ASD Screening with the Child Behavior Checklist/1.5-5 in the Study to Explore Early Development

Susan E. Levy[1,2], **Leslie A. Rescorla**[3], **Jesse L. Chittams**[2], **Tanja J. Kral**[2], **Eric J. Moody**[4], **Juhi Pandey**[1,2], **Jennifer A. Pinto-Martin**[2], **Alison T. Pomykacz**[1], **AnnJosette Ramirez**[2,6], **Nuri Reyes**[4], **Cordelia Robinson Rosenberg**[4], **Laura Schieve**[5], **Aleda Thompson**[2], **Lisa Young**[2,7], **Jing Zhang**[2], **Lisa Wiggins**[5]

[1]Center for Autism Research, Roberts Research Center, Children's Hospital of Philadelphia, 2716 South Street, Room 4361, Philadelphia, PA 19146, USA

[2]University of Pennsylvania, Philadelphia, PA, USA

[3]Bryn Mawr College, Bryn Mawr, PA, USA

[4]Department of Psychiatry, University of Colorado School of Medicine, Aurora, CO, USA

[5]Centers for Disease Control and Prevention, Atlanta, GA, USA

[6]Thomas Jefferson University, Philadelphia, PA, USA

[7]Temple University, Philadelphia, PA, USA

Abstract

We analyzed CBCL/1½–5 Pervasive Developmental Problems (DSM-PDP) scores in 3- to 5-year-olds from the Study to Explore Early Development (SEED), a multi-site case control study, with the objective to discriminate children with ASD (N = 656) from children with Developmental Delay (DD) (N = 646), children with Developmental Delay (DD) plus ASD features (DD-AF) (N = 284), and population controls (POP) (N = 827). ASD diagnosis was confirmed with the ADOS and ADI-R. With a cut-point of T ≥ 65, sensitivity was 80% for ASD, with specificity varying across groups: POP (0.93), DD-noAF (0.85), and DD-AF (0.50). One-way ANOVA yielded a large group effect ($\eta^2 = 0.50$). Our results support the CBCL/1½–5's as a time-efficient ASD screener for identifying preschoolers needing further evaluation.

Keywords

Autism spectrum disorder (ASD); Child Behavior Checklist (CBCL); Developmental delay (DD)

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social interaction/communication and restricted and repetitive behaviors (American Psychiatric Association 2013). Early identification of ASD is crucial because it facilitates early intervention, which is associated with better outcomes (Zwaigenbaum et al. 2015) Because developmental conditions such as ASD may initially present in primary care settings, the American Academy of Pediatrics (AAP) published algorithms to help identify young children at risk for developmental disorders (Council on Children with Disabilities 2006) and a diagnosis of ASD (Johnson et al. 2007). These algorithms include incorporating developmental and behavioral surveillance and screening with standardized tools into primary care visits. The ASD screening algorithm recommends screening with a standardized autism-specific tool at the 9-, 18- and 24-month visits for all children, regardless of risk status. The Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al. 2001) has been widely used in preschool populations (under 3 years) to screen for ASD risk (Chlebowski et al. 2013; Robins et al. 2014; Pandey et al. 2008). In a 2014 sample of 16,071 children (Robins et al. 2014), different cutoffs yielded widely different sensitivity (SENS) /positive predictive value (PPV) results (e.g., from 0.91/0.14 to 0.67/0.51), indicating a sharp trade-off between correct identifications and false positives.

The Child Behavior Checklist for Ages 1½–5 (CBCL/1½–5; Achenbach and Rescorla 2000) is a broad-spectrum assessment instrument. Its 99 items yield scores on seven empirically based syndromes (Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Sleep Problems, Attention Problems, and Aggressive Behavior), three broad-spectrum scales (Externalizing, Internalizing, and Total Problems), and five *DSM*-oriented scales (Affective Problems, Anxiety Problems, Pervasive Developmental Problems (PDP), Attention Deficit/Hyperactivity problems, and Oppositional Defiant Problems). After publication of the *DSM*-5 (American Psychiatric Association 2013), the 2000 *DSM*-PDP scale (with 13 items) was renamed the DSM-Autism Spectrum Problems Scale and one of the 13 items was removed. However, we used the original 13-item *DSM*-PDP scale in the current study because the data were collected and scored before publication of the *DSM*-5.

Several studies have examined the performance of the CBCL/1½–5's *DSM*-PDP scale and Withdrawn syndrome for ASD screening, usually using a $T \geq 65$ cutoff (Rescorla 1988; Rescorla et al. 2015; Sikora et al. 2008; Muratori et al. 2011; Narzisi et al. 2013; Myers et al. 2014). When differentiating children with ASD from typically developing children, the *DSM*-PDP scale's SENS ranged from 80 to 97% and specificity (SPEC) ranged from 42 to 91%. Using the Withdrawn syndrome to differentiate the ASD group from typically developing children, Muratori et al. (2011) reported SENS 89%, SPEC = 92%, and Rescorla et al. (2015) reported SENS = 78% and SPEC = 89%. When differentiating children with ASD from those with other psychiatric disorders, SENS was often unchanged but SPEC was

lower due to more false positives (e.g., PDP/Withdrawn: 55%/63% in Rescorla et al. 2015; 60%/65% in; Muratori et al. 2011). These findings are consistent with other screeners that often do not differentiate autism well from other developmental disorders (Moody et al. 2017). Using the Withdrawn syndrome, Havdahl et al. (2016) reported SENS = 0.63, SPEC = 0.65 when differentiating children with ASD from those with non-ASD developmental and behavioral/emotional problems, and Rescorla et al. (2015) reported SENS = 78% and SPEC = 53% when differentiating children with ASD from those with other developmental disabilities.

