



Validation of Rapid Interactive Screening Test for Autism in Toddlers Using Autism Diagnostic Observation Schedule™ Second Edition in Children at High-Risk for Autism Spectrum Disorder

Xue-Jun Kong^{1,2*}, Hannah Tayla Sherman¹, Ruiyi Tian¹, Madelyn Koh¹, Siyu Liu¹, Alice Chukun Li¹ and William S. Stone²

¹ Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, United States,

² Department of Psychiatry, Massachusetts Mental Health Center, Beth Israel Deaconess Medical Center, Boston, MA, United States

OPEN ACCESS

Edited by:

Costanza Colombi,
Fondazione Stella Maris (IRCCS), Italy

Reviewed by:

Lonnie Zwaigenbaum,
University of Alberta, Canada
Michele Roccella,
University of Palermo, Italy

*Correspondence:

Xue-Jun Kong
xkong1@mgh.harvard.edu

Specialty section:

This article was submitted to
Autism,
a section of the journal
Frontiers in Psychiatry

Received: 07 July 2021

Accepted: 06 September 2021

Published: 01 October 2021

Citation:

Kong X-J, Sherman HT, Tian R, Koh M, Liu S, Li AC and Stone WS (2021) Validation of Rapid Interactive Screening Test for Autism in Toddlers Using Autism Diagnostic Observation Schedule™ Second Edition in Children at High-Risk for Autism Spectrum Disorder. *Front. Psychiatry* 12:737890. doi: 10.3389/fpsy.2021.737890

The Rapid Interactive screening Test for Autism in Toddlers (RITA-T) is a fast and inexpensive early screening measure for autism spectrum disorder (ASD) that was tested previously in children 18–36 months-old; the current validation study compared the RITA-T with the Autism Diagnostic Observation Schedule™ Second Edition (ADOS-2). The hypothesis is to validate the RITA-T with comparison to the ADOS-2. Thirty-five individuals (18–84 months-old) identified as at risk for ASD received the RITA-T and the ADOS-2 during a single visit. Participants were split into two age groups and both whole-group and sub-group data analysis were conducted. With all participants, RITA-T scores correlated significantly with ADOS-2 total scores ($P < 0.001$), social affect (SA) sub-scores ($P < 0.001$), and restrictive and repetitive behavior (RRB) sub-scores ($P < 0.05$). Similarly, ADOS-2 total and SA scores were significantly correlated in both age groups, while the RRB sub-score was only significant in females ($P < 0.05$). Lastly, correlations using subgroups based on ethnicity were only significant in the minority (“Other”) group for ADOS-2 total scores and in the Asian group for SA sub-scores ($P < 0.05$). Our receiver operating characteristic analysis showed that the optimal cut-off score of the RITA-T was consistently at 14, with a sensitivity of 81% and a specificity of 89% in the combined age group with the ADOS-2 and with a sensitivity 74% and specificity 50% with the DSM-5; The area under the curve was 0.84 (95%CI: 0.69–0.99) for ASD classified by ADOS-2 and 0.89 (95%CI: 0.79–0.99) for ASD diagnosed by DSM-5. The RITA-T performed similarly to the ADOS-2 when both were administered in a single visit. Significant correlations between the measures help validate the potential usefulness of the RITA-T as a rapid early screening measure of ASD. This study helps to show that the RITA-T may be used in a larger age range than originally reported and in

different ethnic groups. The study involves human participants and was reviewed and approved by the Institutional Review Board (IRB) of Massachusetts General Hospital (MGH, 2017P0000857).

Keywords: early screening, Asian, social affect (SA), restricted and repetitive behaviors (RRB), autism spectrum disorder (ASD)

INTRODUCTION

Early diagnosis and intervention significantly impact the prognosis of individuals with Autism Spectrum Disorder (ASD) and underscores the importance of easily applied early detection and screening tools (1–7). Standard evaluation methods for ASD, such as Autism Diagnostic Observation Schedule™ (ADOS) and *Autism Diagnostic Interview-Revised* (ADI-R), have been developed to provide an official diagnosis (8, 9), but these standard methods are often lengthy, difficult, and costly. Shortages in resources for such evaluations may thus delay diagnosis and treatment (10, 11). The mean age of diagnosis is still 4–5 years old worldwide despite recent advances and efforts (7). In addition, the difficulties posed by standard evaluation methods are even greater among minority groups (e.g., African American, Hispanic, Asians, and etc.) in the USA due to language and cultural barriers, including a bias against mental diseases or conditions (12–15). To improve this situation, easier, and faster alternative methods have been proposed and applied with various feasibilities and sensitivities (6, 16). As a group, these methods need more study before they can be recommended with confidence to families, educators, and healthcare providers.

The Rapid Interactive screening Test for Autism in Toddlers (RITA-T) is one of the alternative screening methods that may have the potential for widespread use due to its easy administration and rapid interpretation. The RITA-T screener could better stratify the likelihood of ASD, thereby leading to more focused evaluations and faster diagnoses. The RITA-T is intended to function as an ASD screening test (a “level 2 screener”), but not as a general population screener, such as the Modified Checklist for Autism in Toddlers (M-CHAT; a “level 1 screener”). Relative to M-CHAT, which is a freely accessible questionnaire that is completed by either a parent or a caregiver, the RITA-T is a level 2 screener that can be administered by professionals working with young children, including teachers, childcare workers, pediatricians, nurses, family physicians, and pediatric behavioral health professionals. As such, the RITA-T streamlines the access of screening for those who need further evaluation directing a formal diagnosis for ASD. Training for the RITA-T is 3 h. This behavioral evaluation includes 9 interactive games that can be completed in 10 min. The first published paper using the RITA-T compared it to clinical diagnosis ($n = 60$) and reported good sensitivity (100%) and encouraging specificity (84%) (16). A larger study ($n = 239$) recently reported good sensitivity (97%) but lower specificity (71%) with the same cut off score of 14 (17). However, the RITA-T score has not yet been evaluated in the context of the ADOS-2 sub-scores, which includes social affect (SA) and restricted and repetitive behavior (RRB). This is significant because the diagnosis of ASD requires

