The Feasibility and Validity of Autism Spectrum Disorder Screening Instrument: Behavior Development Screening for Toddlers (BeDevel)—A Pilot Study

Guiyoung Bong, Jeong-Heun Kim, Yoohwa Hong, Nan-He Yoon, Hyunjung Sunwoo, Jeong Yun Jang, Miae Oh, Kyung-Sook Lee, Seokjin Jung, and Hee Jeong Yoo

Although early screening is critical for individuals with autism spectrum disorder (ASD) in order to receive early intervention and improve function later in life, screening is often delayed. Limitations of existing screening instruments, and the need for a culturally appropriate early screening tool in Korean children, led us to develop Behavior Development Screening for Toddlers (BeDevel). The BeDevel assessment consists of two parts: BeDevel-Interview, a structured interview measure for parents/primary caregivers; and BeDevel-Play, a play-based semi-structured observational measure in children. To examine the feasibility and validity of BeDevel, 155 children (N = 75 ASD, N = 55 typical development, N = 25 developmentally delayed) aged 18–42 months (M = 31.54 months, SD = 7.60) were examined through parent-reported screening questionnaires, BeDevel, and standard diagnostic assessments. When BeDevel items were analyzed using Cohen's kappa statistics, most items in BeDevel-Interview and all items in BeDevel-Play were reasonably consistent with diagnoses. We identified primary items, which were significantly interacted with actual diagnosis in the chi-squared test ($P < 0.05$, range = 0.000–0.032). Using cutoff numbers of items determined using the receiver operating characteristics curve, BeDevel showed satisfactory levels of sensitivity (83.33%–100%), specificity (81.25%–100%), positive predictive values (80.65%–100%), and negative predictive values (83.87%–100%), as well as high internal consistency (Cronbach's α = 0.866–959). The agreement between BeDevel and most other screening/diagnostic instruments was moderate (k = 0.419–1.000). These results suggest that BeDevel can be a useful instrument for early screening of ASD. ***Autism Res*** *2019, 12: 1112–1128.* © 2019 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary: Although early screening is critical for individuals with autism spectrum disorder (ASD) in order to receive early intervention and improve function later in life, screening is often delayed. Limitations of existing screening instruments and the need for a culturally appropriate early screening tool in Korean children led us to develop Behavior Development Screening for Toddlers (BeDevel). The BeDevel assessment consists of two parts: BeDevel-Interview, a structured interview measure for parents/primary caregivers; and BeDevel-Play, a play-based, semi-structured observational measure in children. In order to test the feasibility and validity of BeDevel, we analyzed preliminary data of total 155 children aged 18–42 months, examined through parent-reported screening questionnaires, BeDevel, and standard diagnostic assessments. When individual items were analyzed, responses of all BeDevel-Interview items and of most BeDevel-Play items well matched actual diagnoses, and we identified primary items, which were particularly useful in differentiating between the ASD group and the non-ASD group. With the optimal screening criteria determined, the BeDevel was able to identify individuals with a diagnosis of ASD and those without it, all at satisfactory levels. Lastly, BeDevel items were closely related as a set, and the BeDevel screening results were reasonably consistent with the results of most other screening/diagnostic instruments. These results suggest that BeDevel can be a useful instrument for early screening of ASD.

Keywords: autism spectrum disorder; early screening; early detection; early sign; toddler

Introduction

Early intervention in autism spectrum disorder (ASD) is critical to improve early deficits and function in later life [Bradshaw, Steiner, Gengoux, & Koegel, 2015; Estes et al., 2015; Mundy & Crowson, 1997]. Identification of ASD at a young age allows early intervention, but early diagnosis is not always possible. According to the literature, early signs of ASD are generally detected and diagnosed as early as 12–24 months of age [Boyd, Odom, Humphreys, & Sam,

From the Department of Psychiatry, Seoul National University Bundang Hospital, Seongnam, South Korea (G.B., J.-H.K., Y.H., H.J.Y.); Department of Health Administration, Hanyang Cyber University, Seoul, South Korea (N.-H.Y.); Department of Psychiatry, Soonchunhyang University Hospital, Seoul, South Korea (H.S.); Sejong Center for Persons with Developmental Disabilities, Sejong, South Korea (J.Y.J.); Department of Psychiatry, Kyung Hee University Hospital, Seoul, South Korea (M.O.); Department of Rehabilitation, Hanshin University, Osan, South Korea (K.-S.L.); Sewon Infant Child Development Center, Seoul, South Korea (K.-S.L., S.J.); Seoul National University College of Medicine, Seoul, South Korea (H.J.Y.)

Received October 22, 2018; accepted for publication April 21, 2019

Address for correspondence and reprints: Hee Jeong Yoo, Department of Psychiatry, Seoul National University Bundang Hospital, Seoul National University College of Medicine, 300 Gumi-ro, Bundang-gu, Seongnam, Gyeonggi 463-707, South Korea. E-mail: hjyoo@snu.ac.kr

Published online 10 May 2019 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.2117

2010; Kleinman et al., 2008; Zwaigenbaum et al., 2009]. However, the average age of diagnosis is much higher in the United States and Korea [Baio et al., 2018; National Institute of Special Education, 2015]. This discrepancy stems from time-consuming evaluations, expensive care systems, long waiting times, shortage of professionals, and low awareness of paraprofessionals [Althouse & Stockman, 2006; Austin et al., 2016; Bisgaier, 2011; Fenikile, Ellerbeck, Filippi, & Daley, 2014; Kalb et al., 2012; Sunwoo, Noh, Kim, Kim, & Yoo, 2017; Wiggins, Baio, & Rice, 2006]. Additionally, in Korea, two major reasons cited for the lag between when parents first suspect symptoms and the time of diagnosis are (a) the belief that symptoms will resolve as the child grows (52.2%) and (b) doctor recommendations to delay diagnosis in young children (39.7%). This highlights the relatively limited awareness of ASD in Korean society [National Institute of Special Education, 2015].

Screening is a pre-diagnostic method to identify early manifestation of disorders, and can be the first step of the diagnostic process [Corbisiero, Hartmann-Schorro, Riecher-Rössler, & Stieglitz, 2017]. Typically, screening for ASD has been performed with screening instruments in the form of caregiver-rated questionnaires or simple observation by trained professionals [Towle & Patrick, 2016]. Examples of existing screening instruments are the Modified Checklist for Autism in Toddlers (M-CHAT) [Robins, Fein, & Barton, 1999; Robins, Fein, Barton, & Green, 2001], the Screening Tool for Autism in Toddlers and Young Children [Stone, Coonrod, & Ousley, 2000; Stone, Coonrod, Turner, & Pozdol, 2004; Stone, McMahon, & Henderson, 2008], the Social Communication Questionnaires (SCQs) [Rutter, Bailey, Lord, & Berument, 2003], and the Social Attention and Communication Study (SACS) [Barbaro & Dissanayake, 2010]. Of these, the M-CHAT and SCQ have been translated into Korean [Kim et al., 2015; Yoo, 2008].

Previous screening instruments for ASD have adopted either cutoff scores or at-risk behavior approaches. Both have advantages, but some studies have suggested that the identification of toddlers with ASD using cutoff scores may be ineffective [Corsello, Cook, & Levanthal, 2003; Newschaffer, Lee, David, & Lee, 2004] because scores can vary across cultures [Chiang et al., 2012; Kim et al., 2015]. Instead, at-risk responses from follow-up protocols in M-CHAT, CHAT, and SACS can identify specific characteristics of ASD and are used to collect additional information and to reduce the occurrence of false-positives [Robins et al., 1999]. Further, parental interviews or questionnaires may also result in reporter bias or subjective responses [Glascoe, 2000; Stone, Hoffman, Lewis, & Ousley, 1994].

Some screening tools are administered at a specific age, and many at-risk children may be missed due to their age [Barbaro & Dissanayake, 2010]. For example, joint attention (JA) is one of the critical factors used to distinguish children at risk of ASD from typically developing children and children with other developmental disabilities (DDs). However, although the children with ASD who were younger than 20 months showed significant delays in response to JA, the children with ASD who were older than 19 months did not show a significant difference in response to JA compared to typically developing children [Mundy, Sigman, & Kasari, 1994]. Further, children with ASD show significantly less JA use than do children with developmental delay (DD) at 24 months, but these two groups show very similar levels of JA at 48 months [Naber et al., 2007]. This indicates that the screening result may differ depending on the age at administration. In addition, as the majority of children with ASD have developmental delays, multiple evaluations at different ages are recommended [Szatmari et al., 2016; Tager-Flusberg, 2010]. As such, we have developed different versions of screening instruments for use at different age ranges. According to previous focus group interviews (FGIs) done by our group, many child healthcare service providers have a low level of understanding of the pathognomonic behaviors of ASD (e.g., abnormalities in JA and eye gaze), and often fail to diagnose ASD based on only a few features [Jang et al., 2016; Sunwoo et al., 2017]. Even if nursery or kindergarten teachers quickly suspect abnormalities in children, they can be hesitant to articulate this to parents due to the potential negative impact of the highly stigmatized term "autism." Therefore, it is essential that screening instruments include easy, clear instructions, and fundamental background information about typical and atypical behaviors in social communication and social interaction in toddlerhood, to help improve understanding of the disorder and assist with communication from caregivers. To this end, we developed a detailed manual, educational handbook, and visual aids alongside the interview questionnaire and play observation tool.

The primary purpose of this study was to develop and validate a culturally appropriate and applicable screening instrument to identify toddlers and children at risk of ASD in the Korean population. To do so, we designed a comprehensive set of screening instruments consisting of a simple caregiver interview, play-based observation, an instruction manual and educational material, which was collectively termed *Behavior Development Screening for Toddlers* (BeDevel). We examined the validity of BeDevel through diagnostic confirmation and selected key items to determine the criteria for screening children at risk of ASD.