While earlier results are promising, the utility of the CBCL/1½–5 as an ASD screener needs testing in a large sample. The Study to Explore Early Development (SEED) (Schendel et al. 2012) sample is ideal for this as (a) it is a multisite population-based case–control study, (b) children were classified with ASD after a comprehensive evaluation with gold standard instruments, and (c) CBCL scores for the large ASD group ($n = 656$) could be compared with those for three comparison groups children with developmental delays with some ASD features (DD-AF, $n = 284$); children with developmental delays without ASD features (DD-noAF, $n = 646$); and children from a population sample (POP, $n = 827$). The overall goal of this study was to examine the utility of the CBCL/1½–5 as a screening tool for a diagnosis of ASD in preschool children. We hypothesized that mean DSM-PDP and Withdrawn scores would be rank-ordered ASD > DD-AF > DD > POP, with significant differences between all groups. Our primary hypothesis was that use of a T score 65 cutpoint on *DSM*-PDP and Withdrawn scores would discriminate the ASD group from the other three groups based on SENS, SPEC, positive predictive value (PPV), and negative predictive value (NPV), with false positive rates varying across the three comparison groups.

Methods

Participants

SEED was conducted in six sites (California, Colorado, Georgia, Maryland, North Carolina, and Pennsylvania), with the study protocol approved by Institutional Review Boards at each site. Parents provided signed consent for their child's participation. Children eligible for this study were born between September 1, 2003 and August 31, 2006 (30–68 months of age), lived in one of the study catchment areas at birth and at the time of the first SEED study contact, and lived with a knowledgeable caregiver who could communicate in English (in California and Colorado, in English or Spanish). Children with a range of developmental disorders were identified from special education programs and healthcare providers in each site's catchment area, including children with and without a previous ASD diagnosis. A population comparison group (POP) was ascertained through random sampling of birth certificates from state vital records. More details about the study design appear elsewhere (Wiggins et al. 2015a; DiGuiseppi et al. 2016).

Data Collection and Final Classification

For study enrollment and triage process, all SEED participants were administered the Social Communication Questionnaire (SCQ; Chandler et al. 2007) to screen for ASD symptoms. Several studies have shown that a SCQ cutoff of 11 maximized SENS and SPEC in young

children (Allen et al. 2007; Wiggins et al. 2007). Therefore, SEED defined a SCQ score of 11 or above as an indicator of risk of ASD. Children with an SCQ lower than 11 and without a previous ASD diagnosis were given the Mullen Scales of Early Learning (MSEL; Mullen 1995). Parents of children with MSEL scores < 78 completed the Vineland Adaptive Behavior Scales-Second Edition (Sparrow et al. 2005). Children with an SCQ score higher than 11 or with a previous ASD diagnosis received a comprehensive developmental evaluation consisting of the Autism Diagnostic Interview-Revised (ADI-R; de Bildt et al. 2015) and the Autism Diagnostic Observation Schedule (ADOS; Le Couteur et al. 2008), the MSEL, and the Vineland-II. If SEED study clinicians suspected ASD, a comprehensive developmental evaluation was administered, even if the children had a SCQ score < 11 and no documented ASD diagnosis. As reported by Schendel et al. (2012), 34% of children who were assigned to receive an ASD clinical evaluation did not have a previous ASD diagnosis but were targeted for full assessment because of an SCQ score > 11 (32%) or because of clinical suspicion during the developmental evaluation (2%). The CBCL/1½–5 and Social Responsiveness Scale (SRS; Constantino et al. 2003; Constantino 2011) were administered to collect additional phenotypic information on enrolled children.

The SEED final classification was based on results of the ADI-R and the ADOS, taking the child's overall developmental status into account. Children classified as ASD met ASD criteria on both the ADOS and ADI-R or met ASD criteria on the ADOS and one of three ADI-R criteria. Children who did not meet the ASD criteria were classified into the DD or POP groups (or an "Incomplete Classification" group, not included in our sample). Children in the DD group had received a developmental diagnosis (e.g., language delay, intellectual disability, or other delays) from a clinical provider, or received early intervention or special education services (Schendel et al. 2012). As noted above, the DD group was further divided into two subgroups (DD-AF and DD-noAF), with the DD-AF group more phenotypically similar than the DD-noAF group to the ASD group because of having some ASD features (Wiggins et al. 2015a, b).

As shown in Fig. 1 there were 3769 children enrolled in the SEED study. After excluding children who did not have sufficiently complete clinic visits to classify them as ASD or non-ASD as well as children who had no CBCL/1½–5 scores, our sample comprised 2413 children (64.0% of SEED sample). We used the same criteria for assignment to our four groups as Wiggins et al. (2015a), except that we eliminated children without CBCL/1½–5 data, resulting in slightly smaller sample sizes in each group than those studied by Wiggins (e.g., ASD: $n = 656$ vs. 707, DD-AF: $n = 284$ vs. 305, DD-noAF: $n = 646$ vs. 690, and POP: $n = 827$ vs. 898).

Measures

We used several measures to assess developmental and behavioral characteristics of our four groups, namely the MSEL to assess developmental levels, the Vineland-II to assess adaptive functioning, and the SCQ, SRS, and ADOS to measure ASD characteristics.

Our primary measure was the CBCL/1½–5, a widely used instrument for identifying a range of behavioral and emotional problems among young children (Achenbach and Rescorla 2000). Written at a 5th grade reading level, it contains 99 items, which the parent rates as *0*

= *Not True, as far as you know*, *1* = *Somewhat or Sometimes True*, or *2* = *Very True or Often True*. As noted above, these 99 items are scored on seven statistically derived syndromes, five *DSM*-oriented scales, and three broad-spectrum scales. As in previous studies examining the CBCL/1½–5 as an ASD screener, we compared our groups on the *DSM*- PDP scale and the Withdrawn syndrome scale (see Table 1). For both scales, a normalized T score of 65 demarcates the clinical + borderline range (Achenbach and Rescorla 2000).