both SA deficits and RRB. If someone only has the SA deficit, for example, but not the RRB deficit, the diagnostic criteria for ASD are not met. The RITA-T is designed to evaluate only SA without assessment or scoring of RRB, which might explain the relatively lower specificity in both published studies. In addition, the RITA-T has also not been assessed across different racial or extended age groups beyond toddlers (16, 17). In this study, we aim to fill these gaps by further evaluating the validity of the RITA-T through a direct comparison with ADOS-2, as part of the larger goal of expanding available resources for early detection of ASD, particularly extending the age group to and beyond preschoolers as the current median age of ASD diagnosis is reported to be 51 months old (7, 18, 19).

MATERIALS AND METHODS

Participants

Thirty-five individuals aged 18–84 months old toddlers and preschoolers (mean = 3.53 years, standard deviation = 1.72 years) participated in this study. During participant recruitment, we reached out to different ethnic groups to encourage a more ethnically heterogeneous study cohort. The enrolled study participants included 15 White (43%), 11 Asian (31%), and 9 other race (26%) subjects. Asian is defined here as an individual who is a descendant of an individual who was born in Southeast Asia, the Far east or the Indian subcontinent. Twenty-three males and twelve females participated in this study. Sex is defined by sex chromosomes composition: males have XY chromosomes, and females have XX chromosomes. Individuals were recruited through clinical care clinics and online recruitment sites. Most participants came from Massachusetts and its surrounding states. All participants were included after being identified as high-risk for ASD by clinicians or caregivers. The high-risk status was confirmed through a phone screening prior to enrollment. Inclusion criteria included one or more of the following: (1) at least one sibling with a clinical diagnosis of ASD; (2) a caregiver or clinician indicated concerns about the child's development of social interaction, play, or other behaviors; and/or (3) the individual scored in the positive range on the M-CHAT. Exclusion criteria included major congenital or genetic disorders or diseases, or behavioral problems that would cause substantial additional stress for the family and/or the child during testing. Individuals with a previous diagnosis of ASD were included, but the examiner was not informed of the diagnosis.

Assessment Instruments and Protocols

The present study involves human participants and was reviewed and approved by the Institutional Review Board (IRB) of Massachusetts General Hospital (MGH, 2017P0000857). Written

informed consent to participate in this study was provided by the participant's legal guardian. RITA-T is a level-2 ASD screening test which includes 9 interactive games (phone block, phone tease, block vision, object constancy, color constancy, face vs. object, rapid joint attention, reaction to emotion, and self-recognition) that takes about 10 min to complete by trained and certified staff. The training and certification only need 3 h to finish and applies to entry level clinical and research staff. The RITA-T focuses on five major areas: joint attention, social awareness, human agency, self-recognition, and fundamental cognitive skills. The total score ranges from 0 to 30 (16). Responses were scored during test administration.

The ADOS-2 is considered the “gold standard” in diagnosis and is a semi-structured, standardized assessment of social interactions, language and communication, repetitive, restricted patterns of behavior and interest, and play and imagination (20–22). It contains five modules that are differentiated by participant's developmental and language levels (Module T, 1, 2, 3, and 4). Every ADOS-2 module ends with a diagnostic algorithm(s) that consists of selected items that have been chosen to maximize diagnostic sensitivity and specificity. In this study, the ADOS-2 was administered by professionally trained investigators, in consultation with a certified ADOS trainer, as needed. Following the standardized algorithm of corresponding modules, the composite score, SA, and RRB sub-scores were all recorded for each subject on a commercially available standard score booklet and scored right after the visit. The RITA-T and the ADOS-2 were both administered in one study visit in a randomized order. The ADOS-2 was administered by two different professionally trained administrators and the RITA-T was administered by three different professionally trained administrators. The overall evaluation time is around 1 h. Diagnostic and Statistical Manual, Fifth Edition (DSM-5) diagnostic criteria (23) were used by trained clinicians to diagnose ASD based on clinical observation, independent of the ADOS-2 and RITA-T scores.

Statistics

We first calculated descriptive statistics for demographic variables including age, sex, and race, and for RITA-T score, ADOS-2 total, and ADOS-2 sub-scores. We also examined distributions of ASD vs. non-ASD subjects categorized by both ADOS-2 and DSM-5.

Since scores in the ADOS-2 from module T and 1 & 2 are not directly comparable, ADOS-2 scores (total score, sub-score SA, and RRB scores) were converted to calibrated severity scores (CSS), including total CSS, SA CSS, and RRB CSS using a previously published and well-recognized conversion method, and were checked for normality via the Shapiro-Wilk test (24–27). Correlations of RITA-T scores with ADOS-2 CSS and the sub-score CSSs were calculated using linear regression via the Kendall rank correlation and the correlation coefficients were compared across age, race, and sex groups.

Receiver operating characteristic (ROC) curve analysis was performed to validate the predictive performance of RITA-T scores by determining its optimal cut-off score compared with that of ADOS-2 and DSM-5. The ROC curves were

TABLE 1 | Demographic features of the participants by age, sex, and race.

Age		18–36 months		37–84 months		Total	
Sex	Male	9	64%	14	67%	23	66%
	Female	5	36%	7	33%	12	34%
Race	White	5	36%	10	48%	15	43%
	Asian	3	21%	8	38%	11	31%
	Others	6	42%	3	15%	9	26%
Total		14	40%	21	60%	35	100%

also compared between age groups (18~36 mo.; 37–84 mo.). Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the RITA-T score in predicting the final diagnosis were computed for each age group and the combined group, and the cut-off scores with the optimal sensitivity and specificity were determined for the entire group and for each age group.