Methods

Developmental Procedures of the Instrument

Development of items. Two board-certified child and adolescent psychiatrists, a developmental psychologist, two clinical psychologists, and a special education professional participated in the design of BeDevel. First, we decided to develop distinct sets of caregiver interviews and play observation for various age groups. We used both

interviews and observations based on prior studies that have suggested that bias may occur when using only one method [Glascoe, 2000; Miller et al., 2011]. The age groups were allocated as (a) 9–11 months, (b) 12–17 months, (c) 18–23 months, (d) 24–35 months, and (e) 36–42 months of age. As some signs of ASD can be observed as early as 12 months, we decided to include the lowest age group to increase the chance of early identification of ASD [Zwaigenbaum et al., 2009]. Given that the identification of ASD is often delayed until 36 months of age or later in Korea, we decided to expand the definition of toddlerhood up to 42 months [National Institute of Special Education, 2015]. In addition, these age groups align with those of the Official National Health Screening Program for Infants and Children in Korea, thus, BeDevel could prove beneficial as an instrument for screening and monitoring those who are found to be at-risk in their regular check-ups [The Korea Centers for Disease Control and Prevention, 2017].

Second, an extensive literature review was conducted on early signs of ASD to determine specific behaviors observed in toddlers with a high risk of ASD in each age range. We selected early signs of ASD that have been proven to be significant in children aged 42 months or younger, and sorted them into three subdomains of deficits in social communication and reciprocal social interaction, and four subdomains of restricted, repetitive patterns of behaviors, interests, or activities according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria of ASD [American Psychiatric Association, 2013].

Lastly, we developed a series of semi-structured caregiver interview questions (BeDevel-Interview Protocol; BeDevel-I) to examine early behaviors, and a series of simple play activities (BeDevel-Play Protocol; BeDevel-P) through which those behaviors can be examined observationally. The preliminary pool of questions and activities were reviewed by a panel of experts, as well as those who work closely with young children in related fields, and modifications were made when needed as discussed in the following sections.

Content validity. *Examination for content validity through expert panel review*. BeDevel-I and the BeDevel-P were reviewed by a panel of experts consisting of 11 child psychiatrists, two special education professionals, two clinical psychologists, and two developmental psychologists who considered content appropriateness, timing, and importance of each item for the purpose of early screening. Overall, items in BeDevel-I and BeDevel-P were evaluated as appropriate for early screening for ASD in infants and toddlers at the targeted ages. However, several suggestions were carefully discussed and revisions were made accordingly. For example, comments for examiners to provide children during BeDevel-P were revised or added, and guidelines for play activities were modified to more clearly explain the tasks. In addition, scoring criteria were clarified by adding more possible examples of responses, and an "intermediate" score (i.e.; score of 2) was added for some items that were previously to be scored dichotomously, such as the use of requesting gestures and facial expressions. For BeDevel-I, a frequency range to obtain each score was specified for social response items (*Response to Name*, *Social Smile*, *Social Referencing*, and *Following Other's Pointing*).

Examination for content validity through childcare worker FGIs. In Korea, 65.6% of infants and toddlers aged 0–3 years attend daycare centers [Ministry of Health and Welfare, 2017; Ministry of the Interior and Safety, 2017]. Thus, in order to examine whether BeDevel is feasible, practical, and field-friendly, we conducted FGIs with 30 daycare teachers in Seongnam City, Korea, who work with children of the target ages of this study on a daily basis. Interviewees consisted of seven regular teachers, nine special education teachers, and 14 head teachers. Fifteen interviewees had 10 or more years of professional experience, 11 had 6–9 years, and 4 had less than 5 years. All were female, reflecting that daycare teachers in Korea are predominantly female. We divided the interviewees in three groups and repeated the interview questions in each group.

The BeDevel-I and BeDevel-P protocols were reviewed item-by-item, and interviewees were asked to evaluate whether the expression of the questions and the behavior/play examples in the protocols were appropriate for use in the field. Collected opinions were carefully discussed by the researchers, and revisions were made if necessary. Notably, descriptions were supplemented with various examples of age-appropriate typical social behaviors suggested by the teachers, so that semi-professionals can easily differentiate typical and atypical (ASD-related) behaviors. In addition, the wording of the questions in the BeDevel-I was revised to make them more clearly understandable to parents/caregivers. For BeDevel-P, the reactive remarks given to children were modified to increase comprehension in infants and toddlers. Suggestions regarding toys that interest infants and toddlers were implemented into the play material set.

BeDevel as a comprehensive package. BeDevel aims to be an early screening instrument that is ready for use not only by professionals specialized in ASD, but also by trained allied professionals who work with young children on a regular basis. For this purpose, we developed a *Comprehensive Package* (comprised of BeDevel-I, BeDevel-P, and a BeDevel Manual) that can be easily utilized by users with different educational backgrounds and experiences. This package includes detailed instructions for both administering and scoring BeDevel, and a BeDevel

Handbook that helps users better understand behaviors related to ASD.

BeDevel-I. BeDevel-I is a structured interview for primary caregivers assessing social communication, reciprocal social interaction and repetitive behavior, and limited interest in toddlers across five age ranges. BeDevel-I consists of questions regarding whether a child shows social behaviors in three domains known to be impaired in children with ASD compared with typically developing children: social interaction, nonverbal communication, and social relationships. In addition, BeDevel-I assesses atypical behaviors including repetitive behaviors, limited interest, and hypersensitivity or hyposensitivity to sensory stimuli. The number of questions varies depending on the age group: for social communication, there are six items for individuals aged 9–11 months, nine items for 12–17 months, 11 items for 18–23 months, 13 items for 24–35 months, and 14 items for 36–42 months. There are five items to assess atypical behaviors and limited interest in all age groups. Each item is rated from 1 to 3 with 1 indicating typical behavior, 2 indicating insufficient or inconsistent behavior, 3 indicating highly limited behavior; and some items are given an additional score, "S," which specifies atypical behavior. BeDevel-I takes 10–15 min to administer.

BeDevel-P. BeDevel-P is a play-based, semi-structured observational measure assessing ASD-related behavior in toddlers. BeDevel-P consists of activities and tasks to observe the three domains of social behavior described above (social interaction, nonverbal communication, and social relationships), atypical repetitive behaviors, limited interest, and response to sensory stimuli. The play was designed to contain appropriate activities for individuals of each age range; the number of items varied for each range: three items for individuals aged 9–11 months, nine items for 12–17 months, 10 items for 18–23 months, 12 items for 24–35 months, and 14 items for 36–42 months. The scoring system is generally identical to those of BeDevel-I, with a score of 1 indicating typical behavior, and 3 representing limited behavior. The score "S" indicates atypical behaviors. The detailed constructs of BeDevel-I and BeDevel-P are summarized in Tables 1 and 2. BeDevel-P also takes 10–15 min to administer.

BeDevel manual. The purpose of the BeDevel Manual is to provide clear instructions for both administering and scoring the assessments to ensure consistency. Provided with each task are the purpose, tools required, detailed instructions for administering and scoring, and a link to more information in the handbook. The "Purpose" section explains the social behaviors and ASD characteristics to look for in each task. The "Tools" section lists the names and the numbers of the tools that should be used in each task in order to avoid arbitrary use of tools by individual examiners. The "Administration Instruction" section gives detailed explanations of how to conduct tasks, including the number of attempts, alternative tasks when necessary, and the extent to which a parent should be allowed to get involved. In addition, it also explains why the specific behavior is important in children's social development. Because the BeDevel Manual is designed to be used by those who do not have previous experience with ASD children or knowledge of ASD in general, introduction to the concept of specific social behavior itself could assist in comprehension of the task and observation. A "Scoring Instruction" section explains what each score means and includes examples. In addition, the "Link" section gives the page number in the BeDevel Handbook where each social behavior of interest is explained in more detail.

BeDevel handbook. The BeDevel Handbook is designed to provide background knowledge of ASD for the examiner, to enhance the understanding and ensure accurate and efficient administration of the screen. For each behavior assessed in the BeDevel-I and BeDevel-P, the handbook describes the definition, developmental implications and importance, developmental trajectories, and differences in manifestation in typically and non-typically developing children. The handbook was collaboratively written by five child psychiatrists, three developmental psychologists, three special education professionals, and two linguistics experts who have expertise in either the diagnosis of and/or intervention in children with ASD. It is written in plain language so that professionals who do not have sufficient knowledge about ASD can easily understand. In addition, we also developed "BeDevel-Visual Aids," 3-D animation video materials showing several key behaviors including JA, social referencing, social smiling, and reciprocal play in a typical child and a child with ASD, as well as atypical (ASD-related) behaviors including a repetitive pattern of behavior (e.g., lining up toys), the use of another's body to communicate, echolalia, and sensory seeking.

Reliability and Diagnostic Validity

Participants. Participants were recruited and referred from multiple sources: the child and adolescent psychiatric clinic, the Pediatrics and Child Rehabilitation clinic at Seoul National University Bundang Hospital, local primary clinics participating in the National Health Screening Program for Infants and Children, ASD treatment institutes, community child mental health care centers, online self-help communities for parents who raise children with DDs, daycare centers, and through advertisements on bulletin boards in multifamily housing complexes in Seongnam City. There were also participants who were introduced through other parents participating in the study. The

Table 1. Construct of BeDevel-Interview

DSM-5 criteria	Items	9 months	12 months	18 months	24 months	36 months
A-1. Deficits in social–emotional reciprocity	Response to name	V	V	V	V	V
	Social smile	V	V	V	V	V
	Social reference	V	V	V	V	V
	Sharing interest			V	V	V
	Back-and-forth conversation					V
A-2. Deficits in nonverbal communicative behaviors used for social interaction	Eye contact	V	V	V	V	V
	Facial expression	V	V	V	V	V
	Gesture		V	V	V	V
	Pointing		V	V	V	V
	Following other's pointing			V	V	V
	Imitation of actions		V	V	V	V
A-3. Deficits in developing, maintaining, and understanding relationships	Social interest and social relationship	V	V	V	V	V
	Interest in peers				V	V
	Shared imaginative play				V	V
B. Restricted, repetitive patterns of behavior, interests, or activities	Stereotyped/repetitive behaviors	V	V	V	V	V
	Insistence on sameness/adherence to routines/restricted pattern of behavior	V	V	V	V	V
	Restricted fixated interest	V	V	V	V	V
	Unusual sensory interest and hyper/hyposensitivity	V	V	V	V	V

inclusion criteria of the study were toddlers between the ages of 9 and 42 months. The exclusion criteria were severe medical conditions, neurological difficulties, severe sensory and motor disorders, and serious rejection that made participation in the assessment difficult. Some children were entered into the study because they had been diagnosed with ASD or communication disorder in advance, whereas others had developmental concerns or were entered into the study simply to identify their developmental characteristics. Therefore, the assessment results led to further treatment if necessary. The study was approved by the Institutional Review Board at Seoul National University Bundang Hospital and all parents of children provided written informed consent.

