Ma *et al. BMC Digital Health (2023) 1:44*
https://doi.org/10.1186/s44247-023-00047-8

BMC Digital Health

RESEARCH

Open Access

A novel telehealth tool using a snack activity to identify autism spectrum disorder

Zenghui Ma[1], Yan Jin[1], Ruoying He[1], Qinyi Liu[2], Xing Su[2], Jialu Chen[2], Disha Xu[1], Jianhong Cheng[1], Tiantian Zheng[1], Yanqing Guo[2*], Xue Li[2*] and Jing Liu[2*]

Abstract

Background The COVID-19 pandemic has caused an unprecedented need for accessible health care services and significantly accelerated the development processes of telehealth tools for autism spectrum disorder (ASD) early screening and diagnosis. This study aimed to examine the feasibility and utility of a time-efficient telehealth tool combining a structured snack time assessment activity and a novel behaviour coding scheme for identifying ASD.

Methods A total of 134 1–6-year-old individuals with ASD (age in months: mean = 51.3, SD = 13.1) and 134 age- and sex-matched typically developing individuals (TD) (age in months: mean = 54, SD = 9.44) completed a 1-min snack time interaction assessment with examiners. The recorded videos were then coded by trained coders for 17 ASD-related behaviours; the beginning and end points and the form and function of each behaviour were recorded, which took 10–15 min. Coded details were transformed into 62 indicators representing the count, duration, rate, and proportion of those behaviours.

Results Twenty indicators with good reliability were selected for group difference, univariate and multivariate analyses. Fifteen behaviour indicators differed significantly between the ASD and TD groups and remained significant after Bonferroni correction, including the children's response to the examiner's initiation, eye gaze, pointing, facial expressions, vocalization and verbalization, and giving behaviours. Five indicators were included in the final prediction model: total counts of eye gaze, counts of standard pointing divided by the total counts of pointing, counts of appropriate facial expressions, counts of socially oriented vocalizations and verbalizations divided by the total counts of vocalizations and verbalizations, and counts of children using giving behaviours to respond to the examiner's initiations divided by the total counts of the examiner's initiation of snack requisitions. The ROC curve revealed a good prediction performance with an area under the curve (AUC) of 0.955, a sensitivity of 92.5% and a specificity of 84.3%.

Conclusion Our results suggest that the snack activity-based ASD telehealth approach shows promise in primary health care settings for early ASD screening.

Keywords Autism spectrum disorder, Telehealth, Early screening, Diagnosis, Snack time

*Correspondence:
Yanqing Guo
guoyanqing1946@bjmu.edu.cn
Xue Li
lixue@bjmu.edu.cn
Jing Liu
ljyuch@bjmu.edu.cn
Full list of author information is available at the end of the article

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by core social communication features and restricted, repetitive sensory–motor behaviours [1]. The prevalence of ASD has increased in recent decades and is currently estimated to be 1/44 in developed countries [2]. ASD substantially burdens individuals, families, and society [3]. Early identification enables children with ASD to access early intervention, which is crucial for them to obtain optimal long-term outcomes [4, 5]. However, there are often significant delays between parents' initial concerns about their child's development and a formal autism spectrum disorder (ASD) diagnosis. While the diagnosis of ASD becomes stable as early as 14 months old [6], the global mean age of ASD diagnosis is over 40 months [7]. One major barrier to diagnostic delay is related to the inadequate availability and sensitivity of early screening tools for ASD, leading to limited usage of these tools in primary health care settings in low- and middle-income countries (LMICs) [8]. Therefore, there is a pressing need to develop cost-effective clinical ASD screening tools, especially for use in under-resourced settings in LMICs.

One promising avenue for addressing the urgent global need for the early identification of ASD is employing telehealth evaluations to shorten the waiting time for clinical appointments and remove geographic barriers. The COVID-19 pandemic has sparked enhanced enthusiasm for developing flexible remote ASD screening and diagnostic tools, adding to the existing prepandemic telehealth repertoire [9]. A recent review summarized 10 screening and 7 diagnostic approaches using telemedicine technology for ASD [10]. In addition to some online questionnaire applications [11, 12], video-based approaches are used in three different settings: 1) a trained provider or examiner interacts with the child in one location with a remote clinician observing and coding in real-time [13, 14]; 2) the caregivers are trained and complete the assessment activities with the child at home and upload the recorded videos for further coding [15, 16]; and 3) clinicians make video appointments with the family to conduct diagnostic evaluations remotely [17, 18]. Although these telehealth approaches for ASD screening and diagnosis seem promising for extensive use in clinical and research services, some concerns arise when considering whether these tools will significantly address time delays in accessing ASD evaluation and diagnostic services, especially for families in LMICs.

The first concern regarding existing ASD telemedicine screening and diagnostic tools is whether these approaches can alleviate the problem of limited ASD specialist capacity due to ever-increasing needs. Experienced ASD specialists are essential for telemedicine tools that require remote real-time communications to make meaningful and accurate observations and ratings for children's behaviours [14, 17], which is even more challenging than in-person evaluations [19]. Other approaches that involve coding children's behaviours in recorded videos also have stringent educational and clinical eligibility requirements for coders [15]. These requirements show the equal importance of specialized training in ASD telehealth applications. Some telehealth tools may not be applicable in LMICs due to the requirements of ADOS-2 or ADI-R administrators [13, 18]. The time cost is also an issue; for example, providers reported that the duration of a telemedicine visit could be as long as 3 h for a family [17], and some applications need coders to watch a 1-h home video to give feedback [20]. Considering these limitations, it is questionable whether the existing telemedicine approaches can effectively shorten the waiting time for families to access formal ASD evaluations and diagnoses. The second concern is about caregiver-mediated assessment, either in real-time or in recorded videos. While this approach resolves geographic barriers, other related problems, including enormous efforts in training caregivers, possible low fidelity in assessment administration and recorded video, potential practice effects for children, and extra tasks for caregivers, may discourage the expanded use of caregiver-mediated telehealth tools [21, 22]. The final concern is considerations for data structure in the era of big data. Almost all behaviour rating methods in existing ASD telehealth tools use Likert scales, relying on the rater's overall observation and impression of the child to form scores. However, this type of categorical dataset does not allow the machine to learn what a targeted autistic behaviour is and achieve automatic behaviour recognition in the future with accumulating database, which would help mitigate the ever-increasing workload for clinician [23].

Given the concerns regarding existing telehealth approaches for ASD screening and diagnosis, it is important to find alternatives to facilitate clinicians in accurately diagnosing ASD with less time and effort in the consulting room, especially in LMICs. Training and supervising practitioners with less experience in administering assessments and coding children's behaviours is an option. Structured and short assessments are more friendly to examiners. Furthermore, training local practitioners instead of caregivers can address geographic barriers and prevent other associated problems. Regarding behaviour measurement, coding behaviours instead of symptoms may lower the eligibility requirements for coders and provide a sound feeding database for automatic behaviour recognition through machine learning [24, 25]. To enable children at risk of ASD to present a wide range of autistic behaviours within a short time with an examiner, the most common daily activities must be used for social interaction contexts, such as snack time and playtime. Since children typically show strong motivations for their favourite food, snack time is a good opportunity to observe children's social interaction, communication behaviours towards snack holders, and potential sensory and motor behaviours and stereotypes [26, 27].

