for the ASD group and 0.366 for the normal group).Correlations between each item and the total score of the questionnaire and also the overall Alphaif an item was excluded are demonstrated in Table 1.

The area under curve (AUC) was calculated as an index of the overall predictive validity and is presented in Table 2 and Figure 2.

Table 1: Correlations between each item and the total score

Items	In all Participants		In the ASD Group		In the Normal Group	
	Correlations between Each Item and the Total Score	Cronbach's Alpha if Item Deleted	Correlations between Each Item and the Total Score	Cronbach's Alpha if Item Deleted	Correlations between Each Item and the Total Score	Cronbach's Alpha if Item Deleted
1. Look When Call Name	.630	.879	.374	.729	.327	.318
2. Eye Contact	.743	.877	.459	.728	.243	.338
3. Line Objects Up *	-.140	.896	-.103	.758	-.144	.416
4. Understand Child's Speech	.711	.875	.558	.717	.206	.331
5. Proto-Imperative Pointing	.290	.887	.465	.718	-.001	.388
6. Proto-Declarative Pointing	.623	.877	.584	.707	.023	.376
7. Interest Maintained by Spinning Object *	.459	.882	.215	.737	.299	.323
8. Number of Words *	.678	.876	.426	.726	.173	.334
9. Pretend Play	.646	.877	.608	.709	.173	.335
10. Follow a Look	.493	.881	.291	.732	.167	.335
11. Sniff/Lick Unusual Objects *	.435	.883	.161	.743	-.001	.378
12. Use of hand as tool *	.511	.881	.025	.754	.105	.353
13. Walk on Tiptoes *	.311	.885	.276	.733	-.061	.385
14. Adapt to Change in Routine activities	.270	.886	.159	.739	-.106	.385
15. Offer Comfort	.585	.878	.406	.724	.006	.378
16. Do the Same Thing Over and Over Again *	.495	.881	.186	.739	.240	.321
17. Typicality of First Words	.566	.879	.225	.737	.080	.359
18. Echolalia *	-.090	.893	-.067	.752	-.023	.383
19. Gestures	.638	.877	.389	.725	.015	.370
20. Unusual Finger Movements *	.446	.882	.157	.741	-.081	.371
21. Check Reaction	.557	.879	.556	.714	.218	.325
22. Maintenance of Interest *	.369	.884	.307	.732	.232	.327
23. Twiddle Objects Repetitively *	.299	.886	.130	.745	.124	.348
24. Oversensitive to Noise *	.386	.884	.050	.749	.183	.331
25. Stare at Nothing with no Purpose *	.567	.879	.339	.729	-.120	.380

*items that were Scored Reversely

Table 2: Area under Curve (AUC) for 25 and 20 Items of Q-CHAT

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
.971	.018	.000	.936	1.006

Table 3: Sensitivity and Specificity of the Total Scores of Q-CHAT

Total score	Sensitivity	Specificity
8.0000	1.000	.000
10.5000	1.000	.020
13.0000	1.000	.060
14.5000	1.000	.080
15.5000	1.000	.100
16.5000	1.000	.140
17.5000	.980	.180
18.5000	.980	.280
19.5000	.980	.360
20.5000	.980	.420
21.5000	.980	.460
22.5000	.980	.540
23.5000	.980	.620
25.0000	.980	.700
26.5000	.980	.780
27.5000	.980	.820
28.5000	.960	.860
30.0000	.960	.900
31.5000	.940	.900
32.5000	.920	.920
33.5000	.920	.940
34.5000	.860	.940
36.0000	.840	.940
38.0000	.820	1.00
40.5000	.760	1.00
43.5000	.720	1.00
46.0000	.680	1.00
47.5000	.660	1.00
48.5000	.640	1.00
49.5000	.600	1.00
50.5000	.560	1.00
52.5000	.540	1.00
54.5000	.500	1.00
55.5000	.420	1.00
56.5000	.360	1.00
57.5000	.320	1.00
58.5000	.300	1.00
59.5000	.280	1.00
60.5000	.260	1.00
62.0000	.240	1.00
63.5000	.160	1.00
64.5000	.100	1.00
65.5000	.080	1.00
67.5000	.040	1.00
70.5000	.020	1.00
73.0000	.000	1.00