Procedures. *Rater training*. The examiners who performed the assessments received 6 × 4-hr training sessions in which they studied the BeDevel protocols and manuals in detail, watched simulations, and discussed ways to overcome practical difficulties and avoid verbal and behavioral mistakes. The entire training process was supervised by the researcher in charge.

Prior to the actual administration, BeDevel-P and BeDevel-I were performed on three groups of children aged 18–23 months, 24–35 months, and 36–42 months, and their parents/caregivers. Four examiners rated each group independently, and inter-rater reliability was calculated. The inter-rater reliability of BeDevel-P scores ranged from 84.2% to 100%, and the inter-rater reliability of the BeDevel-I ranged from 84.6 to 100%. Any disagreements in scoring among examiners were reviewed and discussed to improve the accuracy of scoring. Once the adequate level of reliability was established, each examiner performed their first three to four cases under close supervision; from then onward, score results and video-recorded assessments were reviewed weekly to maintain consistent assessment quality and inter-rater reliability.

Assessment procedure. For this study, all participants were assessed using parent-reported questionnaires, BeDevel-I/P, and diagnostic instruments. All parent-reported questionnaires were sent by ground mail so that they could be completed prior to the assessment visit. Both BeDevel-I/P and the diagnostic assessments were conducted for every participant, regardless of whether they were suspected to have ASD. BeDevel-P and the Autism Diagnostic Observation Schedule (ADOS) were conducted on children, and BeDevel-I, the Korean version of Childhood Autism Rating Scale (K-CARS), and the Autism Diagnostic Interview-Revised (ADI-R) were conducted on parents/caregivers by qualified, trained examiners. To avoid a potential effect of the assessment order, especially the confounding effect of situation familiarity for children and of the stimulation of retrieval in parental interviews, the sequence of BeDevel-P and ADOS on children and BeDevel-I and ADI-R/K-CARS were randomized. All interviews and observational assessments were videotaped and used for scoring and verifying inter-rater reliability. Family history and developmental history were also collected for diagnostic confirmation. The examiners were blinded to the diagnostic characteristics of the participants.

Clinical best estimate diagnoses were determined by qualified diagnostic assessment staff based on all available information collected, and confirmed by experienced clinicians, including two licensed child psychiatrists and special education professionals who have expertise in diagnosing/assessing ASD children. In the 18–23 month group, we refer to those with a high possibility of ASD as

Table 2. Construct of BeDevel-Play

DSM-5 criteria	Behavior	9 months	12 months	18 months	24 months	36 months	Task/activity
A-1. Social reciprocity	Response to name	V	V	V	V	V	Call the child by name
	Sharing interest		V	V	V	V	Point at a sticker
			V	V	V	V	Observe the child's pointing at a sticker
			V				Instruct the child to imitate an action—flipping over a cup
				V			Instruct the child to imitate an action—blowing a pinwheel
					V	V	Instruct the child to imitate an action—sliding a doll
		V	V	V	V	V	Observe the child's sharing interests
	Initiating social interaction		V	V	V	V	Present a toy and observe the child's requesting behavior
	Social reference		V	V			Present an unexpected situation—showing a moving toy
				V	V	V	Present an unexpected situation—hitting blocks to fall
					V	V	Present an unexpected situation—being painful/frustrated
	Joint-attention		V	V	V	V	Point at a sticker
			V	V	V	V	Observe the child's pointing at a sticker
	Imitation of actions		V				Instruct the child to imitate an action—flipping over a cup
				V			Instruct the child to imitate an action—blowing a pinwheel
					V	V	Instruct the child to imitate an action—sliding a doll
	Social play				V		Initiate a catch-ball and observe the child's reciprocal engagement
					V		Initiate a tug-of-war and observe the child's reciprocal engagement
					V		Initiate a *Co–co–co game (Korean traditional game) and observe the child's reciprocal engagement
						V	Initiate an imaginative play and observe the child's reciprocal engagement
	Verbal communication				V	V	Give verbal instructions
						V	Initiate a conversation and observe the child's reciprocal engagement
A-2. Nonverbal communication	Eye contact	V	V	V	V	V	Observe the child's eye-contact
	Social smile	V					Talk to and smile at a child
			V	V	V		Tickle the child
			V	V	V		Play peekaboo with the child
					V	V	Compliment the child
					V	V	Say goodbye to the child
	Facial expression	V	V	V	V	V	Observe the child's facial expressions
	Gesture		V	V	V	V	Present a toy and observe the child's requesting behavior
					V	V	Say goodbye to the child
		V	V	V	V	V	Observe the child's gestures
	Pointing (joint-attention)		V	V	V	V	Observe the child's pointing at a sticker
			V	V	V	V	Present a toy and observe the child's requesting behavior
	Response to pointing		V	V	V	V	Point at a sticker
	Understanding facial expressions		V	V			Present an unexpected situation—showing a moving toy
				V	V	V	Present an unexpected situation—hitting blocks to fall
					V	V	Present an unexpected situation—getting painful/frustrated
					V	V	Initiate a hiding game
	Understanding gestures				V	V	Initiate a hiding game
					V	V	Say goodbye to the child

(Continues)

Table 2. Continued

DSM-5 criteria	Behavior	9 months	12 months	18 months	24 months	36 months	Task/activity
A-3. Social relationship	Social interest	V	V	V	V	V	Observe the child's sharing interests
	Social response	V	V	V	V	V	Call the child by name
		V					Talk to and smile at a child
			V	V			Tickle the child
			V	V			Play peekaboo with the child
					V	V	Compliment the child
					V	V	Say goodbye to the child
B. Interest and behavioral characteristics	Stereotyped/repetitive behaviors	V	V	V	V	V	
	Insistence on sameness/adherence to routines/ restricted pattern of behavior	V	V	V	V	V	
	Restricted fixated interest	V	V	V	V	V	
	Unusual sensory interest and hyper/ hyposensitivity	V	V	V	V	V	

"ASD" for the sake of convenience, though that age is sometimes considered too young for diagnostic confirmation. In addition, any uncertain results were discussed at weekly meetings in which the videotaped assessments and results were reviewed and discussed by the research team staff to confirm the diagnosis.

Measures. *ADOS and Autism Diagnostic Observation Schedule-2.* ADOS and Autism Diagnostic Observation Schedule-2 (ADOS-2) [Lord, Luyster, Gotham, & Guthrie, 2012] are a standardized semi-structured assessment using play-based methods to determine the presence of ASD symptoms. It assesses verbal and nonverbal communication, social interactions and relatedness, play, and imagination. ADOS consists of modules 1 to 4, and ADOS-2 has module T for children younger than 30 months old. Each module has a different combination of activities, but all modules consist of planned social interactions to encourage the examinee's social initiations, social responses, and opportunities to participate in communication. After the assessment, the child receives scores in various domains: communication, social interaction, combined communication and interaction, play and imaginary, and restricted and repetitive behaviors (RRBs) in ADOS; and social interaction and communication, RRBs, total score, and comparison score in ADOS-2. Results fall into three categories: autism, ASD, and non-spectrum combined. The Korean translation of the ADOS-2 [Yoo et al., 2017] was approved by its publisher, Western Psychological Services. This study used ADOS for the children who participated in this study before ADOS-2 was published in South Korea, and ADOS-2 for others; therefore, the ADOS scores were rescored based on the ADOS-2 algorithm for results analysis. For this study, any children aged 18–23 months evaluated as having "mild-to-moderate concerns" and "moderate-to-severe concerns" in module T that had equivalent concerns reported by parents in the ADI-R were classified as having ASD.

Autism Diagnostic Interview-Revised. ADI-R [Rutter, LeCouteur, & Lord, 2003] is a semi-structured interview for caregivers, which assesses the child's communication, social development, play, and restricted, repetitive, and stereotyped behaviors. The administrator gathers information from the interviewees based on the answers to 123 questions that are scored between 0 (socially appropriate level) and 3 (very severe). The interview questions regard three diagnostic domains: social interaction, communication, and repetitive and stereotyped behavior. Each domain has a diagnostic criterion for autism, and all three domains must have scores exceeding this criterion in order for a positive autism diagnosis. ADI-R is applicable to children with a mental age of at least 2 years, but was administered to participants under the age of 2 years to obtain a diagnostic impression and detailed developmental history. The Korean translation of the ADI-R [Yoo et al., 2007] was approved by its publisher, Western Psychological Services.

Korean version of Childhood Autism Rating Scale. The Korean version of the Childhood Autism Rating Scale (CARS) [Schopler, Reichler, & Rochen, 1988] consists of 15 items to measure the presence and severity of symptoms of pervasive developmental disorders. Both the parent's report and the clinician's observation of the child's behavior are used to rate each item. It includes questions in different areas such as socialization, verbal and nonverbal communication, emotional responses, restricted or unusual interests or behaviors, and sensory sensitivities. Each of 15 items is scored from 1 (no impairment observed or reported) to 4 (severe impairment). The child

can be classified as having mild, moderate, or severe autism. Based on the re-adjustment of the cutoff K-CARS score in the Korean population, we used 24 as the cutoff for ASD diagnosis [Kim & Park, 1995; Kwon et al., 2017].