At snack time, when presented with choices of snacks, children are motivated to spontaneously display eye contact, vocalization, pointing, facial expressions, and other requesting behaviours as communication methods to reach out for the preferred food [27]. Snack time also provides a good opportunity to observe children's giving and sharing behaviours under social pressure, such as when the examiner shows interest in the children's snack. Children's responses to the examiner's social initiations are also indicators of reciprocal social interaction behaviours [28]. Several validated ASD screening and diagnostic tools incorporate snack time activities in their assessments with different administration instructions, such as ADOS-2 [29], the Screening Tool for Autism in Two-Year-Olds (STAT) [30], and a telehealth diagnostic application, the Naturalistic Observation Diagnostic Assessment (NODA) [16].

In the current study, a set of structured snack time activities with a novel behaviour coding scheme was developed as a telehealth approach for identifying children with ASD aged 1–6 years old. The snack time activities can be administered within 1 min by an examiner with relatively limited clinical training, and the recorded video is coded by a coder, a process that takes 10–15 min. The new behaviour coding scheme enables the recording of the beginning and ending time points of each targeted behaviour on the video timeline and the form and function of the behaviour. The coded data are then converted into a numerical dataset for further analysis. This study was conducted to establish a cost-effective method for the early identification of ASD, especially in LMICs, through the development of a time-saving telehealth approach.

The present study aimed to (1) explore the feasibility of an ASD telehealth approach combining structured snack time interaction activities administered by examiners with limited clinical training and behaviour coding by trained coders; (2) examine whether this approach can detect behavioural differences between children with ASD and typically developing children, such as eye-contact, vocalization, pointing, requesting, facial expressions, giving, responsiveness, atypical sensory and motor behaviours, and repetitive and repeated behaviours; (3) evaluate the reliability and validity of the tool; and (4) investigate a potential clinical prediction model to distinguish children aged 1–6 years with and without ASD using behavioural indicators from our approach.

Methods

Participants

A total of 436 participants were recruited from January 2022 to August 2022, including 153 individuals with ASD and 283 typically developing individuals (TD). Propensity score matching (PSM) was employed to obtain an age- and sex-matched sample, yielding a total of 268 included participants, 134 per group (see Table 1 for detailed demographic information before and after PSM). The ASD group was recruited from the outpatient clinic at Peking University Sixth Hospital and ASD early intervention centres, and the TD group was enrolled from kindergarten classes. The inclusion criteria for the ASD group were children who were 1–6 years old, had a diagnosis of ASD that was clinically established by experienced child and adolescent psychiatrists following the Diagnostic and Statistical Manual of Mental Disorders, 5th ed (DSM-5) [31], had no physical or neurological diseases, and did

Table 1 Demographics before and after propensity score matching

	Before PMS			After PMS		
	TD ($N=283$)	ASD ($N=153$)	t/χ^2 sig.†	TD ($N=134$)	ASD ($N=134$)	t/χ^2 sig.†
Age in month						
Mean (SD)	58.1 (10.7)	48.5 (14.5)	$<0.001^{***}$	54.0 (9.44)	51.3 (13.1)	=0.053
Median [Min, Max]	58.0 [33.0, 78.0]	48.0 [14.0, 82.0]		53.0 [34.0, 75.0]	50.0 [14.0, 82.0]	
Sex						
Female	130 (45.9%)	23 (15.0%)	$<0.001^{***}$	24 (17.9%)	23 (17.2%)	=1
Male	153 (54.1%)	130 (85.0%)		110 (82.1%)	111 (82.8%)	

† Significant difference ($p<0.05$*, $p<0.01$**, $p<0.001$***) between both groups by two-tailed T-test for continuous variables and by chi-square analysis for categorical variable

not take any psychotropic medication before participating in the study. The inclusion criteria for the TD group were children who were 1–6 years old and had no psychiatric, physical, or neurological disorders.

Data acquisition

Clinical diagnosis

All participants first underwent clinical assessments with senior child and adolescent psychiatrists to confirm their diagnoses using best-estimate clinical diagnoses according to the DSM-5.

Snack time administration and video recording

The snack time assessment scheme was developed by three senior child and adolescent psychiatrists with over 14 years of clinical experience (J.L., Y.G., and X.L.) through clinical experiences and a thorough literature review.

Each participant completed a 1-min standard snack time assessment activity administered by a trained examiner. The purpose of this activity is to create a familiar social context for children to show social interaction and communication behaviours, where structured initiations from examiners function as social cues. For video recording, examiners needed to start and stop recording before and after the assessment. The length of each video was required to be 50–80 s. The setting for the administration was a medium-sized, quiet, and clean room in the hospital, ASD early intervention centre, or kindergarten.

Behaviour coding and data conversion

The behaviour coding system was developed by three clinical experts (J.L., Y.G., and X.L.) through clinical experience and a thorough literature review. To develop the coding scheme, 17 behaviours were first identified as ASD-related behaviours that children might display at snack time based on the three clinicians' clinical experiences and the literature review (see Table S1 for behaviour coding list). The 17 behaviours covered social interaction and communication behaviours and restricted and repetitive behaviours (RRB). A behaviour coding plan was then created to record the duration, form, and function of behaviours present in the video. The duration of the behaviour was recorded by marking the starting and ending points of the behaviour on the video timeline, and multiple behaviour cooccurrences were identified through timeline overlapping. The form and function of the behaviour were recorded to distinguish different behaviour features, for example, socially oriented verbalization and nonsocially oriented verbalization. The coding option of each behaviour was specified in the coding guide to make the coding process transparent and repeatable (see Table S1 for the behaviour coding guide). The 17 behaviours were then converted and expanded to 62 numerical behaviour indicators, including counts, duration, ratio, and rate (see Table S2 for a total of 62 behaviour indicators).

The trained coders coded the 17 behaviours in the snack time videos while blind to the children's diagnostic status. The coded data were then translated into behaviour indicators for the data analysis.

Data reliability

The data reliability of this study was threefold: test–retest reliability, inter-examiner reliability, and inter-coder reliability. Sixty participants were recruited for test–retest reliability, and the test–retest interval was 7–15 days (mean = 11). Sixty-one participants were included for inter-examiner reliability, and inter-examiner assessments were conducted on the same day. Thirty participants were randomly selected for inter-coder reliability by three independent coders. Test–retest reliability, inter-examiner reliability, and inter-coder reliability were calculated for 62 behaviour indicators using intraclass correlation coefficients (ICCs).

To prioritize the reliability of the statistical analysis results, only indicators with ICCs of test–retest reliability, inter-examiner reliability, and inter-coder reliability all greater than 0.4 (indicating good reliability) were included in the next statistical analyses [32].

Data analysis

R (version 4.2.1) and Python (version 3.10.6) were used for the statistical analysis. The Shapiro–Wilk test was used to test whether the variables were normally distributed. The behavioural indicators were first examined for the significance of group differences using the Mann–Whitney U test for nonnormal distribution or two-tailed t tests for normal distribution, and Bonferroni correction was used for multiple comparison correction. The behaviour indicators were also included in the univariate logistic regression analysis to determine how many behaviour indicators would be included in the stepwise logistic regression to construct the prediction model. The stepwise logistic regression method was also used to minimize the multicollinearity among variables [33]. The odds ratio (OR) and 95% confidence intervals (95% CIs) for each indicator selected in the final prediction model were further calculated. Discrimination of the prediction model was assessed using the area under the curve (AUC), and a calibration curve was drawn to estimate the consistency between the probabilities predicted by the model and the observed probabilities. In addition, the Hosmer–Lemeshow goodness-of-fit test was also employed to evaluate the predictive performance of the model. Internal validation was performed using the bootstrapping method with 1000 resamples.