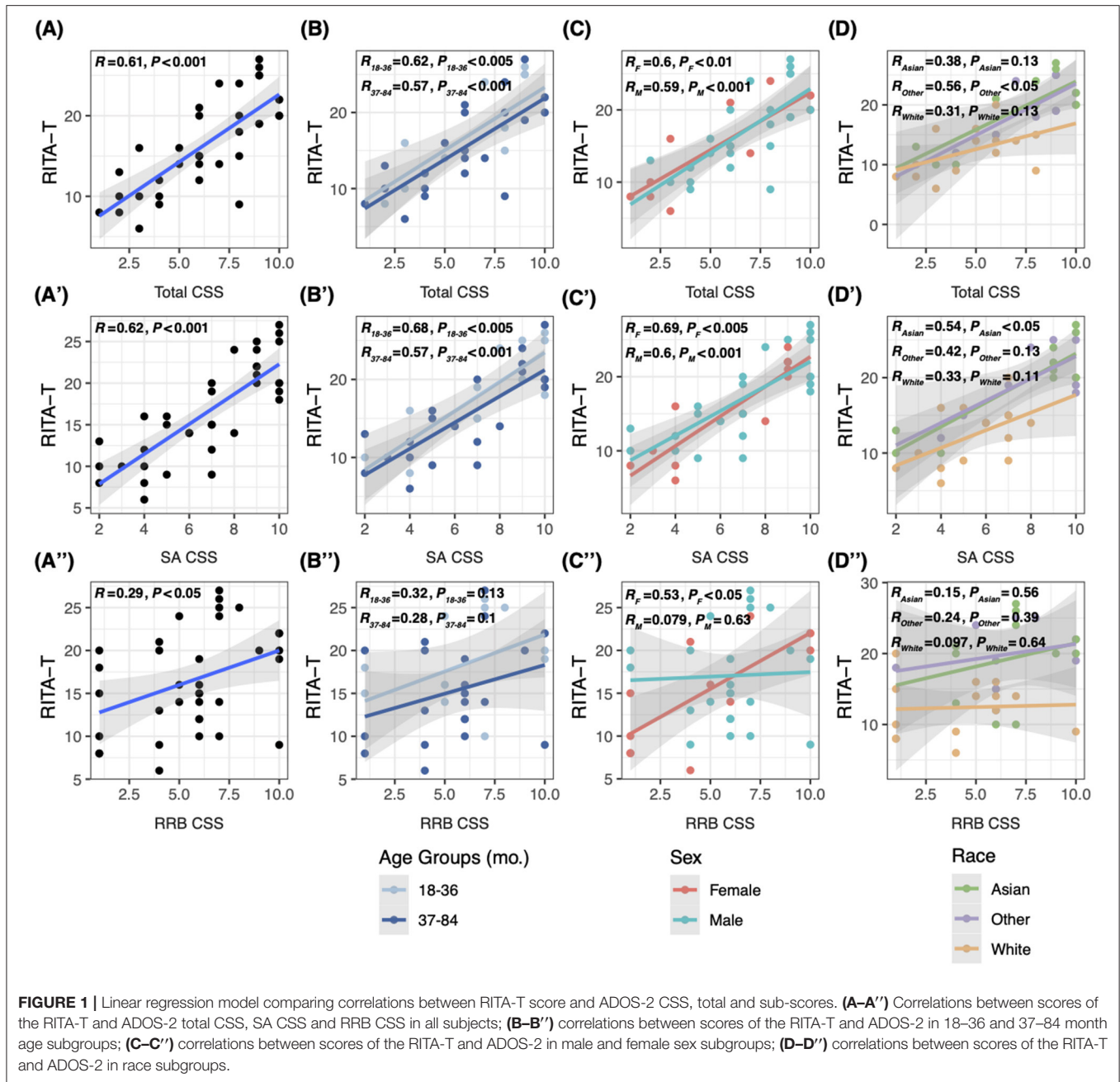
RESULTS

Demographic Features and ADOS-2 Score Compositions of the Participants

Table 1 shows that 40% of the subjects were 18–36 months old and 60% were 37–84 months old; among subjects aged 37–84 months old, 13 subjects were of ages 3.08–5 years old and 8 subjects were of ages 5–7 years old. Sixty-six percent were male and 34% were female. Thirty-one percent were Asian, 43% were White and 26% were in the others category. Among all subjects, 29 out of 35 (82.8%) were classified with ASD by ADOS-2, and 25 out of 35 (71.4%) were diagnosed with ASD using DSM-5 diagnostic criteria. Of those 18–36 months old, 11 out of 14 (78.5%) were classified with ASD by ADOS-2 while in the 37 to 84-month-old group, 18 out of 21 (85.7%) were classified with ASD by the ADOS-2 and 15 out of 21 (71.4%) were diagnosed using DSM-5. Subjects classified with ASD the by the ADOS-2 included 9 out 11 Asians (81.8%), 11 out of 15 Whites (83.3%), and 9 out of 9 others (100%). They also included 21 out 23 males (91.3%) and 8 out of 12 females (66.6%). Among all the participants in this study, 12 of them had a sibling with ASD, 9 of them had positive M-CHAT test, and 33 of them have had concerns from caregivers.