Data Analysis

Data were analyzed using the SAS System for Windows (Version 9.3; SAS Institute, Cary, NC). Maternal and child demographic covariates included maternal education, race/ethnicity, poverty status, child age, and sex. Poverty status was determined using the U.S. Census Bureau 2015 Poverty Thresholds (United States Census Bureau 2015).

One-way analysis of variance (ANOVA) and Chi square tests were used to compare our four groups (ASD, DD-AF, DD-noAF, and POP) on demographic and behavioral/developmental characteristics. We used ANOVA to compare *DSM*-PDP and Withdrawn scores for our four groups, using an alpha level of $p < 0.001$ and effect sizes (η^2) interpreted as small = 1 to 5.9%, medium = 6 to 13.9%, and large 14% (Cohen 1988).

We next used receiver operating characteristics (ROC) analysis to obtain the area under the curve (AUC) statistic, which indicates how well the *DSM*-PDP and Withdrawn scales discriminated children in the ASD group from those in each of the other three groups. Because ROC output shows SENS and 1-SPEC values for a full range of cutpoints, one can decide which cutpoint appears "optimal" for each comparison group (i.e., the one that seems to maximize SENS while minimizing 1-SPEC).

In our primary analysis, we calculated decision statistics for both scales using the commonly used cutpoint of *T* 65, determining SENS, SPEC, PPV, and negative predictive value (NPV) when the ASD group was compared with each of the other three groups (DD-AF, DD-noAF, and POP).

Results

Group Differences in Demographic, Developmental, and Behavioral Characteristics

As seen in Table 2, the four groups did not differ by age, but the ASD group had the lowest percentage of girls and the POP group had the highest, with the two DD groups in between. The DD-AF group had the lowest percentage of children with white mothers, and the highest percentage of mothers who had not finished high school and families in poverty. The ASD group had the lowest scores on the MSEL, $F(3, 2398) = 455.6$, $p < 0.001$, $\eta^2 = 0.36$, and the Vineland-II, $F(3,1205) = 70.4$, $p < 0.001$, $\eta^2 = 0.15$. The ASD group had the highest scores on the SCQ, $F(3,2396) = 1387.6$, $p < 0.001$, $\eta^2 = 0.63$, the SRS, $F(3,162) = 58.8$, $p < 0.001$, $\eta^2 = 0.52$ (preschool), $F(3,2130) = 1017.4$, $p < 0.001$, $\eta^2 = 0.59$ (child total), and the ADOS, $F(1,854) = 1025.4$, $p < 0.001$, $\eta^2 = 0.55$. Student–Newman–Keuls (SNK) pair-wise comparisons indicated that all groups differed significantly from each other on all the

developmental and behavioral measures except for the SRS-preschool, on which the DD-noAF and the POP groups did not differ from each other.

Group Differences in *DSM*–PDP and Withdrawn Mean Scores

One-way ANOVAs yielded significant group effects for both *DSM*-PDP, $F(3,2409) = 797.3$, $p < 0.001$, $\eta^2 = 0.50$, and Withdrawn, $F(3,2409) = 662.2$, $p < 0.001$, $\eta^2 = 0.45$. SNK pair-wise comparisons between groups were all significant, with the pattern ASD > DD-AF > DD-noAF > POP. For both scales, the differences between the ASD and DD-noAF and POP groups were 1.5 *SD*, whereas the difference between the ASD and DD-AF group was about 0.5 *SD.*

ROC Results

The "optimal" cutpoints derived from the ROC analysis and the AUCs for these cutpoints varied by both scale and comparison group, as shown by the AUCs and confidence intervals (CIs) in Table 3. AUCs for the *DSM*-PDP scale were 0.95 (95% CI 0.91–0.95) for ASD vs. POP, 0.90 (95% CI 0.88–0.91) for ASD vs. DD-noAF, and 0.71 (95% CI 0.67–0.75) for ASD vs. DD-AF. Results were very similar for the Withdrawn scale. The high AUCs and narrow CIs (i.e., span of only 0.03 points) indicate that both the PDP and Withdrawn scales discriminate very well between the ASD and the POP and DD-noAF groups. In contrast, the scales discriminate much less well between the ASD and DD-AF groups, as seen in the lower AUCs and wider CIs (i.e., span of 0.08 and 0.07 points). This is to be expected, given that the children in the DD-AF group, by definition, have ASD features.

Decision Statistics Results Using a *T* 65 Cutpoint

Our primary results involved calculation of decision statistics based on a cutpoint of *T* 65 on the DSM-PDP scale or the Withdrawn syndrome. This is a commonly used cutpoint in previous ASD screening studies. Equivalent to the 93rd percentile, *T* 65 is widely used as a criterion for "deviance" for the CBCL/1½–5 on all narrow-band scales and syndromes, as it is the threshold for the "borderline" range. Table 3 shows decision statistics (i.e., SENS, SPEC, PPV, NPV) for *T* 65 when the ASD group was compared with each of the other three groups using the two scales. For *DSM*-PDP, SENS was 0.80 for all three comparisons, but SPEC varied across the POP, DD-noAF, and DD-AF comparison groups (0.93, 0.85, 0.50), as did PPV (0.91, 0.85, 0.79) and NPV (0.86, 0.81, 0.52). The same pattern was found for Withdrawn, but SENS was much lower (0.66) whereas SPEC was somewhat higher (0.96, 0.91, 0.63). We also calculated decision statistics using the criterion of **either** *DSM*-PDP **or** Withdrawn *T* 65, which differed only minimally from the results for only *DSM*-PDP *T* 65, as shown in Table 3. This result is consistent with the fact that 97% of the children in the ASD group who scored *T* 65 on either scale had *T* 65 on the *DSM*-PDP scale, whereas only 80% of the children in the ASD group who scored *T* 65 on either scale had *T* 65 on the Withdrawn scale.