Results

Indicator selection

Only indicators with ICCs of test–retest reliability, inter-examiner reliability, and inter-coder reliability greater than 0.4 were included in the next statistical analyses to prioritize the reliability of the statistical analysis results, leaving 22 behaviour indicators (see reliability results in Table S3). Two indicators (PT2: total counts of standard pointing with coordinated gaze to object; PT3: counts of standard pointing with coordinated gaze to object divided by the total counts of pointing with coordinated gaze to object) were added back because of good test–retest reliability and inter-examiner reliability but unavailable ICC for inter-examiner reliability. Logical repetition was checked for the rest of the indicators. Because the length of the video was relatively fixed, indicators of total counts were retained (EC1: total counts of eye gaze; EC5: counts of eye gaze coordinated with facial expression, vocalization, or gesture; VO1: total counts of vocalizations and verbalizations; VO4: counts of socially oriented vocalizations and verbalizations), and indicators of rate were excluded (EC2: counts of eye gaze divided by the video length; EC6: counts of eye gaze coordinated with facial expression, vocalization, or gesture divided by the video length; VO2: counts of vocalizations and verbalizations divided by the video length; VO5: counts of vocalizations and verbalizations with social intentions divided by the video length). Finally, 20 indicators were included in the statistical analyses (Table 2).

Significance tests of behaviour indicators between the ASD and TD groups

Significance tests between the ASD and TD groups showed that a total of 15 out of 20 behaviour indicators achieved a significance level of $p < 0.05$, and all passed Bonferroni correction (adjusted $p < 0.05$; adjusted p was calculated by multiplying the uncorrected p value by 20) (Table S2). The behaviour indicators that remained significant after Bonferroni correction were RE1 (proportion of times children respond to the examiner's initiation), EC1 (total counts of eye gaze), EC3 (total duration of eye gaze), EC5 (counts of eye gaze coordinated with facial expression, vocalization, or gesture), PT2 (total counts of standard pointing with coordinated gaze to object), PT3 (counts of standard pointing with coordinated gaze to object divided by the total counts of pointing with coordinated gaze to object), FE2 (total counts of appropriate facial expressions), FE3 (counts of appropriate facial expressions divided by the total counts of facial expressions), FE4 (counts of basic facial expressions), VO1 (total counts of vocalizations and verbalizations), VO4 (counts of socially oriented vocalizations and verbalizations), VO7 (counts of socially oriented vocalizations and verbalizations divided by the total counts of vocalizations and verbalizations), GV1 (total counts of giving behaviours), GV3 (counts of children using giving behaviours to respond to the examiner's initiations), and GV4 (counts of children using giving behaviours to respond to the examiner's initiations divided by the total counts of the examiner's initiation of snack requisitions).

Univariate and multivariate analysis

Univariate logistic regression analysis showed that 15 out of 20 behaviour indicators achieved a significance level of $p < 0.05$ (Table 2), and those indicators were included in the stepwise logistic regression to construct the prediction model. The final prediction model consisted of 5 behaviour indicators (Table 2 and Fig. 1): EC1 (total counts of eye gaze) (OR 0.784, 95% CI 0.684 ~ 0.885, $p < 0.001$), PT3 (counts of standard pointing with coordinated gaze to object divided by the total counts of pointing with coordinated gaze to object) (OR 0.012, 95% CI 0.001 ~ 0.072, $p < 0.001$), FE2 (total counts of appropriate facial expressions) (OR 0.675, 95% CI 0.466 ~ 0.92, $p = 0.023$), VO7 (counts of socially oriented vocalizations and verbalizations divided by the total counts of vocalizations and verbalizations) (OR 1.286, 95% CI 1.162 ~ 1.446, $p < 0.001$), and GV4 (counts of children using giving behaviours to respond to the examiner's initiations divided by the total counts of the examiner's initiation of snack requisitions) (OR 0.025, 95% CI 0.006 ~ 0.083, $p < 0.001$).

Construction and validation of a nomogram prediction model

The 5 behaviour indicators identified by stepwise logistic regression were used to construct a model for predicting ASD risk, and the model was displayed using a nomogram (Fig. 2). The ROC curve was used to evaluate the model's predictive performance, revealing a good prediction performance with an area under the curve (AUC) of 0.955 (95% CI 0.932–0.977) (Fig. 3). The optimal cut-off value of the ROC curve was 0.421, with a sensitivity of 92.5% and specificity of 86.6%. The Hosmer–Lemeshow goodness-of-fit test for the model was $\chi 2 = 4.996$, $p = 0.758$, indicating a good model fit. A calibration curve was drawn to test the model's validity, which showed that the model's calibration curve fit well with the ideal standard curve (Fig. 4).

Table 2 Univariate and multivariate analysis

	Behavior indicator	Definition	Calculate units	Univariate analysis			Multivariate regression		
				OR	95% CI	*P*-value	OR	95% CI	*P*-value
Examiner's initiation and children's response									
1	RE1	Proportion of times children respond to the examiner 's initiation	Proportion	0	0–0.001	0			
Eye gaze									
2	EC1	Total counts of eye gaze	Count	0.681	0.619–0.751	0	0.784	0.684 ~ 0.885	< 0.001
3	EC3	Total duration of eye gaze	Duration	0	0–0	0			
4	EC5	Counts of eye gaze coordinated with facial expression, vocalization, or gesture	Count	0.736	0.634–0.855	0			
Pointing with coordinated gaze to object									
5	PT2	Total counts of standard pointing with coordinated gaze to object (pointing to a specific direction using any finger, with flexion of the remaining fingers)	Count	0.184	0.12–0.282	0			
6	PT3	Counts of standard pointing with coordinated gaze to object divided by the total counts of pointing with coordinated gaze to object	Proportion	0.019	0.006–0.061	0	0.012	0.001 ~ 0.072	< 0.001
Facial expression									
7	FE2	Total counts of appropriate facial expressions	Count	0.584	0.467–0.73	0	0.675	0.466 ~ 0.92	0.023
8	FE3	Counts of appropriate facial expressions divided by the total counts of facial expressions	Proportion	0.303	0.179–0.512	0			
9	FE4	Counts of basic facial expressions	Count	0.657	0.541–0.798	0			
Vocalization and verbalization									
10	VO1	Total counts of vocalizations and verbalizations (excluding sneeze, cough, or other vocalizations caused by physiological reflexes)	Count	1.158	1.046–1.283	0.005			
11	VO4	Counts of socially oriented vocalizations and verbalizations	Count	1.158	1.046–1.283	0.005			
12	VO7	Counts of socially oriented vocalizations and verbalizations divided by the total counts of vocalizations and verbalizations	Proportion	1.111	1.054–1.171	0	1.286	1.162 ~ 1.446	< 0.001
13	VO8	Duration of socially oriented vocalizations and verbalizations divided by the total duration of vocalizations and verbalizations	Proportion	1.133	0.683–1.878	0.629			
14	VO9	Counts of vocalization and verbalizations using words	Count	1.042	0.938–1.157	0.442			
15	VO10	Counts of vocalizations and verbalizations using words divided by the total counts of vocalizations and verbalizations	Proportion	1.112	0.67–1.846	0.683			
Giving									
16	GV1	Total counts of giving behaviors	Count	0.313	0.242–0.405	0			
17	GV2	Counts of spontaneous giving behaviors	Count	0.649	0.4–1.053	0.08			
18	GV3	Counts of children using giving behaviors to respond to the examiner's initiations	Count	0.248	0.185–0.333	0			
19	GV4	Counts of children using giving behaviors to respond to the examiner's initiations divided by the total counts of the examiner's initiation of snack requisitions	Proportion	0.015	0.006–0.035	0	0.025	0.006 ~ 0.083	< 0.001
20	GV5	Counts of spontaneous giving behaviors divided by the total counts of giving behaviors	Proportion	0.66	0.149–2.915	0.583			