Social Communication Questionnaire. The SCQ [Rutter et al., 2003] consists of two forms: the "Current Form" and the "Lifetime Form," each of which includes 40 items rated as either "yes" or "no" by caregivers of individuals aged 24 months or above. Items are to assess behaviors of an individual in three domains: social interaction, language and communication, and restricted repetitive behavior. According to a standardization study conducted in Korea by Kim et al. [2015], a cutoff score of 10 is most effective at detecting children under 47 months of age who are at a high risk of ASD. Therefore, this study used that criterion for diagnosis using the SCQ.

Social Response Scale Second Edition. The Social Response Scale Second Edition (SRS-2) [Constantino & Gruber, 2012] is a questionnaire completed by parents/caregivers to screen children at risk of ASD. It consists of measures of the child's social interactions, communication, and stereotyped behaviors. SRS-2 offers four forms: school-age (4–18 years), preschool (2.5–4.5 years), adult (19 years and older), and adult self-report. It has a total of 65 questions, which can be rated from 1 (not at all) to 4 (almost always). If the child scores greater than 75, they are classified as being at high-risk of ASD. We translated and back translated the SRS-2 preschool form and received permission to use it through a contract with the SRS-2 copyright holder, WPS Publish. As SRS has not been standardized in Korea, reliability analysis was performed; Cronbach's α value was 0.96.

Vineland Adaptive Behavior Scales, Second Edition. Vineland Adaptive Behavior Scales, Second Edition [Volkmar, 2013] were designed to address adaptive behavior, personal and social skills necessary for everyday independent living across the life span (birth to 90 years). VABS can be administered to parents, teachers, and other caretakers using survey interview and rating forms. For individuals with ASD, it can be used to determine eligibility for services, planning intervention programs, and tracking and reporting progress following interventions. Domains include communication, daily living skills, socialization, motor skills, and maladaptive behaviors (internalizing and externalizing behaviors). Behaviors are rated on a 0–2 rating scale, with 0 representing a skill that is not used by the individual, 1 representing a skill used sometimes, and 2 representing a skill used most of the time.

Sequenced Language Scale for Infants. Sequenced Language Scale for Infants (SELSI) [Kim, Kim, Yoon, & Kim, 2003] examines development in the areas of receptive language and expressive language in children aged 5–36 months. It gives developmental age, percentile, and standard deviation of each area. It consists of 56 questions and takes 10–20 min to complete, depending on the age of the child. It can also be used to evaluate language development, and provides an approximate development level for children with language delays.

Statistical Analyses

We used Cohen's kappa (k) to check the consistency between the diagnosis of ASD and corresponding response from BeDevel [Watson & Petrie, 2010]. A kappa statistic of 0.21 or higher was considered to have at least fair strength of agreement, and items with the kappa statistic of 0.20 or less were considered to have "slight" or "poor" agreement [Landis & Koch, 1977]. The chi-squared test (χ^2) was also used to check for significant differences between score distributions in the ASD and non-ASD groups (developmental delay without ASD group and typical development group) for each item in BeDevel-I and BeDevel-P. Next, Cronbach's α values were calculated to assess the internal consistency of BeDevel-I/P when nonsignificant items (based on the results of χ^2 test) were removed. We also used receiver operating characteristic curves (ROC) to explore the best criteria to screen for "high risk" individuals. Sensitivity, specificity, positive predictive values, and negative predictive values were calculated to compare several different screening criteria. Lastly, we used kappa values to assess the levels of agreement between BeDevel-I/P and the other measures (ADOS-2, ADI-R, K-CARS, SCQ, and SRS). Statistical analyses were performed using SPSS Statistics 23 (IBM Corp., Armonk, NY, USA). The confidence interval was set at 95%.

Results

Participant Characteristics

A total of 155 children ranging from 18 to 42 months of age participated in the study. The samples of 9–11 months and 12–17 months were too small to be statistically tested and thus were not included in the analysis of the present study. The average age was 31.61 months (SD = 7.52), and the majority were male (n = 109, 70.3%). The number of children in each age range was 18–23 months n = 30; 24–35 months n = 62, and 36–42 months n = 63. Participants were classified by the clinical best estimate diagnosis and divided into three groups: ASD (n = 75), developmental delay without ASD (DD; n = 25), and children with typical development (TD; n = 55). Participants whose language level was below a standard deviation of −1 in SELSI and those for whom the standard score was below 70 in any area of the VABS were considered DD. Of the 155 toddlers who participated in the study, three were previously diagnosed with ASD, and two were diagnosed with communication disorder. Among the 150 participants

who had not been diagnosed in advance, 89 were entered into the study because of concerns about language, social skills, and/or overall developmental delay. The other 51 participants had received a clinical impression of ASD from a child psychiatry clinic, department of rehabilitation, or development center without a confirmed diagnosis. A total of 61 children were entered into the study without any developmental concerns. The demographic characteristics of the participants are summarized in Table 3.

Concurrent Validity

When the agreement between individual item responses and ASD diagnosis was computed using Cohen's kappa, all items in BeDevel-P showed kappa values of 0.21 or higher for all age groups (kappa value range = 0.27–0.80 for individuals aged 18–23 months, 0.27–0.65 for 24–35 months, and 0.38–0.71 for 36–42 months). In BeDevel-I, *Social Smile* (k = 0.13) and *Insistence on Sameness/Adherence to Routines/Restricted Patterns of Behavior* (k = 0.07) in the 18–23 months old group showed kappa values less than 0.20. The rest of the 13 items showed fairly high kappa values, ranging 0.27–0.93. For the 24–35 month group, *Social Smile* (k = 0.03), *Social Interest and Social Relationship* (k = 0.18), *Insistence on Sameness/Adherence to Routines/ Restricted Patterns of Behavior* (k = 0.18), and *Restricted, Fixated Interest* (k = 0.14) showed kappa values less than 0.20. The rest of the 13 items showed fairly high kappa values, ranging 0.24–0.67. For the 36–42 month group, *Social Smile* (k = 0.07) and *Restricted, Fixated Interest* (k = 0.12) showed kappa values less than 0.20. The rest of 16 items showed fairly high kappa values, ranging 0.25–0.62. In the present study, we named the items that showed kappa values of 0.21 or higher *Primary Items* and those with kappa values of 0.20 or less *Secondary Items*. That is, *Primary Items* are those with a higher degree of discrimination in the screening. Although a kappa value of 0.41 or higher is generally accepted as having "moderate" strength of agreement, the questions related to the core risk signs for ASD, replicated multiple times in previous studies, were eliminated when we applied that criteria; therefore, we decided to include items with kappa values of 0.21 or higher in primary items as well. Kappa values of individual items are presented in Table 4.

Whether the selected *Primary Items* could sufficiently discriminate the ASD group from the non-ASD groups was confirmed using the chi-square test. Differences in secondary item responses were not significant in discriminating the ASD group from the non-ASD group ($P > 0.05$, range = 0.056–0.309). In contrast, primary items showed high overall discrimination. For BeDevel-I, the primary items were significant at 0.000–0.032 levels in the 18–23 months, 0.000–0.021 in the 24–35 month group, and 0.000–0.018 in the 36–42 month group. The primary items of the BeDevel-P were also significant, at 0.00–0.032 in 18–23 month group, 0.000–0.021 in the 24–35 month group, and 0.000–0.002 in the 36–42 month group.

Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value

We generated criteria items using ROC analysis, to confirm the cutoff numbers of the *Primary Items*. Based on the results of ROC analysis, the *Cutoff Number of Items* was 2/13 in the 18–23 month, 3/13 in 24–35 month, and 6/16 in the 36–42 month group in BeDevel-I. For BeDevel-P, cutoff numbers were 9/15 in the 18–23 month, 11/18 in 24–35 month, and 9/19 in 36–42 month groups. With this new criterion, both BeDevel-P and BeDevel-I showed satisfactory levels of sensitivity with a range of 83.33%–100%, specificity, 81.25%–100%, positive predictive value (PPV), 80.65%–100%, and negative predictive value (NPV), 83.87%–100%.

When the ASD group was compared with the DD group, sensitivity in both 24–35 month and 36–42 month groups was 90%, which was similar to when the ASD group was compared with TD group. For individuals aged 24–35 months, specificity was 91.67%, which indicates a high ability to discriminate between two groups. Specificity in the 36–42 month group was 76.92%. Sensitivity of BeDevel-P was 80.9% in 24–35 month group and 81.40% in 36–42 month group. Specificity was 66.67% in the 24–35 month group and 53.85% in the 36–42 month group. If children were screened in both BeDevel-I and BeDevel-P, sensitivity range was 76.67–93.33, which was slightly lower, while specificity was higher, ranging from 91.67 to 100. PPV and NPV were moderate at 95.83%–100% and 81.58%–93.33%, respectively. Results are presented in Table 5.

Internal Consistency and Item Discrimination

When the internal consistency of the selected primary items was assessed through Cronbach's α, primary items from BeDevel-P showed high internal consistency, with the average α value of 0.88 (range 0.866–0.889) for 18–23 month, 24–35 month, and 36–42 month groups. The change in α values after removing each primary item was rather small, indicating that no item seriously affects the internal consistency (range: 0.843–0.875 for 18–23 months, 0.843–0.877 for 24–35 months, and 0.876–0.889 for 36–42 months). In addition, correlations between individual primary items and BeDevel-I as a whole were calculated to test item discrimination; all items showed positive correlations, ranging from 0.273 to 0.843 for the 18–23 month group, 0.287–0.724 for the 24–35 month group, and 0.372–0.689 for the 36–42 month group.