Discussion

Our results suggest that the *DSM*-PDP scale of the CBCL/1½–5 may be an effective means of identifying children with ASD features needing further evaluation. The CBCL/1½–5 is

more effective in distinguishing children with ASD from children with typical development and/or drawn from the POP controls or with DD but without ASD features than from children with DD and some ASD symptoms. Since we used a large population sample, state-of-the-art evaluation procedures, and three comparison groups, our study represents a significant advance in testing the performance of the CBCL/1½–5 as an ASD screener.

Between-group differences on the *DSM*-PDP and Withdrawn scales yielded large effect sizes (η^2 of 0.50 and 0.45), and all pairwise comparisons were significant. Mean score differences between the ASD group and the DD-noAF and POP groups were   1.5 *SDs*, whereas the difference between the ASD and DD-AF group was about 0.5 *SDs*. Our ROC decision statistics yielded the strongest discrimination for ASD vs. POP (AUCs = 0.95, 0.94). Very similar results were reported in previous studies when Italian and Korean children with ASD were compared with typically developing children on the *DSM*-PDP and Withdrawn scales (AUCs of 0.93–0.95) (Rescorla 1988; Rescorla et al. 2015; Sikora et al. 2008; Muratori et al. 2011; Narzisi et al. 2013; Myers et al. 2014).

A unique aspect of our study is that the DD group was subdivided into 646 children without ASD features and 284 children with ASD features. Mean scores on the *DSM*-PDP scale were significantly higher for the DD-AF than the DD-noAF group, indicating they were phenotypically quite different from each other. Discrimination was also much better for the ASD vs. DD-noAF comparison than for the ASD vs. DD-AF comparison (AUCs of 0.90 vs. 0.71). It is difficult to discern the degree to which the children with developmental delay in previous CBCL/1½–5 studies had some ASD features. However, the AUCs of 0.68–0.73 reported for the Korean sample (Rescorla et al. 2015) as well as AUCs of 0.74–0.75 reported for a US sample (Havdahl et al. 2016) suggest that these DD groups were more like our DD-AF group than our DD-noAF group. A clinical implication of our findings for the DD-AF vs. DD-noAF groups is that the *DSM*-PDP will have fewer false positives if the children without an ASD diagnosis also have few ASD features.

In our study, a *T* score   65 on DSM-PDP identified 80% of the children in the ASD group regardless of the comparison group. However, as hypothesized, SPEC, PPV, and NPV varied with the comparison group, with the best results for the ASD-POP comparison (0.93, 0.91, 0.86). Our decision statistics results are consistent with those reported in the Italian sample (Muratori et al. 2011; Narzisi et al. 2013) using a *T* score   65 on the *DSM*-PDP scale to compare the ASD and typically developing (TD) groups (SENS = 0.85, SPEC = 0.90, PPV = 0.88, and NPV = 0.92). Our SENS, SPEC, PPV, and NPV were also comparable to those for the Korean sample (Rescorla et al. 2015) using the *T*   65 on the *DSM*-PDP (0.80, 0.87, 0.55, 0.96), although their PPV was lower because the ASD group was rather small compared to the TD group (46 vs. 228).

Using a *T*   65 on the *DSM*-PDP, our SENS, SPEC, PPV, and NPV decision statistics were also very good for our ASD-DD-noAF comparison (0.80, 0.85, 0.85, 0.81) and stronger than the DD group results for the Korean sample (Rescorla et al. 2015) with respect to SPEC (0.60) and PPV (0.36). Compared to results reported by Myers et al. (2014) (SENS = 0.79, SPEC = 0.48) and Sikora et al. (2008) (SENS = 0.80, SPEC = 0.42), our DD-noAF SENS was comparable but our SPEC was much higher. Our decision statistics for the DD-noAF

group were also stronger than those Havdahl et al. (2016) reported for the CBCL Withdrawn syndrome with a comparison group of children with developmental and behavioral/ emotional problems (SENS = 0.63, SPEC = 0.65 for a cutpoint 65). Our results for the DD-AF comparison were very similar to those of Myers et al. and Sikora et al. (SENS = 0.80, SPEC = 0.50); our SENS was higher but SPEC was lower than those reported by Havdahl et al. The consistently lower SPEC in the DD-AF group is consistent with previous research showing high false positive rates when other behavioral or developmental challenges exist (Moody et al. 2017) This highlights the need to develop screeners that accurately detect ASD from other developmental challenges.

Consistent with Muratori et al. (2011) and Rescorla et al. (2015), our results were quite similar for the *DSM*-PDP scale and the Withdrawn syndrome. This is not surprising given that the scales have five overlapping items. However, the *DSM*-PDP scale includes some important ASD-like behaviors not included among the Withdrawn items (i.e., *63. Repeatedly rocks head or body, 80. Strange behavior*) and does not include behaviors less specific to ASD that are included in the Withdrawn syndrome (i.e., *1. Acts young for age; 62. Refuses to play active games*). Furthermore, in our study, the *DSM*-PDP scale had better SENS than the Withdrawn syndrome (0.80 vs. 0.66), with only slightly lower SPEC for the typically developing group (0.93 vs. 0.96), which is the comparison group most important for Level 1 screening. Additionally, using the criterion of either *DSM*-PDP or Withdrawn T 65 barely changed the decision statistics results from those obtained using the criterion of only *DSM*-PDP T 65. Furthermore, 97% of the children with ASD who met the "either" criterion met the DSM-PDP criterion, but only 80% of them met the Withdrawn criterion. For all these reasons, our results suggest that the *DSM*-PDP scale is a better screening tool for ASD than the Withdrawn syndrome.