OR Odds ratio, *CI* Confidence Interval

Discussion

The current study supported the use of a time- and cost-efficient telehealth approach using snack time with a novel behaviour coding scheme to provide clinical practitioners with detailed information on children's behaviour to assist early ASD screening and diagnosis. The

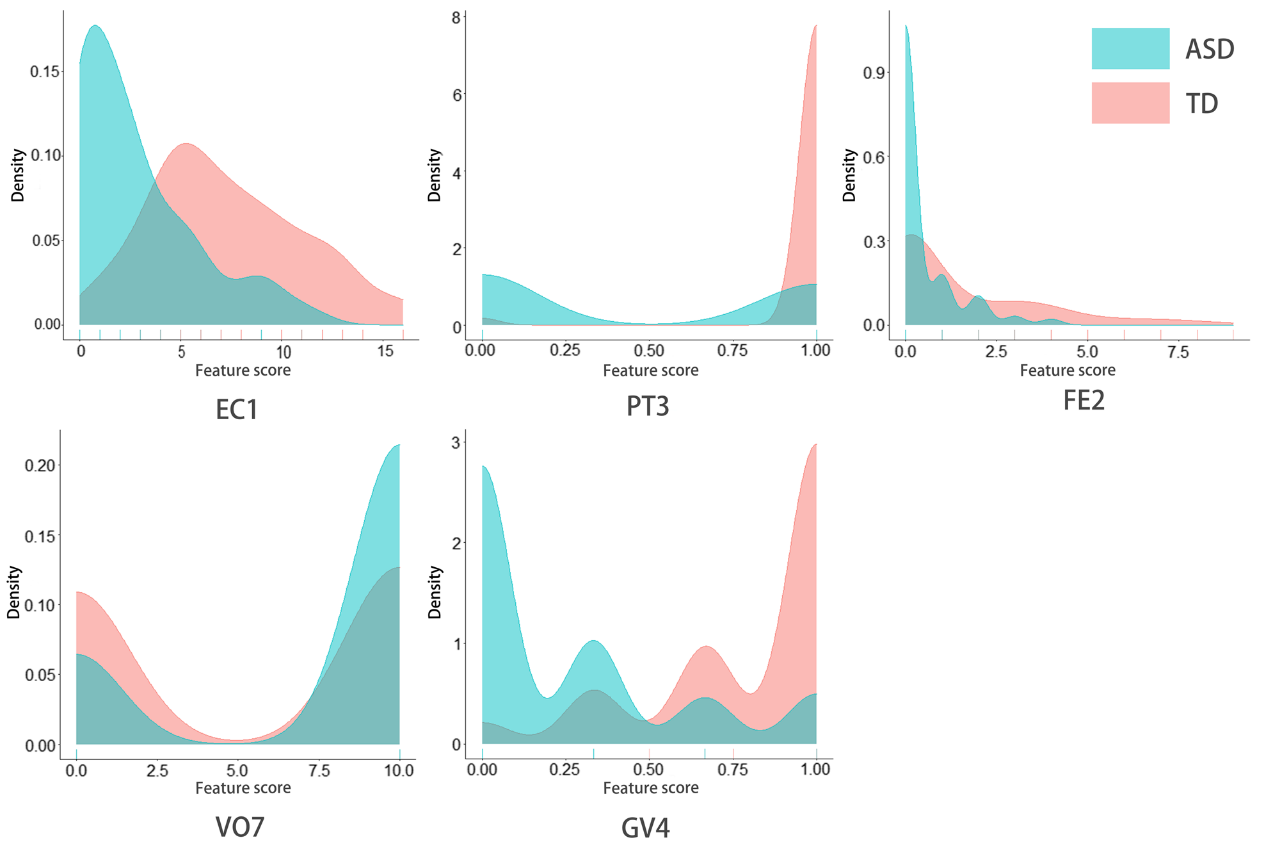

Fig. 1 Density plot for behavior indicators included in the final prediction model

telemedicine tool is timesaving because the examiner is only required to interact with the children for 1 min, and it takes only 10–15 min for the coder to code the 1-min snack time video. The approach is also specialist-saving because it requires limited specialty training for both examiners and coders. Although time-limited, the tool can detect significant behavioural differences between ASD and TD groups, such as children's response to the examiner's initiation, eye gaze, pointing, facial expressions, vocalization and verbalization, and giving behaviours. A clinical prediction model was built based on the behaviour indicators identified, and this model demonstrated good prediction performance and model fit. The sensitivity and specificity reached 92.5% and 86.6%, respectively. As shown here, this ASD telehealth tool shows good potential for future use in clinical practice, especially in resource-limited low- and middle-income countries and regions.

Behaviour and indicator analyses

Children with ASD showed higher severity than those in the TD group in six ASD-related social interaction and communication behaviours during the 1-min snack time interaction activity. Unlike the Likert rating scales used by most existing ASD screening and diagnostic tools, the behaviour coding system used in this study enabled the recording of beginning and ending points on the video timeline of each aimed behaviour. Therefore, counts, rates, durations, and proportions were calculated for these ASD-related behaviours, enabling a deeper understanding of ASD traits.

Significant differences in social interaction and communication behaviours were found between the ASD and TD groups using the developed tool. In the snack time activity setting, examiners were instructed to initiate three sets of social interactions, involving showing the snack, asking the children to point to the snack they want, and showing interest in the children's snack. Whether children responded to the examiner's initiations was coded for each round. Our results revealed that the ASD group showed impairments in those responding behaviours, such as how many times children responded to the examiner's initiation in total, how many times children pointed to the snack as the examiner asked, and how many times children gave the snack to the examiner as requested. In addition, social communication behaviours, such as eye contact, vocalization and verbalization, and facial expressions, were found to be deficient in the ASD groups.