Primary items from BeDevel-P showed high internal consistency as well, with the average α value of 0.953 (range: 0.947–0.959) for the 18–23 month, 24–35 month,

Table 3. General Characteristics of Subject

	18–23 month			24–35 month				36–42 month			
	ASD (*n* = 15) Mean (*SD*)	TD (*n* = 15) Mean (*SD*)	*P*	ASD (*n* = 30) Mean (*SD*)	DD (*n* = 12) Mean (*SD*)	TD (*n* = 20) Mean (*SD*)	*P*	ASD (*n* = 30) Mean (*SD*)	DD (*n* = 13) Mean (*SD*)	TD (*n* = 20) Mean (*SD*)	*P*
Age	20.07 (1.71)	19.73 (1.79)	0.606	29.97 (3.02)	30.58 (3.58)	29.2	0.531	38.8 (2.01)	39.23 (1.83)	38.95 (1.82)	0.79
Gender (M:F)	13:7	2:8		21:9	9:3	12:8		24:6	9:4	14:6	
ADOS-2[a] Module T	*n* = 15	*n* = 15		*n* = 15	*n* = 5	*n* = 14					
SA	17.33 (2.47)	2.87 (2.17)	0.000	14.47 (3.56)	7.6 (4.72)	4.57 (4.03)	0.000	–	–	–	–
RRB	1.40 (1.18)	0.27 (0.59)	0.003	1.87 (1.41)	1.00 (1.00)	0.07 (0.27)	0.000	–	–	–	–
ADOS-2 module 1				*n* = 14	*n* = 7	*n* = 2		*n* = 27	*n* = 10	*n* = 3	
SA	–	–	–	14.71 (2.70)	4.14 (2.19)	7.50 (0.71)	0.000	15.44 (2.84)	6.70 (4.42)	2.00 (2.64)	0.000
RRB	–	–	–	2.50 (1.95)	0.86 (1.07)	1.50 (2.12)	0.145	2.70 (1.96)	0.90 (1.73)	0.33 (0.58)	0.013
ADOS-2 module 2				*n* = 1	*n* = 0	*n* = 4		*n* = 3	*n* = 3	*n* = 17	
SA	–	–	–	4.00	–	2.25 (0.96)	–	10.00 (4.36)	3.00 (1.73)	2.29 (2.02)	0.000
RRB	–	–	–	1.00	–	1.00 (0.82)	–	3.00 (0.00)	0.33 (0.58)	0.65 (0.79)	0.000
ADI-R											
SI	14.53 (3.18)	4.20 (3.80)	0.000	14.87 (3.73)	5.92 (3.60)	4.80 (2.91)	0.000	15.33 (4.32)	7.31 (4.35)	3.05 (2.58)	0.000
NC	9.33 (2.74)	1.73 (1.91)	0.000	8.97 (2.90)	3.00 (2.52)	2.65 (2.41)	0.000	9.47 (2.54)	4.70 (3.16)	1.00 (0.97)	0.000
RRB	3.40 (1.64)	1.00 (1.07)	0.000	3.9 (2.16)	1.83 (1.47)	1.25 (1.16)	0.000	5.77 (1.85)	1.92 (1.44)	1.20 (1.06)	0.000
K-CARS	30.83 (3.81)	15.93 (1.29)	0.000	31.3 (4.47)	21.83 (5.23)	18.13 (3.78)	0.000	31.83 (4.17)	19.77 (5.05)	16.63 (1.56)	0.000
SCQ	15.47 (4.93)	7.00 (4.33)	0.000	15.33 (5.47)	7.41 (4.94)	7.00 (3.23)	0.000	18.30 (5.74)	10.62 (4.99)	4.25 (2.63)	0.000
SRS	63.80 (11.29)	49.07 (5.96)	0.000	62.43 (9.80)	49.33 (9.86)	46.85 (5.54)	0.000	62.27 (10.24)	59.31 (8.32)	45.10 (6.81)	0.000
VABS	76.07 (6.34)	97.40 (9.27)	0.000	72.97 (11.01)	81.36 (8.94)	100.85 (15.79)	0.000	64.37 (10.42)	73.07 (14.13)	100.75 (13.56)	0.000
SELSI											
Receptive (month)	10.50 (5.23)	24.77 (3.24)	0.000	16.79 (8.34)	23.25 (7.02)	28.25 (3.08)	0.000	18.88 (6.87)	24.18 (5.85)	–[b]	0.000
Expressive (month)	8.40 (4.60)	19.77 (2.62)	0.000	15.60 (6.44)	15.67 (6.12)	27.35 (5.18)	0.000	18.00 (8.05)	22.80 (7.35)	–[b]	0.000
Refer (%)											
Outpatient	46.7	6.7		46.7	41.7	5.0		46.7	7.7	0.0	
Local hospital, community center	0.0	6.7		13.3	8.3	15.0		10.0	23.1	5.0	
Nursery	6.7	20.0		0.0	8.3	30.0		3.3	23.1	15.0	
Online, poster	20	0.0		16.7	16.7	10.0		6.7	15.4	35.0	
Introduction form other	26.7	66.7		23.3	25.0	40.0		33.3	30.8	45.0	

Abbreviations: ADOS-2, Autism Diagnostic Observation Schedule and Autism Diagnostic Observation Schedule-2; SA, social affect; RRB, restricted repetitive behavior; ADI-R, Autism Diagnostic Interview-Revised; SI, social interaction; NC, nonverbal communication; K-CARS, Korean Child Autism Rating Scale; SCQ, Social Communication Questionnaire; SRS, Social Response Scale; VABS, Vineland Adaptive Behavior Scales; SELSI, Sequenced Language Scale for Infants.

[a]Because ADOS-2 was not published in Korea until July, 2017, the results collected before that were assessed by ADOS and thus were rescored based on the algorithm of ADOS-2.

[b]The upper age limit in SELSI is 36 months, and all 36–42 month olds in the TD group showed the highest equivalent age score available when assessed by the instrument.

Table 4. Concurrent Validity by Individual Item

Item description	18 month		24 month		36 month	
	Kappa value	χ^2	Kappa value	χ^2	Kappa value	χ^2
BeDevel-I						
Response to name	0.53	10.91**	0.41	12.40***	0.38	11.26**
social smile	0.13	2.14	0.03	1.08	0.07	2.27
Social reference	0.60	12.86***	0.41	11.81**	0.42	12.89***
Imitation of actions	0.33	6.00*	0.27	8.27**	0.45	13.68***
Pointing	0.80	20.00***	0.58	23.57***	0.38	12.03**
Response to pointing	0.53	10.91**	0.44	16.01***	0.42	13.68***
Sharing interest	0.73	16.43***	0.48	16.66***	0.55	20.35***
Conversation	–	–	–	–	0.56	20.20***
Eye contact	0.47	9.13**	0.35	8.37**	0.45	14.03***
Facial expression	0.40	6.14*	0.24	5.98*	0.35	10.47**
Gesture	0.60	11.63*	0.28	5.35*	0.61	25.40***
Social interest and social relationship	0.27	4.62*	0.18	3.64	0.25	5.64*
Level of Play	–	–	0.51	17.15***	0.62	24.72***
Interest in peers	–	–	0.54	19.69***	0.62	25.25***
Stereotyped/repetitive behaviors	0.93	26.25***	0.67	30.25***	0.62	24.08***
Insistence on sticking to routines and rituals	0.07	1.03	0.18	2.86	0.42	12.32***
Limited interest	0.47	9.13**	0.14	2.23	0.12	1.27
Sensory characteristics	0.53	9.60**	0.28	5.35*	0.55	17.18***
BeDevel-P						
Social reference—hitting blocks	0.53	8.89**	0.42	10.87**	0.49	15.15***
Response to name	0.67	13.39***	0.58	21.16***	0.40	9.91**
Understanding simple words	–	–	0.58	20.86***	0.59	22.45***
Social reference—a moving toy/frustrated	0.80	19.29***	0.52	16.74***	0.71	32.26***
Requesting—a distant toy	0.60	11.63**	0.52	17.60***	0.50	16.04***
Joint-attention-response	0.27	4.66*	0.51	21.11***	0.38	12.03**
Joint-attention-initiation	0.73	17.37***	0.52	17.09***	0.58	21.99***
Imitation of action	0.80	19.29***	0.27	5.34*	0.47	14.30***
Social reference-hiding game	–	–	0.55	20.19***	0.62	24.26***
Social play	0.60	11.00**	0.55	22.25***	0.56	21.34***
Back-and-forth Conversation	–	–	–	–	0.56	24.73***
Smile response—peek-a-boo/compliment	0.60	11.00**	0.49	18.63***	0.53	19.17***
Smile response—goodbye	–	–	0.43	12.46***	0.47	15.65***
Eye contact	0.80	19.29***	0.62	24.35***	0.56	19.55***
Use of facial expression	0.53	8.57**	0.65	26.11***	0.68	29.32***
Use of gesture	0.60	10.96**	0.59	22.21***	0.69	36.63***
Sharing interest	0.67	13.39***	0.52	16.47***	0.62	24.08***
Social relationship	0.73	16.13***	0.61	23.78***	0.53	17.90***
Restricted, repetitive patterns of behavior, interests, or activities	0.33	4.69*	0.55	19.15***	0.49	15.47***

*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

and 36–42 month groups. The change in α values after removing each primary item was rather small, again indicating that no item seriously affects the internal consistency (range: 0.939–0.949 in 18–23 month, 0.947–0.954 in 24–35 month, and 0.955–0.960 in 36–42 month groups). In addition, correlations between individual primary items and BeDevel-P were calculated to test item discrimination; all items showed generally strong positive correlations, ranging from 0.451 to 0.896 for the 18–23 month group, 0.352–0.866 for the 24–35 month group, and 0.475–0.852 for the 36–42 month group.