Correlation of RITA-T and ADOS-2 Total Scores/Sub-scores by Age, Race, and Sex

Linear regression was used to evaluate correlations of RITA-T score with ADOS-2 total CSS, sub-scores SA CSS and RRB CSS on all subjects and in both age, race, and sex groups. The RITA-T was found to be significantly positively correlated with ADOS-2 total CSS ($P < 0.001$, $R = 0.61$), SA CSS ($P < 0.001$, $R = 0.62$), and RRB CSS ($P < 0.05$, $R = 0.29$, **Figures 1A–A''**). Additionally, the RITA-T was significantly positively correlated with total CSS ($P_{18-36} < 0.005$, $R_{18-36} = 0.52$, $P_{37-84} < 0.001$, $R_{37-84} = 0.57$) and SA CSS ($P_{18-36} < 0.005$, $R_{18-36} = 0.68$, $P_{37-84} < 0.001$, $R_{37-84} = 0.57$), while RRB ($P_{18-36} = 0.13$,

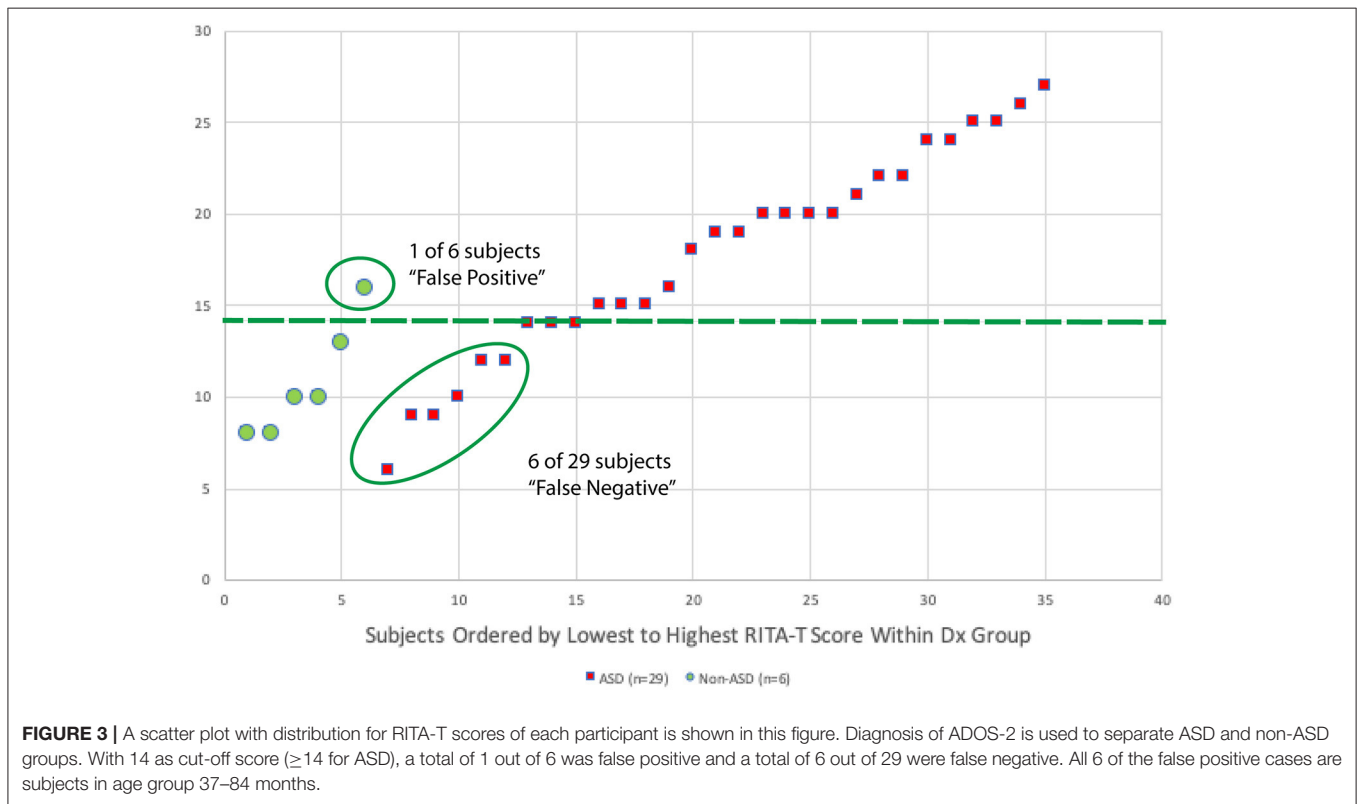
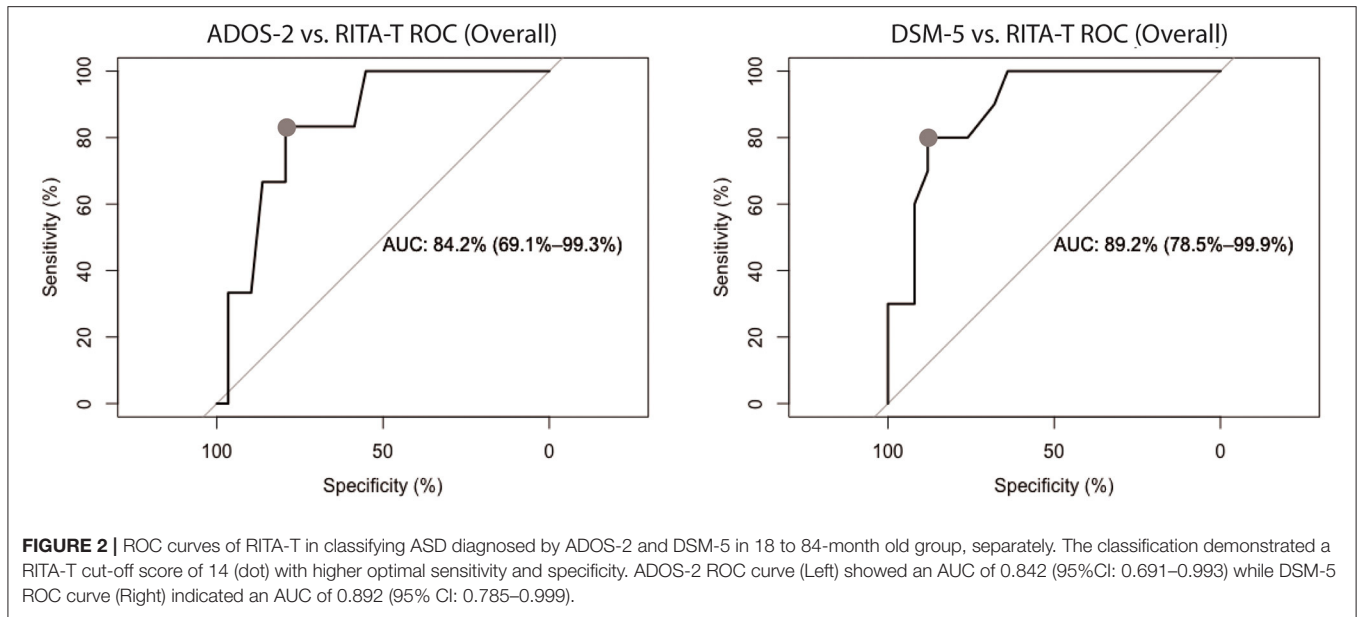