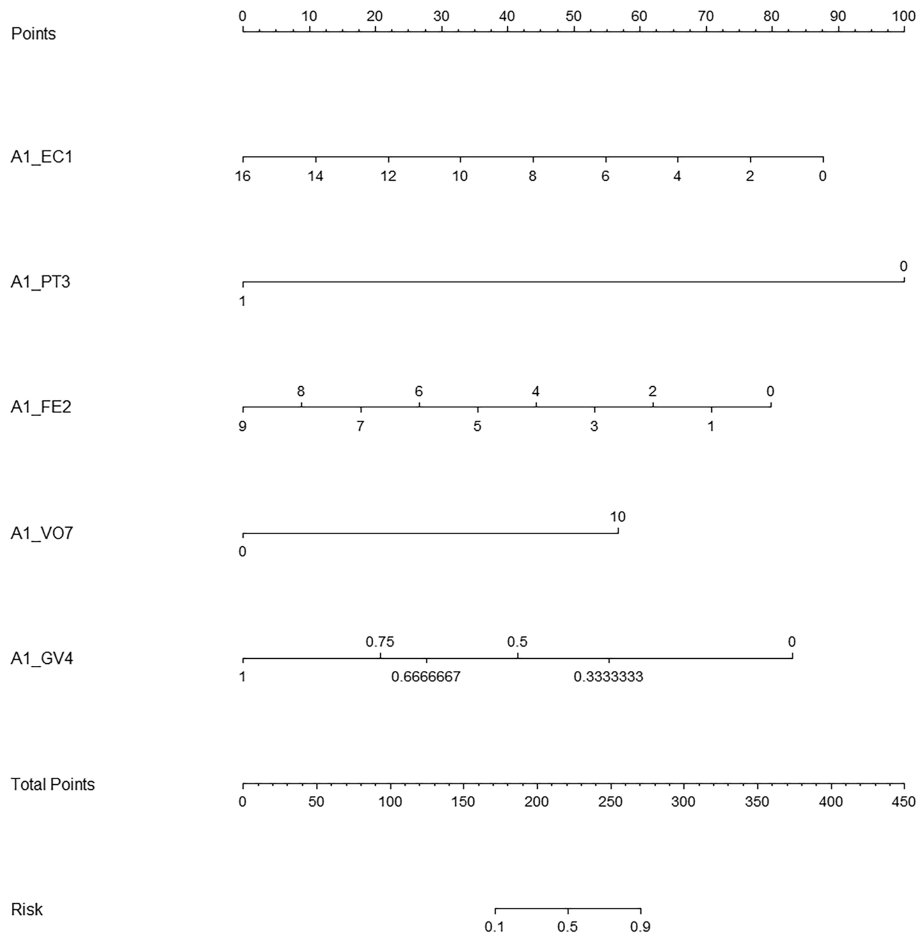

Fig. 2 Nomogram for prediction of ASD risks based on behavior indicators. As shown in the nomogram, a virtual straight line perpendicular to the top points line can be drawn for each behavior indicator line, and the value of each behavior indicator can be converted to a point from the top point line. The points of all behavior indicators in the nomogram can be added up to obtain a total point, and the ASD risk can be predicted by drawing a straight line perpendicular to the risk axis from the bottom total points line

Impaired social interaction and communication are defining features of ASD and emerge early in the affected population [34, 35]. Reduced and disrupted engagements of sustained attention to social stimuli were found as early as 6 months in infants who were later diagnosed with ASD [36]. Communicative behaviours, such as pointing, gestures, showing, and giving usually emerge by the end of infants' first year of life (9–12 months) to actively direct adult attention [37]. Children with ASD lose various communicative behaviours between 1 and 2 years of age [38–40]. Clear eye contact is found in 5- to 6-month-old infants, and eye-contact problems can be detected early in children with ASD [34]. Children with ASD also display fewer and shorter facial expressions than those without ASD [41]. In addition, social communication coordination, which refers to the combination of various communications, for example, gaze or vocalization or pointing with or without other communicative behaviours, was reported to be limited in individuals with ASD versus that in those without ASD [42, 43]. Therefore, this telehealth approach showed promise in identifying deficits in social interactions and communication in ASD.

Based on the reliability of the analysis results, behaviour indicators with reliability ICC values between 0 and 0.4, less than or equal to 0, and those that could not be calculated were excluded from the subsequent statistical analyses. Therefore, gestures, requesting, showing, social overtures, and RRB-related behaviour indicators were all excluded at this step from further analysis. When inspecting the original dataset, we found that 27.4%-35.5% of indicators could not be calculated or had ICC values less than or equal to 0. The reason why some indicators could not be calculated was that all original data were 0. The explanation for ICC = 0 was nonexistent reliability, and for ICC less than 0, the intragroup difference was greater than the intergroup difference [32]. We

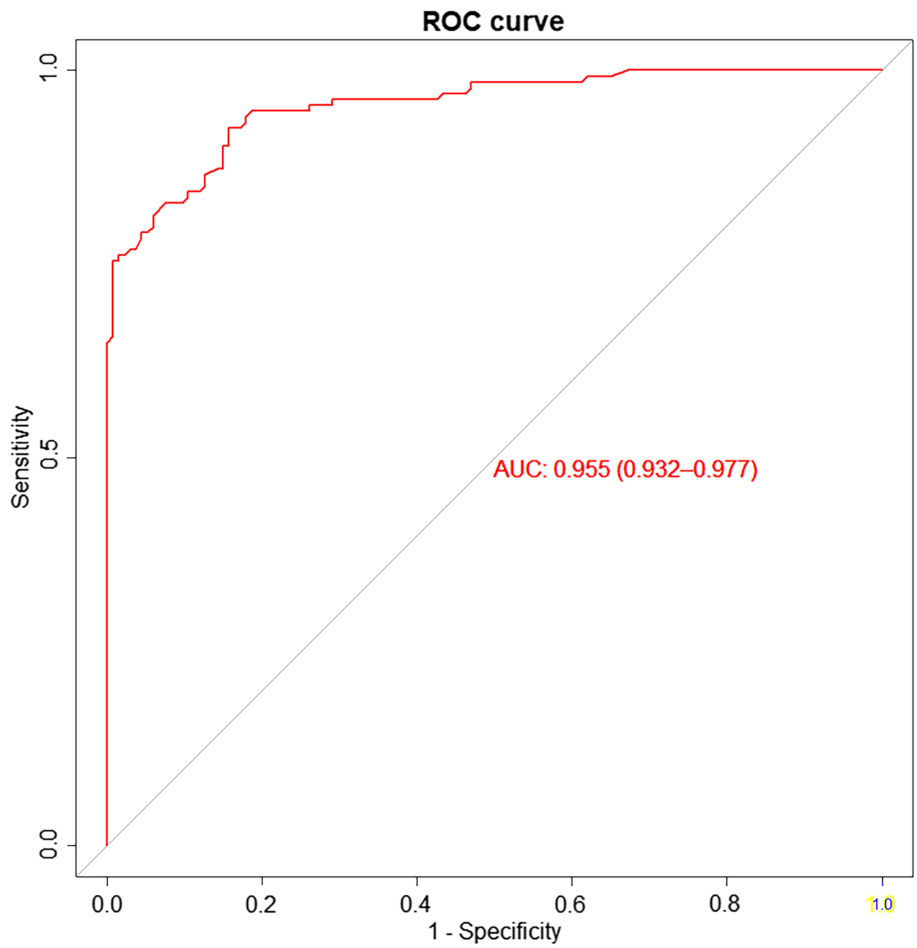

Fig. 3 Receiver operating characteristic curve of the prediction model

checked those indicators with ICCs less than or equal to zero; another possible reason could be too many 0 values for the indicator. Most behaviours related to those unusual ICC values were found to be less frequent behaviours in videos, for example, gestures, complex facial expressions, showing behaviours, and social overtures to examiners. We supposed that the structured activity, 1-min limitation, and children's strong motivation to snack left children with inadequate opportunities to use gestures, show complex facial expressions, and initiate social interactions with examiners, partially explaining why those behaviours were relatively rare in recorded videos. For our future analyses, we may need to expand or modify our activity settings to enable children to have more chances to show those behaviours and explore the extent to which those behaviours contribute to ASD screening and diagnosis.