Our ROC analyses indicated that the "optimal cutpoint" varied somewhat by scale and by comparison group. Furthermore, variable cutpoints are not practical for use by clinical practitioners (e.g., pediatricians, nurse practitioners, psychologists, psychiatrists) who are screening individual children for ASD risk. Fortunately, a cutpoint of T 65 yielded strong SENS, especially for the *DSM*-PDP scale, regardless of the comparison group. It also has the advantage of demarcating the borderline + clinical ranges (93rd percentile) on the CBCL/1½–5's *DSM*-oriented and syndrome scales and is the cutpoint commonly used in previous CBCL/1½–5 ASD screening studies.

Our study had numerous strengths. SEED is a multisite case–control study in six states. Our sample was very large (N = 2413), with an ASD group of 656 children. An additional strength is that we subdivided our DD group into those with no ASD features (n = 646) and those with some ASD symptoms but not enough to meet study criteria for ASD (n = 284). We had a large population based comparison group (POP) (n = 827). However, limitations of our study must also be acknowledged. These include that all groups were not recruited and assessed in the same fashion and that some SEED participants (36%) were excluded because they lacked CBCL data or a definitive diagnostic classification. Also, children in the DD-AF group tended to come from low socioeconomic status, and the ASD sample was composed of mostly boys. An additional limitation is that we do not have specific information about diagnoses other than ASD or DD for our sample, but Wiggins et al. (2015a) reported some

information about diagnoses for the full SEED sample (including those children without CBCL/1½–5 scores, whom we excluded for our study). For example, Wiggins found that 53% of parents/caregivers reported that their child had at least one diagnosis other than ASD. Diagnoses of language delay, sensory integration disorder, and motor delay were most common in the ASD group. Children classified in the DD-AF group tended to have an ADHD diagnosis, whereas those in the DD-noAF group tended to have a Down syndrome diagnosis. As would be expected, children in the POP group had the fewest parent-reported diagnoses.

Conclusions

Although current AAP guidelines recommend use of an autism-specific screening tool at the 9-, 18- and 24-month visits for all children, our results suggest that a broad-spectrum instrument such as the CBCL/1½–5 might also have several advantages as an ASD screener, particularly after the age of 30 months. First, the M-CHAT's age range is 18 to 30 months, whereas the CBCL/1½–5 is normed for children up to age 6, thereby encompassing the age range in which ASD may be typically diagnosed. Second, the 99-item CBCL/1½–5 (which can be completed by parents in < 15 min either on paper or online) provides normed scores on 15 distinct scales and thus screens for a wide range of preschool behavioral/emotional problems. Third, the CBCL/1½–5 does not require health professionals to administer or score and quickly produces scored output that can be incorporated into an electronic medical record. Fourth, since the CBCL is a broad-spectrum instrument tapping many kinds of problems, a parent's pre-existing beliefs about if her/his child has ASD may be less likely to influence ratings than a scale with only ASD features. These advantages plus the strong screening results found for the CBCL/1½-in the current study and in previous studies (Rescorla 1988, 2015; Sikora et al. 2008; Muratori et al. 2011; Narzisi et al. 2013; Myers et al. 2014) suggest that the CBCL/1½-,5 has strong potential for use as an ASD screener in pediatric settings, particularly in children > 30 months of age, when the use of the M-CHAT may no longer be appropriate. Future research comparing results for the CBCL/1½–5 with autism-specific screeners in children > 30 months, would be useful in exploring this potential.

Despite these many advantages, there are also some drawbacks to using the CBCL/1½–5 in pediatric settings. Because the forms are not in the public domain, there is a cost, albeit modest, for purchasing/printing each form. Additionally, there is a cost for either the web-based per-unit scoring utility or for the PC-based unlimited use scoring program. Furthermore, entering and scoring each form takes about 5 min for a clerical worker; however, if parents complete the form online, scoring is even faster. The scoring results can be stored in electronic medical records (EMRs), but some system set up may be required initially to ensure compatibility, given the many platforms used for EMRs.

The current results provide evidence for the *DSM*-oriented PDP scale's potential as an ASD screener to identify children with ASD features who need further evaluation. Because of our unique study sample, this evidence is an important advance relative to previous small-scale studies. If the CBCL/1½–5 suggests that a child is at risk, a parallel form can be completed and scored online by teachers/caregivers, providing valuable multi-informant assessment

information not easily obtained using other instruments. We therefore conclude that the CBCL/1½–5 is a promising mechanism to identify young children at risk for ASD and other problems who can be referred for more detailed assessment.

Acknowledgments

References

Achenbach TM, & Rescorla LA (2000). Manual for the ASEBA preschool forms & profiles: An integrated system of multi-informant assessment; Child Behavior Checklist for Ages 1 1/2–5; Language Development Survey; Caregiver-teacher Report Form. Burlington: University of Vermont.

Allen CW, Silove N, Williams K, & Hutchins P (2007). Validity of the social communication questionnaire in assessing risk of autism in preschool children with developmental problems. Journal of Autism and Developmental Disorders, 37(7), 1272–1278. 10.1007/s10803-006-0279-7. [PubMed: 17080270]

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Arlington: American Psychiatric Publishing.

Chandler S, Charman T, Baird G, Simonoff E, Loucas T, Meldrum D, et al. (2007). Validation of the social communication questionnaire in a population cohort of children with autism spectrum disorders. Journal of the American Academy of Child & Adolescent Psychiatry, 46(10), 1324–1332. 10.1097/chi.0b013e31812f7d8d. [PubMed: 17885574]

Chlebowski C, Robins DL, Barton ML, & Fein D (2013). Largescale use of the modified checklist for autism in low-risk toddlers. Pediatrics, 131(4), e1121–e1127. 10.1542/peds.2012-1525.