$R_{18-36} = 0.32$, $P_{37-84} = 0.01$, $R_{37-84} = 0.28$) did not show any significant correlation (Figures 1B–B''). For race subgroups, RITA-T scores are significantly positively correlated with ADOS-2 total CSS among the minority (“Other”) subgroup ($P < 0.05$, $R = 0.56$, Figure 1D) and with SA CSS in the Asian subgroup ($P < 0.05$, $R = 0.54$, Figure 1D'), while other subgroups based on ethnicity did not show statistical significance in correlations with ADOS-2 total CSS, SA CSS, and RRB CSS (Figures 1D–D''). In subgroups based on sex, both males and females demonstrated significant positive correlations with total CSS ($P_M < 0.001$, $R_M = 0.59$, $P_F < 0.01$, $R_F = 0.6$, Figure 1C) and SA CSS ($P < 0.005$, $R = 0.69$, Figure 1C'). However, correlations between RITA-T

scores and RRB CSS are significant among females ($P < 0.05$, $R = 0.53$, Figure 1C''), but not males ($P = 0.63$, $R = 0.079$).

Correlation of RITA-T and ADOS-2/DSM-5 Classification/Diagnosis

Among all 35 subjects, 6 (17.1%) are not classified with ASD by ADOS-2 and 10 (28.6%) are not diagnosed/classified with ASD by DSM-5. RITA-T’s sensitivity, specificity, PPV and NPV compared across ADOS-2 and DSM-5 for children 18–84 months old are shown in Supplementary Table 1. The cut-off score with most optimal sensitivity and specificity combination is 14 in this age group for ADOS-2 (sensitivity 81%, specificity 89%, PPV 96%,



NPV 57%) and is also 14 for DSM-5 (sensitivity 74%, specificity 50%, PPV 74%, NPV 50%).

ROC curves for comparing RITA-T with ADOS-2 and RITA-T with DSM-5 are shown separately in **Figure 2**. The ADOS ROC curve showed an area under the curve (AUC) of 0.842 (95%CI: 0.691–0.993), while DSM-5 ROC curve indicated an AUC of 0.892 (95% CI: 0.785–0.999).

A scatter plot with the distribution of RITA-T scores is shown in **Figure 3**. Classification of ADOS-2 was used to separate ASD and non-ASD groups. With 14 as the cut-off score (≥ 14 for ASD), a total of only 1 out of 6 was false positive and a total of 6 out of 29 were false negative. All 6 of the false negative cases are subjects were in the 37 to 84-month-old age group.

TABLE 2 | Comparison of ADOS-2 CSS and RITA-T scores between true positive subjects and false negative subjects.

Groups	RITA-T	ADOS-2 CSS		
	Total	Total	SA	RRB
True positive (mean ± SD)	19.78 ± 4.16	7.83 ± 1.67	8.26 ± 1.68	6.22 ± 2.75
False negative (mean ± SD)	9.67 ± 2.25	4.83 ± 1.83	5.17 ± 1.47	6.00 ± 2.19
T-Test (<i>p</i> -value)	<0.0001	0.004	0.0008	0.429

The true positive and false negative analyses (**Table 2**) showed significant differences in RITA-T ($P < 0.0001$), ADOS-2 total score ($P < 0.005$), and ADOS-2 sub-score SA ($P < 0.001$), but did not show a significant difference in the ADOS-2 sub-score RRB ($P = 0.429$).

ROC analysis was performed twice using ADOS-2 and DSM-5 diagnosis/classification independently as golden standard to determine the RITA-T cut-off score of 14. RITA-T's sensitivity, specificity, PPV, and NPV compared cross ADOS-2 and DSM-5 for those 18–84 months old is shown in **Supplementary Table 1**. The cut-off score with most optimal sensitivity and specificity combination is 14 in this age group for ADOS-2 (sensitivity 81%, specificity 89%, PPV 96%, NPV 57%) and 14 for DSM-5 (sensitivity 74%, specificity 50%, PPV 74%, NPV 50%).

Based on the distribution of RITA-T scores for each participant, classification of ADOS-2 is used to classify ASD and non-ASD groups. With 14 as cut-off score (≥ 14 for ASD), a total of 1 out of 6 was false positive and a total of 6 out of 29 were false negative. All 6 of the false positive cases are subjects in age group 37–84 months.

As presented in **Table 2** for true positive subjects and false negative subjects, mean and standard variance for ADOS-2 CSS and RITA-T score are calculated. *T*-test is performed to compare the difference between the two groups. True positive and false negative groups have significant difference in ADOS-2 SA CSS, total CSS, and RITA-T score, but have no significant difference in ADOS-2 RRB CSS.