Regarding requesting behaviour-related indicators, many were found to have low reliability (ICC value between 0 and 0.4). We reported significant differences in counts of pointing, counts of socially oriented vocalizations and verbalizations, and counts of eye gaze between the ASD and TD groups. Requesting behaviours could be different combinations of those behaviours. For example, children could appropriately request a specific snack from the examiner by pointing to the snack or saying the name of the snack, with or without eye contact with the examiner. Previous research has shown inconsistent results regarding whether children with ASD have impairments in requesting skills [44–46]. We considered that this could be because different studies used distinct definitions for requesting behaviour. Definitions for requesting could include object-centred strategies, such as reaching out for the object; strategies using the person as a self-propelling agent with contacting the person, such as putting an adult's hand on a toy; strategies using the person as a self-propelling agent without contacting the person, such as pointing or verbal request; and strategies using the person as perceiving subject, such as pointing with coordinated eye contact [45]. The current study defined requesting

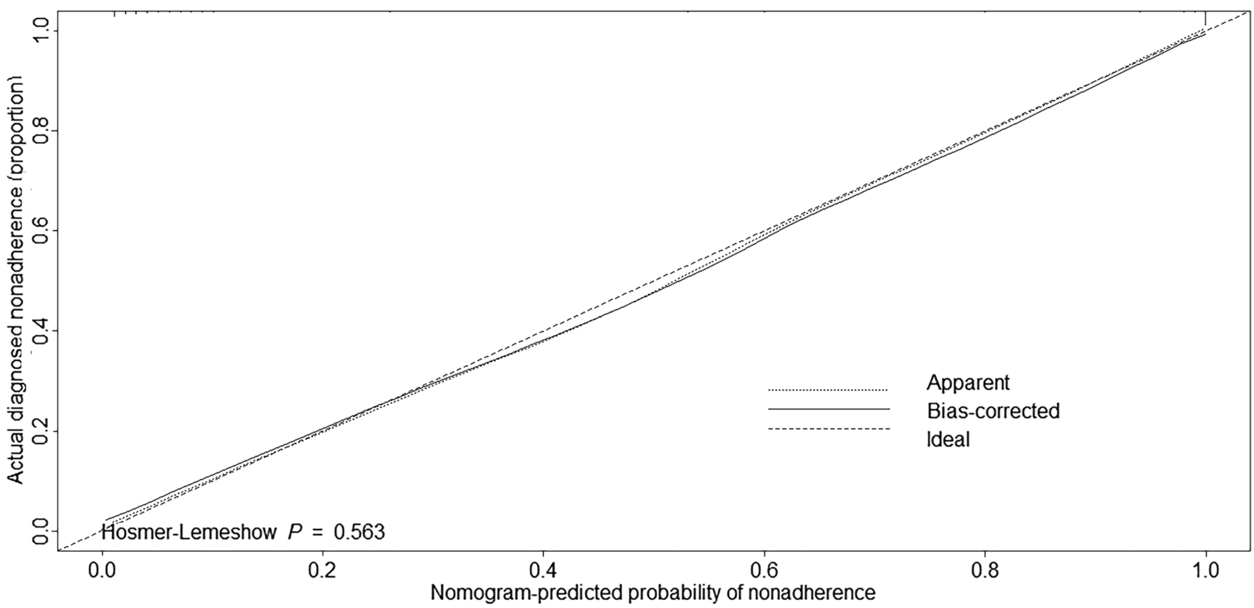

Fig. 4 Calibration curve of the nomogram prediction model for identifying ASD. The x-axis shows the nomogram-predicted probability of having an ASD diagnosis, and the y-axis exhibits the actual probability of being diagnosed with ASD

as children using pointing or vocalizations, with or without eye contact, to communicate the need for a certain snack to the examiner. Object-centred strategies, such as only reaching out for the snack, were excluded from this study's definition of request. When checking the reliability data of requesting behaviour indicators, some participants demonstrated inconsistent results between the 1st and 2nd tests in test–retest assessments and inter-examiner assessments. Specifically, in test–retest assessments, 21.7% of participants exhibited requesting behaviour in the 1st test but did not request in the 2nd test with the same examiner; 26.7% of participants had the opposite situation: requested in the 2nd test but not the 1st test. In the inter-examiner assessments, 26.2% of participants showed requesting behaviour with the 1st examiner but not with the 2nd examiner on the same day; 23% of participants had an adverse situation: requested with the 2nd examiner but not with the 1st examiner. These inconsistent situations revealed that the probability of children displaying requesting behaviours in test–retest and inter-examiner assessments was somewhat uncertain, which could be part of the reason for poor reliability in requesting. For our next analysis, we may consider adjusting the definition of requesting behaviours and activity setting to further understand the role of requesting in ASD screening or diagnosis in our research.

We found largely inconsistent ICCs for RRB indicators in different reliability tests. For example, ICC for total counts of immediate echolalia was 0.299 (ICC < 0.4) in test-test reliability but was 0.842 (ICC > 0.75) in inter-examiner reliability; ICC for total counts of unusual sensory-seeking behaviours was 0.669 (ICC > 0.4) in test-test reliability but was 0.189 (ICC < 0.4) in inter-examiner reliability; and similar situations for stereotyped uses of language and unusual and repetitive hand/body/limb movements. Largely inconsistent reliability excluded those RRB indicators from further analyses. RRB is one of the defining features of ASD in DSM-5. RRBs usually have a relatively low frequency of occurrence in both clinical and home settings. Therefore, most research relies on parents' reports to capture RRB instead of using direct observation methods [47, 48]. In addition, RRB is highly heterogeneous among ASD children, which could also contribute to reliability inconsistency in our research. This finding is consistent with our observation that during the snack time interaction, some children with ASD repeated the examiner's last few words immediately and showed some unusual and repetitive hand movements and some unusual smelling, mouthing, and seeing behaviours. However, those RRBs were not observed in all children with ASD in the 1-min video and were not always observed repeatedly in the same child. We may consider changing the indicator selection criteria and expanding the reliability sample size in the future to further explore how to include those RRB indicators in our prediction models.

Clinical prediction model

A clinical prediction model consisting of five behaviour indicators was identified in the current study with good prediction performance and good model fit. The 5-indicator model included 2 social interaction indicators, namely, counts of socially oriented vocalizations and verbalizations divided by the total counts of vocalizations and verbalizations and counts of children using giving behaviours to respond to the examiner's initiations divided by the total counts of the examiner's initiation of snack requisitions. The other 3 indicators were social communicative indices, including total counts of eye gaze, counts of standard pointing with coordinated gaze to object divided by the total counts of pointing with coordinated gaze to object, and total counts of appropriate facial expressions. The data-driven approach enables researchers to identify a subset of salient features for ASD identification, achieving the same diagnostic outcomes. Several previous studies using the ADOS database to conduct ASD feature selection showed promising results [49, 50]. The most important contribution of feature selection methods may be helping to shorten the time needed to collect children's information and conduct the assessment, thereby shortening the diagnostic process. The current study was a preliminary exploration of a clinical prediction model using the behavioural indicators collected by this telehealth tool, and further validation is required with a larger sample.