Comparison with the existing measurements for ASD

We analyzed the agreement of the results of the BeDevel-P and BeDevel-I with those of preexisting screening instruments for ASD, including SCQ, SRS, ADOS, ADI-R, and K-CARS. For SCQ and K-CARS, we applied the following cutoff scores for the Korean population, which were previously validated by our team: 10 for SCQ (instead of 15 in original version) for toddlers and children younger than 48 months old, and 24 for K-CARS (instead of 30 in original version) [Kim et al., 2015; Kwon et al., 2017]

In BeDevel-I and BeDevel-P of all age levels, agreements with ADOS were substantial (k = 0.677–1.000) and the level of agreement was almost perfect for individuals aged 18 months. Both showed substantial levels of agreements with ADI-R for 36–42 month and 18–23 month (k = 0.685–0.777) groups, and had a moderate level of agreements in 24–35 month groups (k = 0.419–0.533). A higher level of agreement was apparent when the diagnostic results from the social interaction and communication

Table 5. Sensitivity, Specificity, PPV, and NPV Based on BeDevel-I and BeDevel-P

	18–23 months	24–35 months		36–42 months	
	ASD vs. non-ASD (n = 15 vs. n = 15)	ASD vs. DD (n = 30 vs. n = 12)	ASD vs. TD (n = 30 vs. n = 32)	ASD vs. DD (n = 30 vs. n = 13)	ASD vs. TD (n = 30 vs. n = 33)
BeDevel-I					
Number of primary items	13	13		16	
Cut off numbers of item	2	3		5	
Agreement of diagnosis and BeDevel	100	90.48	87.10	86.05	88.69
Sensitivity (%)	100	90.00	90.00	90.00	90.00
Specificity (%)	100	91.67	84.38	87.88	76.92
PPV (%)	100	96.43	84.38	90.00	87.10
NPV (%)	100	78.57	90.00	76.92	90.63
BeDevel-P					
Number of primary items	15	18		19	
Cut off numbers of item	9	11		9	
Agreement of diagnosis and BeDevel	93.33	80.95	85.48	81.40	87.30
Sensitivity (%)	93.33	86.67	86.67	93.33	93.33
Specificity (%)	93.33	66.67	84.38	53.85	81.82
PPV (%)	93.33	86.67	83.87	82.35	82.35
NPV (%)	93.33	66.67	87.10	11.18	93.10
BeDevel-Total[a]					
Agreement of diagnosis and BeDevel	96.67	80.95	87.10	86.05	90.48
Sensitivity (%)	93.33	76.67	76.67	83.33	83.33
Specificity (%)	100.00	91.67	96.88	92.31	96.97
PPV (%)	100.00	95.83	95.83	96.15	96.15
NPV (%)	93.33	61.11	81.58	70.59	86.49

[a]BeDevel-Total: Screened on both BeDevel-I and BeDevel-P.

areas only of the ADI-R were compared to BeDevel (k = 0.613–1.000). Agreements for all ages were substantial for both instruments when compared to K-CARS (k = 0.742–0.933). For SCQ, the 24–35 month-old group and the 36–42 month-old group showed a substantial level of agreement (k = 0.623–0.710) with both BeDevel-I and BeDevel-P. The agreement with SRS was relatively low compared to that of other instruments, but was still fair to substantial (k = 0.395–0.649). The agreement between BeDevel-P and BeDevel-I showed 96.7% of agreement in 18–23 month (k = 0.933), 75.8% in 24–35 month (k = 0.516), and 79.5% in 36–42 month (k = 0.588) groups. Results are presented in Table 6.

Discussion

In our pilot analyses of BeDevel-I and BeDevel-P, both showed a satisfactory level of concurrent validity, sensitivity, specificity, positive and negative predictive values, and agreement with existing instruments for diagnosing ASD. These results suggest that both can be useful instruments for screening ASD-related behavior in children up to 42 months of age. It was relatively easy for semi-professional raters to reach sufficient inter-rater reliability with a modest level of training. We hypothesize that this might be due to the fact that BeDevel translates the diagnostic criteria of ASD as defined by DSM-5 criteria into observable, specific, real-life behaviors in young children, whereas DSM-5 criteria define the core components of ASD-related behaviors in an operational manner [Rogers, Goddard, Hill, Henry, & Crane, 2016; Skellern, Schluter, & Mcdowell, 2005].

In the item analyses, our primary items (those that showed good correlations with the diagnosis) were generally comparable with previous studies looking for early signs of ASD, based on Cohen's kappa values. *Response to Name, Use of Facial Expression, Imitation*, and *Joint Attention-Response* showed relatively low kappa value of 0.21–0.40. When we calculated sensitivity, specificity, PPV, NPV, and agreement of diagnosis and BeDevel using the items with kappa values of 0.41 or higher, the results were similar to what we got with our primary items (range: 83.87%–100% in sensitivity, 81.82%–100% in specificity, 82.35%–100% in PPV, and 84.38%–100% in NPV, 83.87%–100% in agreement of diagnosis and BeDevel). However, we decided not to get rid of the items because those behaviors have been repeatedly and consistently replicated as early significant signs of ASD in multiple researches [Baird et al., 2000; Barbaro & Dissanayake, 2010; Boyd et al., 2010; Macaril et al., 2012; Stenberg et al., 2014]. Those items showed high correlation with ASD diagnosis in χ^2 test ($P < 0.05$). Also, as the tendency of relatively lower kappa was more pronounced in BeDevel-I, we assumed that it might be influenced by individual parents' ability to accurately understand what the interview questions are asking, as well as their understanding of child development and should be re-examined through the expanded samples in further study. In the

Table 6. Agreement with Existing Instruments

Age	Tool		ADOS	ADI-R	ADI-R (s + c)	K-CARS	SCQ	SRS[a]
18–23 month (n = 30)	BeDevel-I	Agreement (%)	100	86.70	96.70	96.70	–	–
		Kappa value	1.00	0.73	0.93	0.93	–	–
	BeDevel-P	Agreement (%)	93.40	86.70	96.70	90	–	–
		Kappa value	0.87	0.73	0.93	0.80	–	–
	BeDevel-total	Agreement (%)	96.70	90	100	93.30	–	–
		Kappa value	0.93	0.80	1.00	0.87	–	–
24–35 month (n = 62)	BeDevel-I	Agreement (%)	88.70	72.60	85.50	87.10	85.40	68.40
		Kappa value	0.77	0.46	0.71	0.74	0.71	0.40
	BeDevel-P	Agreement (%)	83.80	71	80.70	88.70	80.60	73.70
		Kappa value	0.68	0.42	0.61	0.77	0.61	0.48
	BeDevel-total	Agreement (%)	85.50	79.10	88.70	87.10	85.50	79
		Kappa value	0.71	0.53	0.76	0.74	0.71	0.54
36–42 month (n = 63)	BeDevel-I	Agreement (%)	87.30	88.90	87.30	88.90	82.50	82.50
		Kappa value	0.75	0.78	0.76	0.78	0.65	0.65
	BeDevel-P	Agreement (%)	88.90	84.20	82.60	87.30	80.90	71.40
		Kappa value	0.78	0.69	0.65	0.75	0.62	0.44
	BeDevel-total	Agreement (%)	88.90	90.50	88.90	90.60	77.80	77.80
		Kappa value	0.78	0.81	0.78	0.82	0.57	0.54

Abbreviations: ADOS-2, Autism Diagnostic Observation Schedule and Autism Diagnostic Observation Schedule-2; ADI-R, Autism Diagnostic Interview-Revised; ADI-R(s + c); Autism Diagnostic Interview-Revised (Social domain and Communication domain); K-CARS, Korean Child Autism Rating Scale; SCQ, Social Communication Questionnaire; SRS, Social Response Scale.

[a]In 24–35 months, children over 30 months were included in the results analysis (n = 38).

BeDevel-I, the reliability of *Social Smile* (all age groups) and *Social Interest and Social Relationship* (in 24–35-month old individuals) were limited, and Repetitive and Restricted Behaviors and Interests items showed inconsistent results, although the kappa values of BeDevel-P were acceptable for all items. This might be attributed to the fact that awareness of *Social Smil*e is affected by cultural differences. Although social smiling is a universal social behavior, cultural differences have been found regarding the duration and frequency of social smiles [Wörmann, Holodynski, Kärtner, & Keller, 2012]. Thus, it is possible that cultural differences have made Korean parents less sensitive at detecting differences in social smiles in their children, and that direct observation is more informative than parents' reports for this item. Also, because 24 month of age is a time of transition into daycare in Korea, behaviors pertaining to *Social Interest and Social Relationshi*p begin to change as children transition from relationships primarily with caregivers/parents to those outside of the family. Thus, it might be difficult for caregivers/parents to judge the age-appropriateness of their children's social interest and behavior.

Despite the usefulness of direct observation through BeDevel-P for some items such as *Social Smile* and *Use of Facial Expression*, BeDevel-P revealed very low specificity for diagnosis, while sensitivity was generally satisfactory. We introduced the *Cutoff Number of Item* method, in which a subject is regarded to have a high risk of ASD when he/she meets high-risk criteria in multiple risk items. Both BeDevel-P and BeDevel-I showed high sensitivity as well as specificity, along with high PPV and NPV. This implies that the combination of risk items and cutoff number methods might be useful when applied in screening in a community and clinical setting. However, the specificity that distinguishes between ASD and DD was not excellent. This may be partially due to the fact that that toddlers and young children with ASD and DD usually have language delays and limitations or difficulties in using nonverbal communication skills [Osterling, Dawson, & Munson, 2002; Ventola et al., 2007]. As the sample size of the DD group was relatively small, further analyses with larger sample sizes would be needed to clarify the specificity issue. When combining BeDevel-I and BeDevel-P in screening, specificity both in individuals aged 24–35 months and 36–42 months was higher. We recommend that those who are screened using both BeDevel-I and BeDevel-P are classified as at-risk of ASD in order to differentiate ASD and DD. Although direct observation is more useful for assessing some domains, it may be necessary to place more emphasis on parent interviews to increase specificity.