DISCUSSION

This study compared performance on the RITA-T and ADOS-2 to further validate the RITA-T as a screening tool for ASD. ADOS-2 requires comprehensive training and strict certification, and DSM-5 diagnosis could only be done by licensed professionals or specialists. The limited resources and availabilities of these providers are unable to meet the rapidly increasing demands, which could only delay the diagnosis. Found to be correlated with ADOS-2 scores, the RITA-T has its value in further stratifying those individuals at high-risk of ASD, which could allow them to initiate the service before a formal diagnosis could be achieved. Our findings demonstrated that RITA-T scores correlated significantly with the ADOS-2 total score and both the SA and RRB sub-scores in the study subjects, but not

with the RRB sub-score in age-, sex-, and race-based subgroups. The RITA-T also showed good sensitivity (81%) and specificity (89%) with a cutoff score of 14. When the RITA-T is compared to the secondary outcome of DSM-5 at the same cut off score of 14 the sensitivity is 74% and the specificity is 50%. These results add to the validation of the RITA-T as a rapid screening measure for ASD in children.

Our results showed significant associations between the RITA-T and the ADOS-2 in both the 18–36 and the 37 to 84-month-old groups, which indicates the possibility of extending the age range in which these relationships have been studied (16, 17). This may also extend the validity of the RITA-T to an older age group, which is notable because despite ongoing efforts at early detection, ASD diagnosis is usually delayed to a median of 51 months old (7, 18, 19). Furthermore, one recent study showed, that only 44% of children who were diagnosed with ASD received an evaluation by 36 months of age (18). Despite on-going efforts in reducing the mean age of diagnosis, we have observed minimal changes in the average age of diagnosis. By extending the RITA-T age range, we are hopeful that it can further cover these older children who may not have gotten a diagnosis at a younger age. Importantly, RITA-T is still suitable for those chronologically older children with younger development age such as all subjects aged <4 years old in this study. The benefit of using the RITA-T is that anyone can be trained to administer it, including teachers and healthcare provider; this further suggests that a primary healthcare provider could administer it during a routine visit if warranted, as the training time required for its certified administration is only 3 h. This could allow an easy access to bridge for earlier diagnoses and interventions, which has been previously shown to improve outcomes of ASD individuals (7, 18, 19). Therefore, the RITA-T may be a particularly useful screening measure for children in the 37 to 84-month-old age range.

We also found similar relationships between RITA-T and ADOS-2 scores in different ethnic groups. Notably, correlations between RITA-T and ADOS-2 total was uniquely significant in the minority group (denoted as “Other”) whereas SA sub-scores were uniquely significant in the Asian subgroup but not in the other nor White groups. This is notable because, in addition to delays in ASD noted above, additional delays in diagnosis are common in minority groups (12–15). Several factors likely contribute to these delays (e.g., low socioeconomic status, cultural/language barriers, and shortage of accessible therapeutic resources or diagnostic testing, including the ADOS). Moreover, foreign language versions of ADOS-2, including Chinese, have not been validated formally (28). These findings suggest the RITA-T may be particularly useful in minority groups.

Consistent with our findings concerning age and race, our findings concerning sex showed that both male and female subjects showed strong correlations between RITA-T scores and the ADOS-2 total score and the SA sub-scores. One sex-related difference did emerge. While male subjects did not show significant correlations with the ADOS-2 RRB sub-scores, female subjects did. Since the RITA-T does not include RRB assessment, this may reflect a false positive finding. However, the finding may also provide insight into how sex-related differences in autism overlap with the assessment of autism. First, the male to

female ratio of ASD is up to 4.3:1 (18, 19). It was about 2:1 in the current study, which makes the female group smaller ($n = 12$), less reliable and more subject to the influence of extreme scores. Second, females may, on average, show higher social and linguistic skills and fewer atypical repetitive behaviors (29). RRB behaviors are typically more severe in males than females (19, 30, 31), though at least one study reported that the differences might be small (32). Related to this point, sex differences have been reported in subcategories of RRB domains, with female showing more insistence on sameness and compulsivity, and males showing more stereotyped and restrictive behaviors (33). It is thus possible in the present study that mild, more restricted or atypical RRB behaviors in females showed enough overlap with non-RRB ADOS-2 items to contribute to a significant RITA-T—ADOS-2 RRB correlation, even in the absence of actual RITA-T RRB items.

Several limitations should be noted. First, the number of subjects in the study was modest ($n = 35$), which limits the effects we were able to detect. Future investigation will require larger sample size to officially validate these findings. Related to this point, and second, the small number of subjects allowed for the specific assessment of white and Asian subjects, but African-American, Hispanic, and multiracial subjects were combined in an “other” category. Given the potential effectiveness of the RITA-T in minority subjects, it will be important to recruit enough African-American and Hispanic subjects in future studies to study them in separate groups. Third, although the RITA-T was well-aligned with the ADOS-2 total score and SA sub-scores, it did not contain items related to the RRB sub-score and thus, generally did not correlate significantly with that sub-score. This reflects a design weakness in the RITA-T that will hopefully be corrected in future revisions. Fourth, the same group of examiners administered the screens and diagnostic measures and were not blinded to the findings. Consequently, we cannot rule out the possibility that rating and diagnostic decisions were fully independent of each other, although the scoring was usually done afterwards by the main examiner with inputs from other co-examiners and observers of the tests, review of video and discussion to resolve scoring agreements. Fifth, the RITA-T was administered before the ADOS-2, except when siblings came in at the same time (roughly 30% of the time), when the order was reversed for just one of them. It is thus possible that performance on the ADOS-2 was influenced by performance on the RITA-T (although this influence could be small because the content of the tests do not overlap and the RITA-T is a short test). Future investigations will counterbalance test order and apply the order consistently to avoid performance variability related to order of administration. Sixth, we did not correct statistically for familial influences on performance, though its low frequency reduced the likelihood it altered outcome measures.