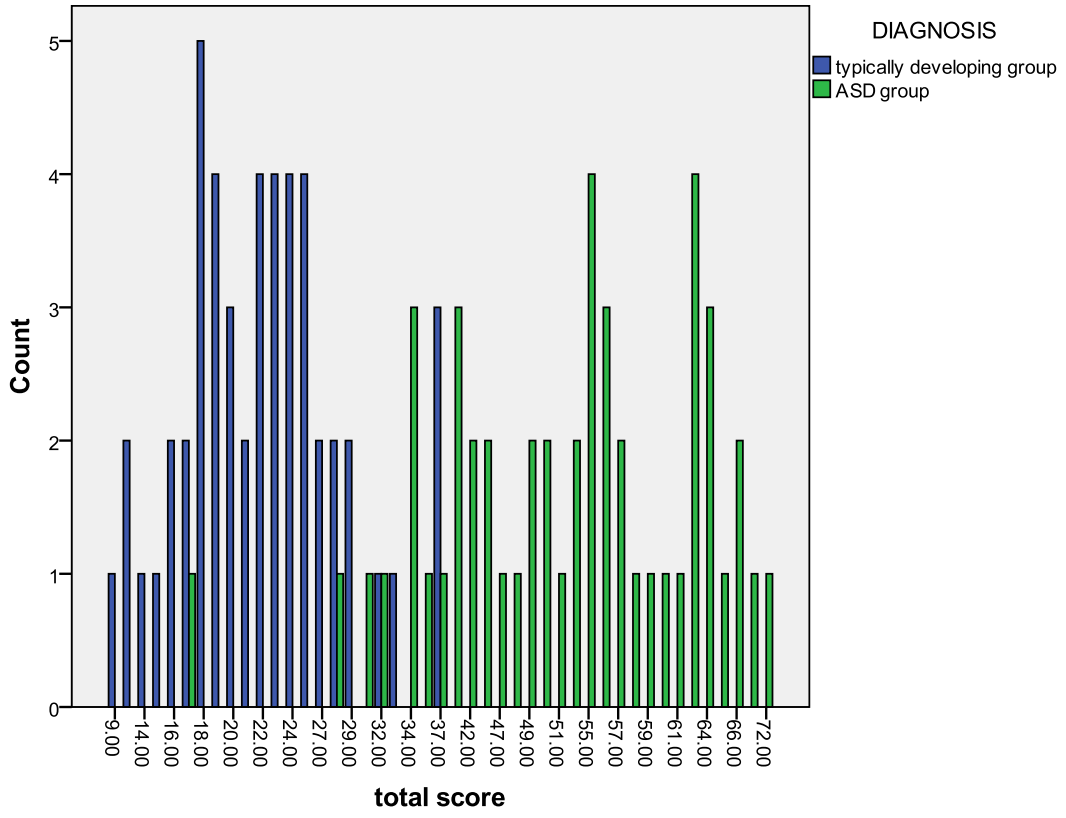

Figure 1: Comparison of Score Distributions of the Two Groups

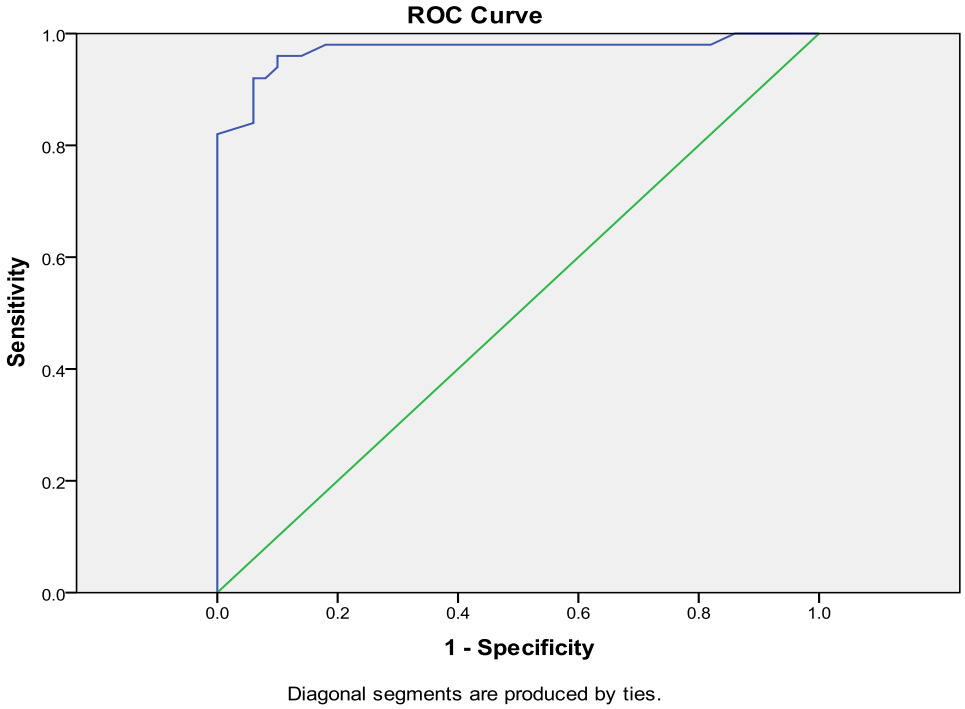

Figure 2: Area Under Curve (AUC)

Discussion

The aim of this study was to investigate the adequacy of the Persian translation of Q-CHAT. However, the checklist was made to screen toddlers between 18-24 months. The mean age of the ASD group in Allison et al. study was more than ours (the mean age was 44 months in our study). The mean age of children in the ASD group was closer to this number (the mean age was 29.62 months).