Cohen J (1988). Statistical Power for the Behavioral Sciences. New Jersey: Lawrence Erlbaum Associates.

Constantino JN (2011). (SRS™) Social Responsiveness Scale™. Torrence: WPS.

Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, et al. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. Journal of Autism and Developmental Disorders, 33(4), 427–433. [PubMed: 12959421]

Council on Children with Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee & Medical Home Initiatives for Project Advisory Committee. (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. Pediatrics, 118(1), 405–420. 10.1542/peds.2006-1231. [PubMed: 16818591]

de Bildt A, Sytema S, Zander E, Bolte S, Sturm H, Yirmiya N, et al. (2015). Autism Diagnostic Interview-Revised (ADI-R) algorithms for toddlers and young preschoolers: Application in a Non-US sample of 1,104 children. Journal of Autism and Developmental Disorders, 45(7), 2076–2091. 10.1007/s10803-015-2372-2. [PubMed: 25682078]

DiGuiseppi CG, Daniels JL, Fallin DM, Rosenberg SA, Schieve LA, Thomas KC, et al. (2016). Demographic profile of families and children in the Study to Explore Early Development (SEED): Case–control study of autism spectrum disorder. Disability and Health Journal, 9(3), 544–551. 10.1016/j.dhjo.2016.01.005. [PubMed: 26917104]

Havdahl KA, von Tetzchner S, Huerta M, Lord C, & Bishop SL (2016). Utility of the Child Behavior Checklist as a screener for autism spectrum disorder. Autism Research, 9(1), 33–42. 10.1002/aur.1515. [PubMed: 26140652]

Johnson CP, Myers SM, & American Academy of Pediatrics, Council On Children With Disabilities (2007). Identification and evaluation of children with autism spectrum disorders. Pediatrics, 120(5), 1183–1215. 10.1542/peds.2007-2361. [PubMed: 17967920]

Le Couteur A, Haden G, Hammal D, & McConachie H (2008). Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: The ADI-R and the ADOS. Journal of Autism and Developmental Disorders, 38(2), 362–372. 10.1007/s10803-007-0403-3. [PubMed: 17605097]

Moody EJ, Reyes N, Ledbetter C, Wiggins LD, Diguiseppi C, Alexander-Shardel A, et al. (2017). Screening for autism with the SRS and SCQ: Variations across demographic, developmental and behavioral factors in preschool children. Journal of Autism and Developmental Disorders 47, 3550–3561 [PubMed: 28856480]

Mullen EM (1995). Mullen Scales of Early Learning. San Antonio: Pearson.

Muratori F, Narzisi A, Tancredi R, Cosenza A, Calugi S, Saviozzi I, et al. (2011). The CBCL 1.5–5 and the identification of preschoolers with autism in Italy. Epidemiology and Psychiatric Sciences, 20(4), 329–338. [PubMed: 22201210]

Myers CL, Gross AD, & McReynolds BM (2014). Broadband behavior rating scales as screeners for autism? Journal of Autism and Developmental Disorders, 44(6), 1403–1413. 10.1007/s10803-013-2004-7. [PubMed: 24287879]

Narzisi A, Calderoni S, Maestro S, Calugi S, Mottes E, & Muratori F (2013). Child Behavior Checklist (11/2–5) as a tool to identify toddlers with autism spectrum disorders: A case–control study. Research Development and Disability, 34(4), 1179–1189. 10.1016/j.ridd.2012.12.020.

Pandey J, Verbalis A, Robins DL, Boorstein H, Klin AM, Babitz T, et al. (2008). Screening for autism in older and younger toddlers with the modified checklist for autism in toddlers. Autism, 12(5), 513–535. 10.1177/1362361308094503. [PubMed: 18805945]

Rescorla L (1988). Cluster analytic identification of autistic preschoolers. Journal of Autism and Developmental Disorders, 18(4), 475–492. [PubMed: 3215877]

Rescorla L, Kim YA, & Oh KJ (2015). Screening for ASD with the Korean CBCL/1(1/2)-5. Journal of Autism and Developmental Disorders, 45(12), 4039–4050. 10.1007/s10803-014-2255-y. [PubMed: 25239178]

Robins DL, Casagrande K, Barton M, Chen CM, Dumont-Mathieu T, & Fein D (2014). Validation of the modified checklist for Autism in toddlers, revised with follow-up (M-CHATR/F). Pediatrics, 133(1), 37–45. 10.1542/peds.2013-1813. [PubMed: 24366990]

Robins DL, Fein D, Barton ML, & Green JA (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(2), 131–144. [PubMed: 11450812]

Schendel DE, Diguiseppi C, Croen LA, Fallin MD, Reed PL, Schieve LA, et al. (2012). The Study to Explore Early Development (SEED): A multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. Journal of Autism and Developmental Disorders, 42(10), 2121–2140. 10.1007/s10803-012-1461-8. [PubMed: 22350336]

Sikora DM, Hall TA, Hartley SL, Gerrard-Morris AE, & Cagle S (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. Journal of Autism and Developmental Disorders, 38(3), 440–448. 10.1007/s10803-007-0407-z. [PubMed: 17619131]

Sparrow S, Balla DA, & Cicchetti DV (2005). Vineland Adaptive Behavior Scales (2nd edn.). San Antonio: Pearson.

United States Census Bureau (2015). Poverty thresholds by size of family and number of children. Suitland: United States Census Bureau.

Wiggins LD, Bakeman R, Adamson LB, & Robins D (2007). The utility of the social communication questionnaire in screening for autism in screening for autism in children referred for early intervention. Focus on Autism and Other Developmental Disabilities, 22(1), 33–38.

