

The CBCL 1.5–5 and the identification of preschoolers with autism in Italy

F. Muratori*, A. Narzisi, R. Tancredi, A. Cosenza, S. Calugi, I. Saviozzi, E. Santocchi and S. Calderoni

Division of Child Neurology and Psychiatry, Stella Maris Scientific Institute, University of Pisa, Italy

Aims. To study the potential use of child behaviour checklist (CBCL) 1.5–5 scales for the early identification of preschoolers at risk of autism.

Methods. CBCL scores of three groups of preschoolers were compared: (1) an experimental group of 101 preschoolers with autism spectrum disorder (ASD); (2) a control group of 95 preschoolers with other psychiatric disorders (OPD); (3) a control group of 117 preschoolers with typical development (TD). One-way analysis of variance (ANOVA), logistic regression with odds ratio (OR) and receiver operating characteristic (ROC) analyses were performed.

Results. ANOVA revealed that ASD and OPD had significantly higher scores in almost all CBCL scales than TD. ASD presented significantly higher scores than OPD on Withdrawn, Attention Problems and Pervasive Developmental Problems (PDP) scales. Logistic regression analysis demonstrated that these same CBCL scales have validity in predicting the presence of an ASD towards both TD and OPD. ROC analysis indicated high sensitivity and specificity for PDP (0.85 and 0.90) and Withdrawn (0.89 and 0.92) scales when ASD is compared to TD. Specificity (0.60 for PDP and 0.65 for Withdrawn) decreases when comparing ASD and OPD

Conclusions. The PDP and Withdrawn scales have a good predictive validity so that they could be proposed as a first-level tool to identify preschoolers at risk of autism in primary care settings. Problems regarding the lower specificity when comparing ASD *v.* OPD are discussed.

Received 19 November 2010; Revised 9 March 2011; Accepted 9 March 2011

Key words: Autistic disorder, behaviour, early diagnosis, questionnaires.

Introduction

Autism spectrum disorders (ASDs) are a class of neurodevelopmental pathologies with heterogeneous genetic abnormalities that reflect a highly variable severity of behavioural phenotypes (Ruggeri & Tansella, 2009; Tosato & Lasalvia, 2009). The lack of medical tests or biological markers for identifying ASD has led researchers to concentrate on behavioural phenotypes to detect early signs of autism. As recommended by the practice parameters of the American Academy of Neurology and the Child Neurology Society (AACN) (Filipek *et al.* 2000), an appropriate and timely ASD diagnosis requires two different levels of investigation: Level (1) a routine developmental surveillance; Level (2) an exhaustive evaluation restricted to children identified at risk at Level 1.

The principal aim of the present study was to investigate the possible use of child behaviour checklist

(CBCL) 1.5–5 as a Level 1 tool to support non-specialized professionals (e.g. paediatricians) in their ability to detect behaviours that are suggestive of an ASD risk. Primary care providers are in the best position to screen for ASD because they see young children routinely for medical care. Thus, it would be useful to have a tool that allows them to easily identify children at risk for ASD and then refer such children for a Level 2 evaluation. In fact, even if most parents of children with autism first become concerned about their child's development in the first years of life (Chawarska *et al.* 2007; Ozonoff *et al.* 2009), several months or years may elapse before their worry is taken into consideration by paediatricians and a further delay may occur between the paediatric visit and a specialist's ASD diagnosis (Maestro *et al.* 1999; Wiggins *et al.* 2006). The implications of this gap extend far beyond outcome gains associated with an early intervention (Ricciardi *et al.* 2008; Vismara & Rogers, 2010); in fact, most parents of children with autism experience considerable amounts of stress because they are parenting a child with atypical development, and the uncertainty of diagnosis accentuates parental anxiety (Muratori *et al.* 2010).

* Address for correspondence: Professor F. Muratori, Division of Child Neurology and Psychiatry, University of Pisa – Stella Maris Scientific Institute, Via dei Giacinti, 2, I-56018 Calambrone (Pisa), Italy. (Email: f.muratori@inpe.unipi.it)

To help clinicians in pointing out the actual risk for autism and in reducing the gap, standardized caregivers checklists and questionnaires have been used. Such tools should differentiate well between children likely to have a disorder and typically developing children (that is high sensitivity), while their specificity is less of concern because ASD can be differentiated from other conditions in a Level 2 assessment. Some specific instruments were developed to identify very young children with ASD (Pinto-Martin *et al.* 2008), while others (usually applied in epidemiology) have proved to be useful in detecting children at risk of general developmental and behavioural disorders (Briggs-Gowan & Carter 2006; Matson *et al.* 2010). Among these latter instruments, the CBCL is one of the most widely used parent report checklist that measures a broad range of behavioural and emotional problems (Achenbach & Rescorla, 2001). It displays adequate reliability and validity and requires little effort to be completed (it takes 5–10 min for parents and 5 min to score).

Almost 20 years ago, Rescorla was the first researcher to use the CBCL for preschoolers with autism (Rescorla, 1988). In that study, the emergence of an autistic factor suggested that a future use of the CBCL as a possible instrument to recognize children with autism might be fruitful. However, after Rescorla's investigation, only a few studies have applied the CBCL scales to children with autism (Bolte *et al.* 1999; Duarte *et al.* 2003). In more recent years, the preschool form of the CBCL (Achenbach & Rescorla, 2000) was developed and used in different settings (Rescorla, 2005). The 100-problem item of this CBCL 1.5–5 yields seven empirically based syndrome scales and five Diagnostic and Statistical Manual (DSM)-oriented scales. Among the latter is the Pervasive Developmental Problems (PDP) scale. To construct this new DSM-oriented scale, the relationship between DSM-IV-TR diagnostic criteria for PDP and item of CBCL 1.5–5 were studied (Krol *et al.* 2006).