Interestingly, we found different results for each item depending on age. For example, response to name and social-referencing-related items in BeDevel-P showed higher discrimination power in the 18–23 month group than in the 24–35 month and 36–42 month groups. Such differences in social–sensitivity-related items across age groups can be understood by considering that awareness and fear of unfamiliar environments develop around 12 months and last until 24 months, the time when toddlers grow more independent [Jones, Greenberg, & Crowley, 2015]. On the other hand, social behaviors such as JA behaviors and imitation were less sensitive in differentiating ASD from other groups for the 24–35 month and 36–42 month groups compared to

the 18–23 month group in BeDevel-I. Training effect may be one explanation, given that such social behaviors often become the primary targets of early intervention when children show difficulties [Zwaigenbaum et al., 2015]. Considering that the display of social behaviors varies across age, selecting primary items and deciding screening criteria by age group is appropriate. Diagnostic stability of ASD in 18 months has been debated. For example, some reports suggested that the diagnoses of ASD at the average age of 19 months were maintained in the follow-up at an average age of 37 months, while subjects showed improved in Social Affect domain of the ADOS [Guthrie, Swineford, Nottke, & Wetherby, 2013]. Another study found that diagnosis at 18–23 months of age should be confirmed through follow-up assessments because some symptoms may not be apparent at a certain age, and some behaviors might be affected by age [Baird et al., 2000; Macaril et al., 2012; Stenberg et al., 2014]. Further analyses with larger sample sizes would be needed to clarify the characteristics of behavior in 18 months of age.

As for the RRB domain, diagnostic validity was higher among older children in BeDevel-P. *Insistence on Sameness/Adherence to Routines/Restricted Patterns of Behavior* were only significant in the 36–42 month group and not for the younger groups, and *Restricted, Fixated Interest* was not significant, despite the fact that those are recognized as important diagnostic features of ASD. One possible explanation is that such behaviors can be observed in the developmental trajectories of both TD and DD children at a young age, and some children with mild ASD do not display RRBs [Arnott et al., 2010; Evans et al., 1997]. - Nevertheless, because children diagnosed with ASD often display more severe RRBs during infancy and toddlerhood compared to children with TD and DD, and because it remains stable over time, careful evaluation of RRB through parental interviews and child observation remains necessary [Joseph, Thurm, Farmer, & Shumway, 2013; Kim & Lord, 2010; Richler, Huerta, Bishop, & Lord, 2010; Watson et al., 2007]. As both typical and atypical behaviors can change over time through children's developmental trajectories, the age-specific design that considers developmental changes when assessing a child is particularly advantageous in BeDevel [Kasari, Gulsrud, Freeman, Paparella, & Hellemann, 2012; Zwaigenbaum et al., 2015].

When compared with other measures, BeDevel-I and BeDevel-P showed mostly high levels of agreements with ADOS and CARS with ADI-R, but levels of agreement were relatively low in the 24–35 month group. This might be associated with the fact that toddlers younger than 24 months, who have developmental difficulties, including intellectual disabilities, often show underdeveloped social reciprocity and communication skills, which might increase their chance of meeting the diagnostic criteria using ADI-R. In addition, RRB might appear after 24 months of age, suggesting that the ADI-R algorithm might not be suitable for young infants and toddlers [Rutter et al., 2003]. This explanation is supported by higher levels of agreements when only the social and communication domains of ADI-R were applied. Thus, if infants and toddlers under 24 months are identified as at-risk of ASD, immediate intervention and follow-up evaluation would be critical. BeDevel showed relatively low levels of agreement with SCQ, SRS, and other diagnostic tools. This may be attributed to the fact that SRS and SCQ are rated by caregivers/parents, without assessment by clinicians. It can be difficult to make objective judgments about social development characteristics and to correctly identify RRB related to ASD in their own children [Miller et al., 2011; Taylor, Vehorn, Noble, Weitlauf, & Warren, 2014]. Therefore, for more accurate screening, an interview measure administered by an interviewer who has received training and observational measures is important in the assessment.

In conclusion, BeDevel has many important advantages. It takes a top-down approach based on bottom-up process integration, and reflects opinions from FGI with various groups. It was developed as a comprehensive package that includes educational materials about ASD as well as a detailed guide for assessment. It is intended to be used with minimal training by populations who do not already have sufficient knowledge in or experience with ASD. In addition, BeDevel appears to be a feasible screening tool. Although it has relatively low specificity, it can be complemented by applying screening criteria in both BeDevel-I and BeDevel-P. From a screening perspective, the high sensitivity of BeDevel could be strength because it can identify more children at high risk. The application of both BeDevel-I and BeDevel-P is expected to enable accurate screening. Even a small amount of time, only 10–15 min for each tool, can produce relatively accurate results. It is necessary to select the age appropriate version, as each version has different contents and criterion. In addition, even though BeDevel is a screening tool, it allows for more detailed screening through interviews and observation. Therefore, we are considering using BeDevel as secondary step for screening after using questionnaires as a first step. Further data collection and analysis will be needed in the future to apply this step-by-step screening. The major limitation of this study is the small sample size of children in each different age group, especially toddlers under 18 months old. To overcome this limitation, more data collection is actively ongoing. Further, follow-up diagnosis of toddlers under 18 months old has yet to be confirmed.

Acknowledgments

We are grateful to the children and families who participated in this study. This work has been supported by the

Korea Health Industry Development Institute, Social Service R&D project (HI14C2719).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Althouse, L. A., & Stockman, J. A. (2006). Pediatric workforce: A look at pediatric nephrology data from the American Board of Pediatrics. The Journal of Pediatrics, 148(5), 575–576.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.

Arnott, B., McConachie, H., Meins, E., Fernyhough, C., Couteur, A. L., Turner, M., … Leekam, S. (2010). The frequency of restricted and repetitive behaviors in a community sample of 15-month-old infants. Journal of Developmental & Behavioral Pediatrics, 31(3), 223–229. https://doi.org/10.1097/DBP.0b013e3181d5a2ad

Austin, J., Manning-Courtney, P., Johnson, M. L., Weber, R. T., Johnson, H., Murray, D. S., … Murray, M. F. (2016). Improving access to care at autism treatment centers: A system analysis approach. Pediatrics, 137, 149–157.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., … Dowling, N. F. (2018). Prevalence of autism spectrum disorder among children aged 8 years - Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. MMWR Surveillance Summaries, 67(6), 1–23. http://doi.org/10.15585/mmwr.ss6706a1

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

Barbaro, J., & Dissanayake, C. (2010). Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: The Social Attention and Communication Study. Journal of Developmental & Behavioral Pediatrics, 31(5), 376–385.

Bisgaier, J. (2011). Access to autism evaluation appointments with developmental-behavioral and neurodevelopmental subspecialists. Archives of Pediatrics & Adolescent Medicine, 165(7), 673–674. https://doi.org/10.1001/archpediatrics.2011.90

Boyd, B. A., Odom, S. M., Humphreys, B. P., & Sam, A. M. (2010). Infants and toddlers with autism spectrum disorder: Early identification and early intervention. Journal of Early Intervention, 32(2), 75–98.

Bradshaw, J., Steiner, A. M., Gengoux, G., & Koegel, L. K. (2015). Feasibility and effectiveness of very early intervention for infants at-risk for autism spectrum disorder: A systematic review. Journal of Autism and Developmental Disorders, 45 (3), 778–794.

Chiang, C. H., Wu, C. C., Hou, Y. M., Chu, C. L., Liu, J. H., & Soong, W. T. (2012). Development of T-STAT for early autism screening. Journal of Autism and Developmental Disorders, 43 (5), 1028–1037. https://doi.org/10.1007/s10803-012-1643-4

Constantino, J. N., & Gruber, C. P. (2012). Social responsiveness scale–Second edition (SRS-2). Torrance: Western Psychological Services.

Corbisiero, S., Hartmann-Schorro, R. M., Riecher-Rössler, A., & Stieglitz, R. D. (2017). Screening for adult attention-deficit/-hyperactivity disorder in a psychiatric outpatient population with specific focus on sex differences. Frontiers in Psychiatry, 8, 115. https://doi.org/10.3389/fpsyt.2017.00115

Corsello, C., Cook, E., Levanthal, M. (2003). A screening instrument for autistic spectrum disorders. Society for Research in Child Development Biennial Conference.

Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., & Dawson, G. (2015). Long-term outcomes of early intervention in 6-year-old children with autism spectrum disorder. Journal of the American Academy of Child & Adolescent Psychiatry, 54(7), 580–587.

Evans, D. W., Leckman, J. F., Carter, A., Reznick, J. S., Henshaw, D., King, R. A., & Pauls, D. (1997). Ritual, habit, and perfectionism: the prevalence and development of compulsive-like behavior in normal young children. Child Development, 68(1), 58–68.

Fenikile, T. S., Ellerbeck, K., Filippi, M. K., & Daley, C. M. (2014). Barriers to autism screening in family medicine practice: A qualitative study. Primary Health Care Research and Development, 4, 1–11.

Glascoe, F. P. (2000). Evidence-based approach to developmental and behavioural surveillance using parents' concerns. Child: Care, Health and Development, 26, 137–149.

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. Journal of Child Psychology and Psychiatry, 54(5), 582–590. https://doi.org/10.1111/jcpp.12008

Jang. J., Sunwoo, H., Jeon, J., Bong, G., Roh, N., Jeong, S., Ro, D. &, Yoo. H. (2016). Child-care teachers' perception of autism spectrum disorder: Focus group interview analysis. Journal of the Korean Association for Persons with Autism, 16(2), 1–27.

Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. American Journal of Public Health, 105(11), 2283–2290. https://doi.org/10.2105/ajph.2015.302630

Joseph, L., Thurm, A., Farmer, C., & Shumway, S. (2013). Repetitive behavior and restricted interests in young children with autism: Comparisons with controls and stability over two years. Autism Research, 6(6), 584–595.