Clinical application setting

The results of this preliminary study revealed the feasibility, reliability and validity of an ASD telehealth assessment approach combining structured snack time interaction activities with behaviour coders. The COVID-19 pandemic significantly increased parents' needs for accessible ASD screening and diagnostic health care services, which posed challenges to the capability and distribution rationale of current ASD diagnostic resources. This study introduced a new approach to the existing tools to address this challenge and inequality, especially in low- and middle-income countries and regions. This approach designated examiners and coders separate roles to enhance accessibility to different regions. The 1-min assessment administration lowers the specialty training requirement for examiners, enabling the primary care providers to be widely trained to administer the assessment. The video behaviour coding can be finished within 10–15 min, and the coding results can be sent back to the primary care providers in a timely manner. Under clinical experts' training and supervision, this approach largely empowers primary care providers to conduct ASD screening and diagnosis by training them to administrate the assessment and provide a detailed children's behaviour report. The accessibility of ASD screening and diagnosis resources can be enhanced by saving parents time and travel costs. The inequality in medical resource distribution may be alleviated, especially for low- and middle-income countries and regions. Finally, this ASD telehealth approach uses indices such as count, rate, duration, and proportion, which are more suitable for future application scenarios of machine learning automatic identification of behaviours, further facilitating clinical workers in conducting quick and accurate clinical judgements.

Limitations and future directions

The current study must be considered in light of several limitations. First, the sample size was relatively small, and continual data collection is required to further evaluate the reliability and validity of the tool and the clinical prediction model. Due to the limited sample size, analyses among different age groups, language level groups, and ASD severity groups were not conducted. Along with the future enlargement of the sample size, further exploration of behaviour indicator features and prediction models for different age groups, language level groups, and ASD severity groups is required. Furthermore, the present study only included ASD and TD groups without non-ASD groups, such as participants with developmental delay (DD). However, in the real clinical setting, non-ASD populations, such as individuals with DD, are more prevalent than individuals with ASD. Therefore, to test the discrimination performance of this ASD assessment tool, non-ASD groups, such as DD groups, are needed in further research. Although snack time activity in this study was sensitive to capture deficits in social interaction and communication behaviours and RRBs in ASD, some other social interaction behaviours may not have been observed due to the activity setting, such as interactive play, response to names, and response to joint attention. Other assessment activities may need to be added to this telehealth assessment tool to improve its sensitivity and specificity. ICC-based variable selection was used in this study, which excluded indicators with unusual or poor reliability. A total of 22.4% of the whole sample underwent reliability tests, which may be insufficient to represent the whole sample. In addition, ICC-based variable selection also excluded some indicators with unusual or poor reliability that are important for ASD diagnosis, such as RRB-related indicators. How to select variables may need to be further discussed, and the reliability test sample needs to be expanded in the future. Finally, the future clinical application scenario for the tool is primary health care settings; however, this tool has not been validated in primary health care settings because it is still under development.

Administrators in this study part of the research team and were not primary care providers. Our next step is to recruit primary care providers to be trained as examiners to test the feasibility and acceptability of the ASD telehealth approach in primary care settings.

Conclusion

In conclusion, our study showed the feasibility of an ASD telehealth approach combining structured snack time interaction activities administered by examiners with limited clinical training and behaviour coding by trained coders. Significant behavioural differences between children with ASD and typically developing children were detected, including children's response to the examiner's initiation, eye gaze, pointing, facial expressions, vocalization and verbalization, and giving behaviours in snack activities. A clinical prediction model was constructed with good performance to distinguish children aged 1–6 years old with and without ASD. The current study indicates the promise of this novel telehealth tool in primary health care settings for early ASD screening.

Abbreviations

ASD	Autism spectrum disorder
TD	Typically developing individuals
LMICs	Low- and middle-income countries
PSM	Propensity score matching
RRB	Restricted and repetitive behaviours
ICC	Intraclass Correlation Coefficients
OR	Odds ratio
CI	Confidence intervals
AUC	Area under the curve

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s44247-023-00047-8.

Additional file 1: Table S1. Behavior coding guide for snack time. **Table S2.** Group difference analysis. **Table S3.** Reliability test results.

Acknowledgements

We would like to express our deepest gratitude to the families who participated in this study. Their support and willingness were invaluable to our research. We are truly grateful for their time and contributions.

Authors' contributions

J.L. contributed to the study design, recruitment, and data collection, and critically reviewed the statistical analysis and the manuscript. X.L. and Y.G. contributed to the study design, recruitment, and data collection. Z.M. contributed to the study design and data collection, interpreted the data, and wrote the manuscript. Y.J. conceptualized and carried out the statistical analysis. R.H. and J.C. contributed to the statistical analysis. Q.L., X.S., J.C., D.X., and T.Z. contributed to recruitment and data collection. All authors agree to be accountable for all aspects of the work. All authors read and approved the final manuscript.

Funding

This study was not funded by any specific grant from public, commercial or non-profit organizations.

Availability of data and materials

The datasets generated during and/or analysed during the current study are not publicly available due to ethical restrictions regarding data protection issues and the study-specific consent text and procedure, but anonymized data are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

All methods were carried out in accordance with the guidelines and recommendations specified within the Helsinki declaration. The procedure was approved by the Ethics Committee of Peking University Sixth Hospital (approval no. 2021–71), and written informed consent was obtained from the parents/legal guardians of all participants before enrolling in this study.

Consent for publication

Not applicable.

Competing interests

The authors (Z.M., Y.J., J.C., Y.G., X.L., and J.L.) declare that there were two pending patent applications related to this research (No. 202310300588.1 and No. 202310315701.3). R.H., Q.L., X.S., J.C., D.X., and T.Z. have no conflicts of interest with regard to the content of this manuscript.

Author details

[1]ALSO Institute, ALSOLIFE, Beijing, China. [2]Peking University Sixth Hospital, Peking University Institute of Mental Health, NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), 51 Huayuan Road, Haidian District, Beijing 100191, China.

Received: 4 April 2023 Accepted: 11 September 2023
Published online: 07 November 2023

References

1. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. Lancet. 2018;392(10146):508–20.
2. Maenner MJ, Shaw KA, Bakian AV, Bilder DA, Durkin MS, Esler A, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2018. MMWR Surveill Summ. 2021;70(11):1.
3. Li Y-A, Chen Z-J, Li X-D, Gu M-H, Xia N, Gong C, et al. Epidemiology of autism spectrum disorders: Global burden of disease 2019 and bibliometric analysis of risk factors. Front Pediatr. 2022;10:972809.
4. Corsello CM. Early intervention in autism. Infants Young Child. 2005;18(2):74–85.
5. Moore V, Goodson S. How well does early diagnosis of autism stand the test of time? Follow-up study of children assessed for autism at age 2 and development of an early diagnostic service. Autism. 2003;7(1):47–63.
6. 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(6):578–87.
7. van't Hof M, van Berckelear-Onnes I, van Nieuwenhuyzen A, Daniels AM, Deen M, et al. Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. Autism. 2021;25(4):862–73.
8. Bauer K, Morin KL, Renz TE III, Zungu S. Autism assessment in low-and middle-income countries: Feasibility and usability of western tools. Focus Autism Other Dev Disabil. 2022;37:179–88 10883576211073691.
9. Sutherland R, Trembath D, Roberts J. Telehealth and autism: A systematic search and review of the literature. Int J Speech Lang Pathol. 2018;20(3):324–36.
10. Liu M, Ma Z. A systematic review of telehealth screening, assessment, and diagnosis of autism spectrum disorder. Child Adolesc Psychiatry Ment Health. 2022;16(1):79.
11. Barbaro J, Yaari M. Study protocol for an evaluation of ASDetect-a Mobile application for the early detection of autism. BMC Pediatr. 2020;20(1):1–11.