Our results revealed that the scores of Q-CHAT were normally distributed in the typically developing group. This finding is in line with the original report of Q-CHAT(18) and with this hypothesis that autistic traits are normally distributed in the general population(20). As expected, it is also shown that the mean scores of children with ASD is significantly higher than the typically developing children.

Previous studies have shown that boys have higher level of autistic traits than girls (18, 21, 22). We also found a sex difference in the typically developing group in terms of the total scores of Q-CHAT; boys had significantly higher scores than girls. This finding is also in conjunction with the extreme male brain theory introduced by Baron-Cohen(23) to explain autistic behaviors .

To investigate the reliability of the Persian translation of Q-CHAT, we calculated Cronbach's alpha and test-retest reliability in our sample. Our finding revealed that the Persian translation of the Q-CHAT has good internal consistency (0.866) for all participants (n=100) and test-retest reliability (0.997).By calculating Alpha for ASD (n=50) and normal (n=50) groups separately, Alpha was reduced to 0.741 and 0.366, respectively. In Allison et al. study Cronbach's alpha was calculated for the ASD (n=160) and normal (n=779) groups separately as 0.81 and 0.67, respectively. As reported previously, our Cronbach's alpha was lower than that of the Allison et al. study, and especially it was very low and weak in the normal group (0.366).This may be due to our small sample size and due to this fact that Cronbach's alpha is strongly correlated with sample size.

Considering the correlation between each item and the total score, it was found that some items did not have a significant correlation with the total score in three different modes of analysis. As seen in Table 1, when data of all participant were analyzed, five items (items 3, 5, 14, 18 and 23) had low correlation (<0.3) with the total score .By analyzing data of the two groups (ASD and normal) separately, it was revealed that 13 items in the ASD group and almost all items in the normal group had low correlations with the total scores. As mentioned previously, it was no surprise due to the small sample size. Allison et al. have reported that four items in the ASD group and one item in the normal group did not have a significant correlation with the total score; nevertheless, they did not remove any items from the checklist.

The ability to discriminate between the typically developing children and ASD children is an important factor for a screening tool. We used receiver operating characteristic (ROC) curve to explore the discriminative ability of the Persian translation of the Q-CHAT. The area under the curve (AUC) was used as an index of the overall predictive validity. The area under the curve (AUC) in our study was 0.971, which is very good. In the preliminary study of Q-CHAT, authors did not calculate AUC, but in a recent study, Allison et al. have reported area under the curve (AUC) of equal to 0.920.There is a similarity between the result of our study and that of Allison et al. study in terms of the predictive validity of Q-CHAT which is an indicator of good predictive validity of Q-CHAT.

In terms of optimum cutoff point, it should be noted that choosing the best cutoff point depends on the usage of a test. If we use a test for the purpose of screening, we should choose a cutoff point with a higher sensitivity (24). In Table 3, we demonstrated sensitivity and specificity for all cutoff points of Q-CHAT. Considering our small sample size, the results should be carefully interpreted, but it seems that the total score equal to 30 can be a good cut point to determine the toddlers who are at risk of autism. There is a report for optimum cutoff point of the short version of the Q-CHAT (Q-CHAT-10), but we did not find any report about the cutoff point of the full version of the Q-CHAT (Q-CHAT-25). Based on the Allison et al.(19) study, a cut-point of 3 for the Q-CHAT-10, has good sensitivity (0.91) and specificity(0.89).

Limitations

As noted, the most important limitation of this study that could influence the results was the small sample size; this was due to the low prevalence of autism and to our focus on a restricted age range. For better investigation of the Persian translation of the Quantitative Checklist for Autism in Toddlers (Q-CHAT), this study should be replicated with a larger sample.

Conclusion

Our study showed that the Persian translation of the Quantitative Checklist for Autism in Toddlers (Q-CHAT) has good reliability and predictive validity and can be used as a screening tool for detecting children between 18 to 24 months who are at risk of autism. In other words, the Persian translation of the Quantitative Checklist for Autism in Toddlers (Q-CHAT) is a good option to be used in early screening programs in the general population in Iran.

Acknowledgement

We acknowledge the financial support of the Iran University of Medical Sciences through a grant from Mental Health Research Center (MHRC). Wealso wish

to thank all families who agreed to participate in our study.