Despite the limitations, these findings provide encouraging evidence for validation of the RITA-T as a rapid and useful screening tool for ASD in both toddlers and those chronologically older children with developmental ages <4 years old. Compared to existing, less specific measures such

as the M-CHAT, the RITA-T may provide a more useful way to stratify risk for ASD and can thus refer children to the most appropriate diagnostic procedures and bridge the subsequent interventions.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author, X-JK.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Massachusetts General Hospital IRB. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

X-JK was the PI of the study, conceived the concept, designed and led the study and data analysis, drafted the first version, and revised the manuscript. X-JK, HS, SL, and MK contributed to data collections, data management, and analysis. RT contributed to data analysis, create tables, figures, and legend, and the literature search. HS contributed to revision of references, format, manuscript revision, and the literature search. AL contributed to literature search and measurement writing. WS contributed to data interpretation and writing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by internal funding #230361 and #233263 of Massachusetts General Hospital awarded to X-JK.

ACKNOWLEDGMENTS

We highly appreciate Dr. Ellen Hanson for training and critical instruction and consultation of ADOS-2 and Dr. Bruce Rosen for his great support and mentorship. We highly appreciate Dr. Roula Choueiri for her kind instruction of RITA-T and comments about the manuscript. We would like to thank Dr. Yiqing Song and Mr. Kevin Liu for their critical review and suggestions of the statistical analysis and Dr. Cunjian Dong for his critical comments about the manuscript. We appreciate Dr. Jian Kong for his constructive suggestions about data interpretation. We also want to thank Mr. Bryan K. Wang for his English proof of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2021.737890/full#supplementary-material>