Kalb, L., Freedman, B., Foster, C., Menon, D., Landa, R., Kishfy, L., & Law, P. (2012). Determinants of appointment absenteeism at an outpatient pediatric autism clinic. Journal of Developmental and Behavioral Pediatrics, 33(9), 685–697.

Kasari, C., Gulsrud, A., Freeman, S., Paparella, T., & Hellemann, G. (2012). Longitudinal follow-up of children with autism receiving targeted interventions on joint attention and play. Journal of the American Academy of Child and Adolescent Psychiatry, 51(5), 487–495.

Kim, J., Sunwoo, H., Park, S., Noh, D., Jung, Y., Cho, I., … Yoo, H. (2015). A validation study of the Korean version of social communication questionnaire. Journal of Korean Academic Child Adolescent Psychiatry, 26(3), 197–208.

Kim, S., & Lord, C. (2010). Restricted and repetitive behaviors in toddlers and preschoolers with autism spectrum disorders

based on the Autism Diagnostic Observation Schedule (ADOS). Autism Research, 3(4), 162–173.
Kim, T., & Park, R. (1995). Korean version of childhood autism rating scale. Seoul: Special Education.
Kim, Y., Kim, K., Yoon, H., & Kim, W. (2003). Sequenced language scale for infants (SELSI). Seoul: Special Education Publication Co.
Kleinman, J. M., Ventola, P. E., Pandey, J., Verbalis, A. D., Barton, M., Hodgson, S., … Fein, D. (2008). Diagnostic stability in very young children with autism spectrum disorders. Journal of Autism and Developmental Disorders, 38(4), 606–615.
Kwon, H., Yoo, H., Kim, J., Noh, D., Sunwoo, H., Jeon, Y., & Bong, G. (2017). Re-adjusting the cut-off score of the Korean version of the Childhood Autism Rating Scale for high-functioning individuals with autism spectrum disorder. Psychiatry and Clinical Neurosciences, 71(10), 725–732. https://doi.org/10.1111/pcn.12540
Landis, J. R., & Koch, G. G. (1977). The Measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310
Lord, C., Luyster, R., Gotham, K., & Guthrie, W. (2012). Autism diagnostic observation schedule, Second Edition (ADOS-2) Manual. Torrance: Western Psychological Services.
Macaril, S. L., Campbell, D., Gengoux, C. W., Saulnier, C. A., Klin, A. J., & Chawarska, K. (2012). Predicting developmental status from 12 to 24 months in infants at risk for autism spectrum disorder: A preliminary report. Journal of Autism and Developmental Disorders, 42(12), 2636–2647.
Miller, J. S., Gabrielsen, T., Villalobos, M., Alleman, R., Wahmhoff, N., Carbone, P. S., & Segura, B. (2011). The each child study: Systematic screening for autism spectrum disorders in a pediatric setting. Pediatrics, 127, 866–871.
Ministry of Health and Welfare. (2017). *Daycare User Statistics. IOP Publishing Ministry of Health and Welfare*. Retrieved from http://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_15407_NN002&conn_path=I2.
Ministry of the Interior and Safety. (2017). *Population by Age. IOP Publishing Ministry of the Interior and Safety*. Retrieved from http://www.mois.go.kr/frt/sub/a05/ageStat/screen.do.
Mundy, P., & Crowson, M. (1997). Joint attention and early social communication: Implications for research on intervention with autism. Journal of Autism and Developmental Disorders, 27(6), 653–676.
Mundy, P., Sigman, M., & Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in autism. Development and Psychopathology, 6(3), 389–401.
Naber, F. B., Bakermans-Kranenburg, M. J., Ijzendoorn, M. H., Dietz, C., Daalen, E., Swinkels, S. H., … Engeland, H. (2007). Joint attention development in toddlers with autism. European Child & Adolescent Psychiatry, 17(3), 143–152.
National Institute of Special Education. (2015). 2014 Special Education Field Survey. National Institute of Special Education.
Newschaffer, C., Lee, L., David, A., & Lee, N. (2004). Investigating developmental delays: A comparison of the SCQ and PDDST. Poster presented at the International Meeting for Autism Research, Sacramento, CA.
Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. Development and Psychopathology, 14(2), 1239–1251. https://doi.org/10.1017/s0954579402002031
Richler, J., Huerta, M., Bishop, S. L., & Lord, C. (2010). Developmental trajectories of restricted and repetitive behaviors and interests in children with autism spectrum disorders. Development and Psychopathology, 22(1), 55–69. https://doi.org/10.1017/S0954579409990265
Robins, D. L., Fein, D., & Barton, M. (1999). Modified checklist for autism in toddlers follow up (M-CHAT/F). Self-published.
Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. Journal of Autism and Developmental Disorders, 31, 131–144.
Rogers, C. L., Goddard, L., Hill, E. L., Henry, L. A., & Crane, L. (2016). Experiences of diagnosing autism spectrum disorder: A survey of professionals in the United Kingdom. Autism, 20 (7), 820–831. https://doi.org/10.1177/1362361315611109
Rutter, M., Bailey, A., Lord, C., & Berument, S. K. (2003). Social communication questionnaire. Los Angeles, CA: Western Psychological Services.
Rutter, M., LeCouteur, A., & Lord, C. (2003). Autism diagnostic interview – Revised. Los Angeles: Western Psychological Services.
Schopler, E., Reichler, R., & Rochen, R. (1988). The childhood autism rating scale. Los Angeles: Western Psychological Services.
Skellern, C., Schluter, P., & Mcdowell, M. (2005). From complexity to category: Responding to diagnostic uncertainties of autistic spectrum disorders. Journal of Pediatrics and Child Health, 41(8), 407–412. https://doi.org/10.1111/j.1440-1754.2005.00634
Stenberg, N., Bresnahan, M., Gunnes, N., Hirtz, D., Hornig, M., Lie, K. K., … Stoltenberga, C. (2014). Identifying children with autism spectrum disorder at 18 months in a general population sample. Paediatric and Perinatal Epidemiology, 28(3), 255–262.
Stone, W.,. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. Journal of Autism and Developmental Disorders, 34, 691–701.
Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Screening tool for autism in two-year-olds (STAT): Development and preliminary data. Journal of Autism and Developmental Disorders, 30, 607–612.
Stone, W. L., Hoffman, E. L., Lewis, S. E., & Ousley, O. Y. (1994). Early recognition of autism: Parental reports vs clinical observation. Archives of Pediatrics & Adolescent Medicine, 148(2), 174–179.
Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in two-year-olds (STAT) for children under 24 months. Autism, 12, 557–573.
Sunwoo, H., Noh, D., Kim, K., Kim, J., & Yoo, H. (2017). A study on practitioner's perceptions on early screening of autism spectrum disorder. Journal of the Korean Academy of Child Adolescent Psychiatry, 28(2), 96–105. https://doi.org/10.5765/jkacap.2017.28.2.96
Szatmari, P., Chawarska, K., Dawson, G., Georgiades, S., Landa, R., Lord, C., … Halladay, A. (2016). Prospective longitudinal studies of infant siblings of children with autism: Lessons learned and future directions. Journal of the American Academy of Child and Adolescent Psychiatry, 55(3), 179–187.
Tager-Flusberg, H. (2010). The origins of social impairments in autism spectrum disorder: Studies of infants at risk. Neural Networks, 23(8–9), 1072–1076.

Taylor, C. M., Vehorn, A., Noble, H., Weitlauf, A. S., & Warren, Z. E. (2014). Brief report: Can metrics of reporting bias enhance early autism screening measures? Journal of Autism and Developmental Disorders, 44(9), 2375–2380. https://doi.org/10.1007/s10803-014-2099-5

The Korea Centers for Disease Control and Prevention. (2017). Korean Developmental Screening Test for infants and children.

Towle, P. O., & Patrick, P. A. (2016). Autism spectrum disorder screening instruments for very young children: A systematic review. Autism Research and Treatment, 2016, 1–29. http://doi.org/10.1155/2016/4624829

Ventola, P., Kleinman, J., Pandey, J., Wilson, L., Esser, E., Boorstein, H., … Fein, D. (2007). Differentiating between autism spectrum disorders and other developmental disabilities in children who failed a screening instrument for ASD. Journal of Autism and Developmental Disorder, 37(3), 425–436.

Volkmar, F. R. (2013). Vineland adaptive behavior scales, second edition, encyclopedia of autism spectrum disorders. New York: Springer.

Watson, L. R., Baranek, G. T., Crais, E. J., Reznick, J. S., Dykstra, J., & Perryman, T. (2007). The first year inventory: retrospective parent responses to a questionnaire designed to identify one-year-olds at risk for autism. Journal of Autism and Developmental Disorders, 37, 49–61.

Watson, P., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. Theriogenology, 73(9), 1167–1179. https://doi.org/10.1016/j.theriogenology.2010.01.003

Wiggins, L. D., Baio, J., & Rice, C. (2006). Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. Journal of Developmental & Behavioral Pediatrics, 27, 78–87.

Wörmann, V., Holodynski, M., Kärtner, J., & Keller, H. (2012). A cross-cultural comparison of the development of the social smile. A longitudinal study of maternal and infant imitation in 6- and 12-week-old infants. Infant Behavior & Development, 35, 335–347.

Yoo, H. (2008). Korean version of social communication questionnaire. Seoul: Hakjisa.

Yoo, H., Bong, G., Kwak, Y., Lee, M., Cho, S., Kim, B., … Kim, S. (2017). Korean version of autism diagnostic observation schedule-second edition (ADOS-2). Seoul: Inpsyt.

Yoo, H., Park, K., Kwak, Y., Cho, S., Ban, G., & Kim, B. (2007). Korean version of autism diagnostic interview-revised (ADI-R). Seoul: Inpsyt.

Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., … Natowicz, M. R. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: Recommendations for practice and research. Pediatrics, 136, 60–81.

Zwaigenbaum, L., Bryson, S., Lord, C., Rogers, S., Carter, A., Carver, L., … Yirmiya, N. (2009). Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. Pediatrics, 123(5), 1383–1391.