References

1. American Psychiatric Association. , American Psychiatric Association. DSM-5 Task Force. Diagnostic and statistical manual of mental disorders : DSM-5. 5th. City: American Psychiatric Association; 2013.
2. Harris SL , Handleman JS. Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. Journal of autism and developmental disorders 2000; 30: 137-142.
3. Dababnah S, Parish SL, Turner Brown L , Hooper SR. Early screening for autism spectrum disorders: A primer for social work practice. Children and Youth Services Review 2011; 33: 265-273.
4. Daniels AM , Mandell DS. Children's Compliance with American Academy of Pediatrics' Well-Child Care Visit Guidelines and the Early Detection of Autism. Journal of autism and developmental disorders 2013; 43: 2844-2854.
5. National Research Council (U.S.). Committee on Educational Interventions for Children with Autism. Educating children with autism. City: National Academy Press; 2001.
6. Oosterling IJ, Wensing M, Swinkels SH, Van Der Gaag RJ, Visser JC, Woudenberg T, et al. Advancing early detection of autism spectrum disorder by applying an integrated two-stage screening approach. Journal of Child Psychology and Psychiatry 2010; 51: 250-258.
7. Johnson CP. Recognition of autism before age 2 years. Pediatrics in Review 2008; 29: 86-96.
8. Johnson CP , Myers SM. Identification and evaluation of children with autism spectrum disorders. Pediatrics 2007; 120: 1183-1215.
9. Dumont-Mathieu T , Fein D. Screening for autism in young children: The Modified Checklist for Autism in Toddlers (M-CHAT) and other measures. Mental retardation and developmental disabilities research reviews 2005; 11: 253-262.
10. Chawarska K, Klin A, Paul R , Volkmar F. Autism spectrum disorder in the second year: Stability and change in syndrome expression. Journal of Child Psychology and Psychiatry 2007; 48: 128-138.
11. Steiner AM, Goldsmith TR, Snow AV , Chawarska K. Practitioner's guide to assessment of autism spectrum disorders in infants and toddlers. Journal of autism and developmental disorders 2012; 42: 1183-1196.
12. Barton ML, Dumont-Mathieu T , Fein D. Screening young children for autism spectrum disorders in primary practice. Journal of autism and developmental disorders 2012; 42: 1165-1174.
13. Pierce K, Carter C, Weinfeld M, Desmond J, Hazin R, Bjork R, et al. Detecting, studying, and treating autism early: the one-year well-baby check-up approach. The Journal of pediatrics 2011; 159: 458-465. e456.
14. Baron-Cohen S, Allen J , Gillberg C. Can autism be detected at 18 months? The needle, the haystack, and the CHAT. The British Journal of Psychiatry 1992; 161: 839-843.
15. Robins DL, Fein D, Barton ML , Green JA. The Modified Checklist for Autism in Toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. Journal of autism and developmental disorders 2001; 31: 131-144.
16. Dietz C, Swinkels S, van Daalen E, van Engeland H , Buitelaar JK. Screening for autistic spectrum disorder in children aged 14–15 months. II: population screening with the Early Screening of Autistic Traits Questionnaire (ESAT). Design and general findings. Journal of autism and developmental disorders 2006; 36: 713-722.
17. Reznick JS, Baranek GT, Reavis S, Watson LR , Crais ER. A parent-report instrument for identifying one-year-olds at risk for an eventual diagnosis of autism: the first year inventory. Journal of autism and developmental disorders 2007; 37: 1691-1710.
18. Allison C, Baron-Cohen S, Wheelwright S, Charman T, Richler J, Pasco G, et al. The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): A normally distributed quantitative measure of autistic traits at 18–24 months of age: Preliminary report. Journal of autism and developmental disorders 2008; 38: 1414-1425.
19. Allison C, Auyeung B , Baron-Cohen S. Toward brief "red flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. Journal of the American Academy of Child & Adolescent Psychiatry 2012; 51: 202-212. e207.
20. Baron-Cohen S, Wheelwright S, Skinner R, Martin J , Clubley E. The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. Journal of autism and developmental disorders 2001; 31: 5-17.
21. Leekam S, Tandos J, McConachie H, Meins E, Parkinson K, Wright C, et al. Repetitive behaviours in typically developing 2-year-olds. Journal of Child Psychology and Psychiatry 2007; 48: 1131-1138.
22. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, et al. 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 2003; 33: 427-433.
23. Baron-Cohen S. The extreme male brain theory of autism. Trends in cognitive sciences 2002; 6: 248-254.
24. Doi SAR , Williams GM. Methods of clinical epidemiology. City: Springer; 2013.