Wiggins LD, Levy SE, Daniels J, Schieve L, Croen LA, DiGuiseppi C, et al. (2015a). Autism spectrum disorder symptoms among children enrolled in the Study to Explore Early Development (SEED). Journal of Autism and Developmental Disorders, 45(10), 3183–3194. 10.1007/s10803-015-2476-8. [PubMed: 26048040]

Wiggins LD, Reynolds A, Rice CE, Moody EJ, Bernal P, Blaskey L, et al. (2015b). Using standardized diagnostic instruments to classify children with autism in the study to explore early development.

Journal of Autism and Developmental Disorders, 45(5), 1271–1280. 10.1007/s10803-014-2287-3. [PubMed: 25348175]

Zwaigenbaum L, Bauman ML, Choueiri R, Fein D, Kasari C, Pierce K, et al. (2015). Early identification and interventions for autism spectrum disorder: Executive summary. Pediatrics, 136(Suppl 1), S1–S9. 10.1542/peds.2014-3667B. [PubMed: 26430167]
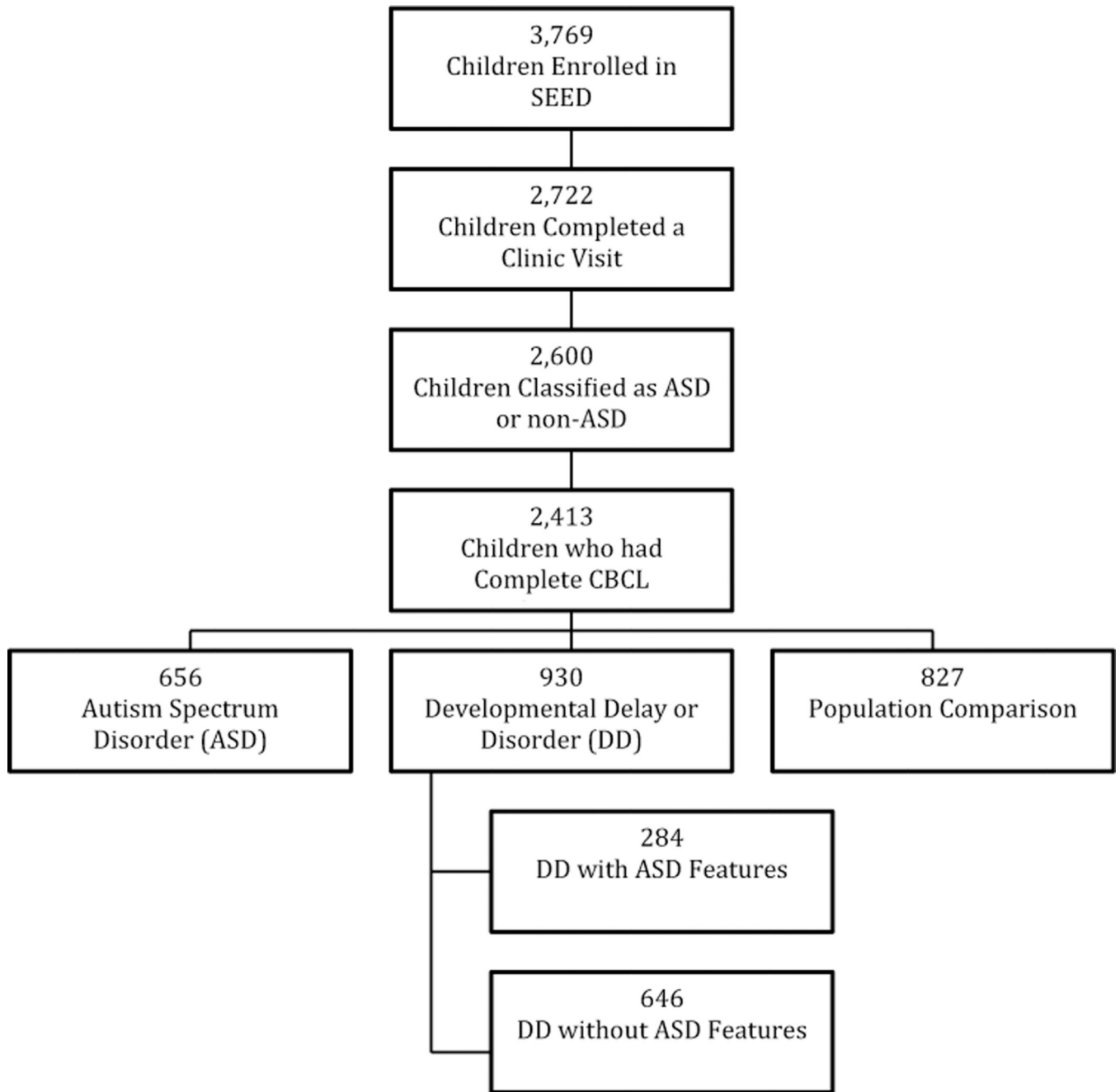
Fig. 1.
Participant selection flow diagram

Table 1

CBCL/1½–5 Withdrawn syndrome and DSM-PDP scale items

Withdrawn syndrome	DSM-PDP scale
Item 1. *Acts young for age*	
	Item 3. *Afraid to try new things*
Item 4. *Avoids looking others in the eye*	Item 4. *Avoids looking others in the eye*
	Item 7. *Can't stand things out of place*
	Item 21. *Disturbed by any change in routine*
Item 23. *Doesn't answer when people talk to him/her*	Item 23. *Doesn't answer when people talk to him/her*
	Item 25. *Doesn't get along with other children*
Item 62. *Refuses to play active games*	
	Item 63. *Repeatedly rocks head or body*
Item 67. *Seems unresponsive to affection*	Item 67. *Seems unresponsive to affection*
Item 70. *Shows little affection toward people*	Item 70. *Shows little affection toward people*
Item 71. *Shows little interest in things around him/her*	
	Item 76. *Speech problem*
	Item 80. *Strange behavior*
	Item 92. *Upset by new people or situations*
Item 98. *Withdrawn, doesn't get involved with others*	Item 98. *Withdrawn, doesn't get involved with others*