12. Anand V, Carroll AE, Biondich PG, Dugan TM, Downs SM. Pediatric decision support using adapted Arden Syntax. Artif Intell Med. 2018;92:15–23.
13. Juárez AP, Weitlauf AS, Nicholson A, Pasternak A, Broderick N, Hine J, et al. Early identification of ASD through telemedicine: Potential value for underserved populations. J Autism Dev Disord. 2018;48(8):2601–10.
14. Dow D, Holbrook A, Toolan C, McDonald N, Sterrett K, Rosen N, et al. The brief observation of symptoms of autism (BOSA): development of a new adapted assessment measure for remote telehealth administration through COVID-19 and beyond. J Autism Dev Disord. 2022;52(12):5383–94.
15. Megerian JT, Dey S, Melmed RD, Coury DL, Lerner M, Nicholls CJ, et al. Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. NPJ Digit Med. 2022;5(1):1–11.
16. Smith CJ, Rozga A, Matthews N, Oberleitner R, Nazneen N, Abowd G. Investigating the accuracy of a novel telehealth diagnostic approach for autism spectrum disorder. Psychol Assess. 2017;29(3):245.
17. Wagner L, Corona LL, Weitlauf AS, Marsh KL, Berman AF, Broderick NA, et al. Use of the TELE-ASD-PEDS for Autism Evaluations in Response to COVID-19: Preliminary Outcomes and Clinician Acceptability. J Autism Dev Disord. 2021;51(9):3063–72.
18. Reese RM, Braun MJ, Hoffmeier S, Stickle L, Rinner L, Smith C, et al. Preliminary Evidence for the Integrated Systems Using Telemedicine. Telemed J E Health. 2015;21(7):581–7.
19. Solomon D, Soares N. Telehealth approaches to care coordination in autism spectrum disorder. Interprofessional Care Coordination for Pediatric Autism Spectrum Disorder: Translating Research into Practice. 2020. p. 289–306.
20. Dow D, Day TN, Kutta TJ, Nottke C, Wetherby AM. Screening for autism spectrum disorder in a naturalistic home setting using the systematic observation of red flags (SORF) at 18–24 months. Autism Res. 2020;13(1):122–33.
21. Corona LL, Weitlauf AS, Hine J, Berman A, Miceli A, Nicholson A, et al. Parent perceptions of caregiver-mediated telemedicine tools for assessing autism risk in toddlers. J Autism Dev Disord. 2021;51(2):476–86.
22. Jones MK, Zellner MA, Hobson AN, Levin A, Roberts MY. Understanding caregiver satisfaction with a telediagnostic assessment of autism spectrum disorder. Am J Speech Lang Pathol. 2022;31(2):982–90.
23. Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, et al. Integrating machine learning and multiscale modeling—perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. NPJ digital medicine. 2019;2(1):1–11.
24. Manfredonia J, Bangerter A, Manyakov NV, Ness S, Lewin D, Skalkin A, et al. Automatic recognition of posed facial expression of emotion in individuals with autism spectrum disorder. J Autism Dev Disord. 2019;49(1):279–93.
25. Deng A, Yang T, Chen C, Chen Q, Neely L, Oyama S. Problem behaviors recognition in videos using language-assisted deep learning model for children with autism. arXiv preprint arXiv:2211.09310. 2022.
26. Frost KM, Koehn GN, Russell KM, Ingersoll B. Measuring child social communication across contexts: Similarities and differences across play and snack routines. Autism Res. 2019;12(4):636–44.
27. Pasco G, Gordon RK, Howlin P, Charman T. The classroom observation schedule to measure intentional communication (COSMIC): An observational measure of the intentional communication of children with autism in an unstructured classroom setting. J Autism Dev Disord. 2008;38(10):1807–18.
28. Lee SH, Lee LW. Promoting snack time interactions of children with autism in a Malaysian preschool. Topics Early Child Spec Educ. 2015;35(2):89–101.
29. Lord C, Rutter M, DiLavore P, Risi S, Gotham K, Bishop S. Autism diagnostic observation schedule: ADOS-2. Western Psychological Services. 2012.
30. Stone WL, Coonrod EE, Turner LM, Pozdol SL. Psychometric properties of the STAT for early autism screening. J Autism Dev Disord. 2004;34(6):691–701.
31. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. 2013.
32. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.
33. Chong I-G, Jun C-H. Performance of some variable selection methods when multicollinearity is present. Chemom Intell Lab Syst. 2005;78(1–2):103–12.
34 Johnson CP, Myers SM, American Academy of Pediatrics Council on Children With D. Identification and evaluation of children with autism spectrum disorders. Pediatrics. 2007;120(5):1183–215.
35. Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioral manifestations of autism in the first year of life. Int J Dev Neurosci. 2005;23(2–3):143–52.
36. Jones EJ, Venema K, Earl R, Lowy R, Barnes K, Estes A, et al. Reduced engagement with social stimuli in 6-month-old infants with later autism spectrum disorder: a longitudinal prospective study of infants at high familial risk. J Neurodev Disord. 2016;8:1–20.
37. Camaioni L, Perucchini P, Bellagamba F, Colonnesi C. The role of declarative pointing in developing a theory of mind. Infancy. 2004;5(3):291–308.
38. Sparaci L. Beyond the point: A basic guide to literature on pointing abilities in children with autism. Humana Mente Journal of Philosophical Studies. 2013;6(24):177–202.
39. Ramos-Cabo S, Vulchanov V, Vulchanova M. Gesture and language trajectories in early development: An overview from the autism spectrum disorder perspective. Front Psychol. 2019;10:1211.
40. Charman T. Why is joint attention a pivotal skill in autism? Philos Trans R Soc Lond B Biol Sci. 2003;358(1430):315–24.
41. Keating CT, Cook JL. Facial expression production and recognition in autism spectrum disorders: a shifting landscape. Child Adolesc Psychiatr Clin. 2020;29(3):557–71.
42. Delehanty A, Wetherby AM. Coordinated social communication in toddlers with and without autism spectrum disorder during a home observation. Autism Dev Lang Impair. 2022;7:23969415221121090.
43. Wetherby AM. Understanding and measuring social communication in children with autism spectrum disorders. Social and communication development in autism spectrum disorders: Early identification, diagnosis, and intervention. 2006. p. 3–34.
44. Mundy P, Sigman M, Kasari C. Joint attention, developmental level, and symptom presentation in autism. Dev Psychopathol. 1994;6(3):389–401.
45. Phillips W, Gómez JC, Baron-Cohen S, Laá V, Rivière A. Treating people as objects, agents, or "subjects": How young children with and without autism make requests. J Child Psychol Psychiatry. 1995;36(8):1383–98.
46. Paparella T, Goods KS, Freeman S, Kasari C. The emergence of nonverbal joint attention and requesting skills in young children with autism. J Commun Disord. 2011;44(6):569–83.
47 Harrop C, McConachie H, Emsley R, Leadbitter K, Green J, Consortium P. Restricted and repetitive behaviors in autism spectrum disorders and typical development: Cross-sectional and longitudinal comparisons. J Autism Dev Disord. 2014;44:1207–19.
48. Turner M. Annotation: repetitive behaviour in autism: a review of psychological research. J Child Psychol Psychiatry Allied Discip. 1999;40(6):839–49.
49. Kosmicki J, Sochat V, Duda M, Wall D. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Transl Psychiatry. 2015;5(2):e514.
50. Bone D, Bishop SL, Black MP, Goodwin MS, Lord C, Narayanan SS. Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. J Child Psychol Psychiatry. 2016;57(8):927–37.

Publisher's Note