REFERENCES

- Dow D, Guthrie W, Stronach ST, Wetherby AM. Psychometric analysis of the Systematic Observation of Red Flags for autism spectrum disorder in toddlers. *Autism*. (2017) 21:301–9. doi: 10.1177/1362361316636760
- Kantzer A-K, Fernell E, Westerlund J, Hagberg B, Gillberg C, Miniscalco C. Young children who screen positive for autism: stability, change and “comorbidity” over two years. *Res Dev Disabil*. (2018) 72:297–307. doi: 10.1016/j.ridd.2016.10.004
- Pierce K, Gazestani VH, Bacon E, Barnes CC, Cha D, Nalabolu S, et al. Evaluation of the diagnostic stability of the early autism spectrum disorder phenotype in the general population starting at 12 months. *JAMA Pediatr*. (2019) 173:578–87. doi: 10.1001/jamapediatrics.2019.0624
- Raza S, Zwaigenbaum L, Sacrey L-AR, Bryson S, Brian J, Smith IM, et al. Brief report: evaluation of the short quantitative checklist for autism in toddlers (Q-CHAT-10) as a brief screen for autism spectrum disorder in a high-risk sibling cohort. *J Autism Dev Disord*. (2019) 49:2210–8. doi: 10.1007/s10803-019-03897-2
- Sacrey L-AR, Bryson S, Zwaigenbaum L, Brian J, Smith IM, Roberts W, et al. The Autism Parent Screen for Infants: predicting risk of autism spectrum disorder based on parent-reported behavior observed at 6–24 months of age. *Autism*. (2018) 22:322–34. doi: 10.1177/1362361316675120
- Towle PO, Patrick PA. Autism spectrum disorder screening instruments for very young children: a systematic review. *Autism Res Treat*. (2016) 2016:4624829. doi: 10.1155/2016/4624829
- Zwaigenbaum L, Penner M. Autism spectrum disorder: advances in diagnosis and evaluation. *BMJ*. (2018) 361:k1674. doi: 10.1136/bmj.k1674
- de Bildt A, Sytema S, Zander E, Bölte S, Sturm H, Yirmiya N, et al. Autism Diagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers: application in a non-US sample of 1,104 children. *J Autism Dev Disord*. (2015) 45:2076–91. doi: 10.1007/s10803-015-2372-2
- Rutter M, Le Couteur A, Lord C. *Autism Diagnostic Interview-Revised*. Vol. 29. Los Angeles, CA: Western Psychological Services (2003). p. 30.
- Monteiro SA, Dempsey J, Broton S, Berry L, Goin-Kochel RP, Voigt RG. Early intervention before autism diagnosis in children referred to a regional autism clinic. *J Dev Behav Pediatr*. (2016) 37:15–19. doi: 10.1097/DBP.0000000000000241
- Rotholz DA, Kinsman AM, Lacy KK, Charles J. Improving early identification and intervention for children at risk for autism spectrum disorder. *Pediatrics*. (2017) 139:e20161061. doi: 10.1542/peds.2016-1061
- Bilaver LA, Havlicek J. Racial and ethnic disparities in autism-related health and educational services. *J Dev Behav Pediatr*. (2019) 40:501–10. doi: 10.1097/DBP.0000000000000700
- Sritharan B, Koola MM. Barriers faced by immigrant families of children with autism: a program to address the challenges. *Asian J Psychiatr*. (2019) 39:53–57. doi: 10.1016/j.ajp.2018.11.017
- Stahmer AC, Vejnoska S, Iadarola S, Straiton D, Segovia FR, Luelmo P, et al. Correction to: caregiver voices: cross-cultural input on improving access to autism services. *J Racial Ethn Health Disparities*. (2019) 6:774. doi: 10.1007/s40615-019-00587-8
- Zelege WA, Hughes TL, Drozda N. Disparities in diagnosis and service access for minority children with ASD in the United States. *J Autism Dev Disord*. (2019) 49:4320–31. doi: 10.1007/s10803-019-04131-9
- Choueiri R, Wagner S. A new interactive screening test for autism spectrum disorders in toddlers. *J Pediatr*. (2015) 167:460–6. doi: 10.1016/j.jpeds.2015.05.029
- Lemay J-F, Amin P, Langenberger S, McLeod S. Experience with the rapid interactive test for autism in toddlers in an autism spectrum disorder diagnostic clinic. *J Dev Behav Pediatr*. (2020) 41:95–103. doi: 10.1097/DBP.0000000000000730
- Maenner MJ. Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2016. *MMWR Surveill Summ*. (2020) 69:1–12. doi: 10.15585/mmwr.ss6904a1
- Zablotsky B, Black LI, Maenner MJ, Schieve LA, Blumberg SJ. Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 National Health Interview Survey. *Natl Health Stat Rep*. (2015) 13:1–20.
- Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop S. (*ADOSTM-2*) *Autism Diagnostic Observation ScheduleTM, Second Edition*. WPS (2012). Available online at: [http://www.wpspublish.com/ados-2-autism-diagnostic-observation-schedule-second-edition?utm_term=%2Bados%20%2B2&utm_campaign=Search\\$+%7C\\$+%\\$Autism\\$+\\$\\$+Champion&utm_source=adwords&utm_medium=ppc&hsa_net=adwords&hsa_tgt=aud-325280925571:kw-340728300176&hsa_ad=394983746385&hsa_acc=6243382947&hsa_grp=61741303410&hsa_mt=b&hsa_cam=1698309095&hsa_kw=%2Bados%20%2B2&hsa_ver=3&hsa_src=g&gclid=CjwKCAjwjQ5BRAPEiwAJlBuBcUZMhr-MOMJEdrt2Psi_cOgBFgnaX4w2h8Fqw8Z1ZvON836DbON4hoCy1wQAvD_BwE](http://www.wpspublish.com/ados-2-autism-diagnostic-observation-schedule-second-edition?utm_term=%2Bados%20%2B2&utm_campaign=Search$+%7C$+%$Autism$+$$+Champion&utm_source=adwords&utm_medium=ppc&hsa_net=adwords&hsa_tgt=aud-325280925571:kw-340728300176&hsa_ad=394983746385&hsa_acc=6243382947&hsa_grp=61741303410&hsa_mt=b&hsa_cam=1698309095&hsa_kw=%2Bados%20%2B2&hsa_ver=3&hsa_src=g&gclid=CjwKCAjwjQ5BRAPEiwAJlBuBcUZMhr-MOMJEdrt2Psi_cOgBFgnaX4w2h8Fqw8Z1ZvON836DbON4hoCy1wQAvD_BwE) (accessed August 4, 2020).
- Falkmer T, Anderson K, Falkmer M, Horlin C. Diagnostic procedures in autism spectrum disorders: a systematic literature review. *Eur Child Adolesc Psychiatry*. (2013) 22:329–40. doi: 10.1007/s00787-013-0375-0
- Luyster R, Gotham K, Guthrie W, Coffing M, Petrak R, Pierce K, et al. The autism diagnostic observation schedule—toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. *J Autism Dev Disord*. (2009) 39:1305–20. doi: 10.1007/s10803-009-0746-z
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. 5th ed. Washington, DC: American Psychiatric Association (2013).
- Esler AN, Bal VH, Guthrie W, Wetherby A, Weismer SE, Lord C. The autism diagnostic observation schedule, toddler module: standardized severity scores. *J Autism Dev Disord*. (2015) 45:2704–20. doi: 10.1007/s10803-015-2432-7
- Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord*. (2009) 39:693–705. doi: 10.1007/s10803-008-0674-3
- Hedley D, Nevill R, Uljarević M, Butter E, Mulick JA. ADOS-2 Toddler and Module 1 standardized severity scores as used by community practitioners. *Res Autism Spectr Dis*. (2016) 32:84–95. doi: 10.1016/j.rasd.2016.09.005
- Hus V, Gotham K, Lord C. Standardizing ADOS domain scores: separating severity of social affect and restricted and repetitive behaviors. *J Autism Dev Disord*. (2014) 44:2400–12. doi: 10.1007/s10803-012-1719-1
- Sun X, Allison C, Auyeung B, Matthews FE, Zhang Z, Baron-Cohen S, et al. Comparison between a Mandarin Chinese version of the childhood autism spectrum test and the Clancy Autism Behaviour Scale in mainland China. *Res Dev Disabil*. (2014) 35:1599–608. doi: 10.1016/j.ridd.2014.02.005
- Lai M-C, Szatmari P. Sex and gender impacts on the behavioural presentation and recognition of autism. *Curr Opin Psychiatry*. (2020) 33:117–23. doi: 10.1097/YCO.0000000000000575
- Messinger DS, Young GS, Webb SJ, Ozonoff S, Bryson SE, Carter A, et al. Early sex differences are not autism-specific: a Baby Siblings Research Consortium (BSRC) study. *Molecular Autism*. (2015) 6:32. doi: 10.1186/s13229-015-0027-y
- Supekar K, Menon V. Sex differences in structural organization of motor systems and their dissociable links with repetitive/restricted behaviors in children with autism. *Mol Autism*. (2015) 6:50.
- Knutsen J, Crossman M, Perrin J, Shui A, Kuhlthau K. Sex differences in restricted repetitive behaviors and interests in children with autism spectrum disorder: an Autism Treatment Network study. *Autism*. (2019) 23:858–68. doi: 10.1177/1362361318786490
- Antezana L, Factor RS, Condy EE, Strega MV, Scarpa A, Richey JA. Gender differences in restricted and repetitive behaviors and interests

in youth with autism. *Autism Res.* (2019) 12:274–83. doi: 10.1002/aur.2049

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kong, Sherman, Tian, Koh, Liu, Li and Stone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.