CBCL, Child Behavior Checklist; *DSM*-PDP, *DSM*-Pervasive Developmental Problems Scale

Table 2

Demographic and behavioral characteristics of the four SEED study groups

Characteristic	ASD (n = 656)	DD-AF (n = 284)	DD-noAF (n = 646)	POP (n = 827)
Children				
	N (%)	N (%)	N (%)	N (%)
Age (months)	59.5 (6.7)	60.1 (7.1)	59.7 (6.9)	59.8 (6.9)
Female ***	120 (18.3%)	71 (25.0%)	237 (36.7%)	378 (45.7%)
	M (SD)	M (SD)	M (SD)	M (SD)
MSEL total score ***	66.3 (21.0)[a]	78.6 (19.6)[b]	89.8 (20.9)[c]	101.9 (14.8)[d]
Vineland-II total ***	304.0 (56.1)[a]	343.6 (51.1)[b]	322.5 (54.2)[c]	374.6 (42.6)[d]
SCQ total score ***	17.3 (6.1)[a]	13.5 (5.0)[b]	4.9 (2.9)[c]	4.2 (3.4)[d]
SRS child total ***	94.0 (30.1)[a]	67.4 (27.7)[b]	35.9 (21.7)[c]	26.1 (17.2)[d]
SRS preschool total ***	81.9 (29.6)[a]	58.5 (25.2)[b]	33.2 (16.5)[c]	29.8 (15.4)[c]
ADOS total score ***	7.1 (1.6)[a]	2.9 (2.1)[b]	NA	NA
DSM-PDP ***	71.5 (9.4)[a]	64.3 (9.8)[b]	55.9 (7.3)[c]	53.1 (5.4)[d]
Withdrawn ***	70.0 (10.4)[a]	62.3 (9.6)[b]	54.8 (6.6)[c]	53.1 (5.2)[d]
Mothers				
	N (%)	N (%)	N (%)	N (%)
Race/ethnicity, n (%)				
White	361 (57.7)	130 (47.6)	423 (68.5)	552 (69.4)
African American	112 (17.9)	76 (27.8)	69 (11.2)	89 (11.2)
Other	134 (21.4)	48 (17.6)	106 (17.2)	139 (17.5)
Hispanic	19 (3.0)	19 (7.0)	20 (3.2)	16 (2.0)
Education, n (%)				
Less than high school	42 (6.5%)	44 (17.2%)	36 (5.8%)	25 (3.1%)
High school	92 (14.3%)	62 (24.2%)	67 (10.7%)	69 (8.4%)
Some college/higher	509 (79.2%)	150 (58.6%)	522 (83.5%)	727 (88.6%)
Below federal poverty level, n (% yes)	66 (10.4%)	65 (24.7%)	39 (6.3%)	50 (6.3%)

ASD, Autism Spectrum Disorder group; DD-AF, Developmental Disabilities with Autistic Features; DD-noAF, Developmental Disabilities without Autistic Features; POP, General Population group; N, number; M, mean; SD, standard deviation; MSEL, Mullen Scales of Early Learning; SCQ, Social Responsiveness Scale; Vineland II (Vineland Adaptive Behavior Scale, Second Edition); SCQ, Social Communication Scale; SRS, Social Responsiveness Scale; ADOS, Autism Diagnostic Observation Schedule; DSM-PDP, Diagnostic and Statistical Manual-Pervasive developmental problems

***
$p < 0.001$

a,b,c,dGroups with different superscripts differ significantly by Student–Newman–Keuls post-hoc tests

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

CBCL ROC and decision statistics results by group with cutpoints of T ≥ 65

	AUC (95% CI)	SENS	SPEC	PPV	NPV
ASD vs. POP					
DSM-PDP ROC (optimal T = 63)	0.95 (0.92–0.95)	0.86	0.90		
Withdrawn ROC (optimal T = 60)	0.94 (0.93–0.96)	0.88	0.85		
DSM-PDP T ≥ 65		0.80	0.93	0.91	0.86
Withdrawn T ≥ 65		0.66	0.96	0.93	0.78
DSM-PDP T ≥ 65 OR Withdrawn T ≥ 65		0.82	0.92	0.89	0.87
ASD vs. DD-no AF					
DSM-PDP ROC (optimal T = 66)	0.90 (0.88–0.91)	0.80	0.85		
Withdrawn ROC (optimal T = 63)	0.90 (0.88–0.91)	0.79	0.84		
DSM-PDP T ≥ 65		0.80	0.85	0.85	0.81
Withdrawn T ≥ 65		0.66	0.91	0.88	0.72
DSM-PDP T ≥ 65 OR Withdrawn T ≥ 65		0.82	0.84	0.84	0.82
ASD vs. DD-AF					
DSM-PDP ROC (optimal T = 59)	0.71 (0.67–0.75)	0.91	0.29		
Withdrawn ROC (optimal T = 60)	0.71 (0.67–0.74)	0.88	0.38		
DSM-PDP T ≥ 65		0.80	0.50	0.79	0.52
Withdrawn T ≥ 65		0.66	0.63	0.81	0.44
DSM-PDP T ≥ 65 OR Withdrawn T ≥ 65		0.82	0.46	0.78	0.53

CBCL, Child Behavior Checklist; *ASD*, Autism Spectrum Disorder group; DD-AF, Developmental Disabilities with Autistic Features; DD-noAF, Developmental Disabilities without Autistic Features; POP, General Population group; ROC, receiver operating characteristics; AUC, area under the curve; SENS, sensitivity; SPEC, specificity; PPV, positive predictive value; NPV, negative predictive value