For the last several years, we have used this CBCL form for preliminary assessment in our second-level neuropsychiatric clinic. Our impression from this use of the CBCL 1.5–5 was that clinically significant elevations on the PDP scale were in good agreement with clinical ASD diagnosis. Our observation is now supported by some recent papers that applied this CBCL form to children with ASD (Sikora *et al.* 2008; Hartley *et al.* 2008, 2009) and evaluated adequacy of CBCL/1.5–5 factor model in a well-characterized sample of preschoolers with ASD (Pandolfi *et al.* 2009). Sikora *et al.* (2008) have described a better performance of the CBCL/1.5–5 compared to a specific instrument as the Gilliam Autism Rating Scale (GARS), in identifying young children (aged 36–71 months) with ASD. Especially, in this study Withdrawn and PDP subscales

of the CBCL were higher among children with autism than among non-spectrum children, and these subscales had better sensitivity and specificity in identifying children with autism than the GARS.

The present research addressed the following aims: (a) to identify CBCL scales that deviated significantly preschoolers with ASD from typical developing and other psychiatric disorders (OPDs) children; (b) to test the effects of age and gender on CBCL capacity in discriminating among the three groups; (c) to provide more detailed understanding of predictive properties of a broad-scale rating instrument such as the CBCL 1.5–5.

Method

Participants

A total of 313 children aged 24–60 months (mean age 43.8 months; s.d. 12.5) were included in the study. Participants belong to three groups: (1) an experimental group of 101 children affected by an ASD (85 M; 16 F; mean age: 44 months; s.d. 12.3; range 24–60 months); (2) a control group of 95 children with OPD (43 M; 52 F; mean age: 40 months; s.d. 12.7; range 24–56 months); (3) a second control group of 117 preschoolers with typical development (TD) (65 M; 52 F; mean age: 47 months; s.d. 12; range 26–60 months). The whole sample (ASD, OPD and TD) was composed of Caucasian children of Italian descent. All the ASD subjects were consecutively admitted to the Department of Child Neuropsychiatry of the University of Pisa, Stella Maris Scientific Institute (a suburban public academic hospital providing care to patients of all socioeconomic levels, coming from all over Italy) between September 2006 and June 2009 and diagnosed on the base of DSM-IV-TR criteria coupled with clinical judgments made by a senior child psychiatrist with expertise in autism Raffaella Tancredi (RT) and confirmed by Autism Diagnostic Observation Schedule (ADOS) (applied by a certificate MD: Elisa Santocchi (ES)). Laboratory tests to rule-out medical causes of autism included audiometry, standard karyotyping, fragile X testing and metabolic screening; structural brain Magnetic Resonance Imaging (MRI) and Electro Encefalo Gram (EEG) were performed in case of clinical indication and were all assessed as normal.

OPD subjects were selected from a large database of children with psychiatric disorders diagnosed according to the DSM-IV-TR (APA, 2000) or – for children under 3 years of age – to the DC: 0–3R (Zero to Three, 2005). For these children, an ASD diagnosis was actively excluded by a senior child psychiatrist and supported by a score under 15 (that is the cut-off point for an ASD) on the social communication

questionnaire (Rutter *et al.* 2003). Moreover, in order to avoid possible confusing overlap with ASD, also children with a DSM-IV-TR diagnosis of attention deficit hyperactivity disorder (ADHD), and children with multi-system developmental disorder (MSDD) or regulatory disorder (RD), according to DC0-3R criteria, were not included in OPD sample. Final diagnosis of the OPD children was affective disorders for 59 subjects, oppositional defiant disorder for 25 subjects and mixed disorders for 11 subjects.

Children with TD were collected in three urban kindergartens in Pisa; subjects with whatever internalizing problems or/and some parent or teacher concern about child development were excluded.

The study was approved by the research ethics boards of the Stella Maris Scientific Institute.

Measures

The CBCL 1.5–5 (Achenbach & Rescorla, 2000; Frigerio *et al.* 2006) is a 100-item parent-report measure designed to record the problem behaviours of preschoolers. Each item describes a specific behaviour and the parent is asked to rate its frequency on a three-point Likert scale (0, not true; 1, somewhat or sometimes true; 2, very true or often true). The scoring gives a summary profile (including internalizing, externalizing and total problems scores), a syndrome profile (emotionally reactive, anxious/depressed, somatic complaints, withdrawn, sleep problems, attention problems and aggressive behaviour) and five different DSM-oriented scales (affective problems, anxiety problems, pervasive developmental problems, attention deficit/hyperactive problems and oppositional defiant problems). A T-score of 63 and above for summary scales and of 70 and above for syndrome and DSM-oriented scales, are generally considered clinically significant; values between 60 and 63 for summary scales or between 65 and 70 for syndrome and DSM-oriented scales, identify the borderline clinical range; values under 60 or under 65 are considered not-clinical. Scores and profiles for each child were obtained, thanks to a computer scoring software. Each profile has an easy-reading layout that allows one to immediately understand whether the scores are in normal, borderline or clinical range. Even if specific interest was devoted to the 13-item PDP scale (see Appendix), this paper focuses on the potential diagnostic role of all the CBCL scales.

Procedures

Parents (mother when possible) of the 313 children filled the CBCL 1.5–5. In ASD and OPD, CBCL 1.5–5 was completed at the beginning of a multidisciplinary clinical

observation at the Scientific Institute Stella Maris; parents of TD filled the CBCL 1.5–5 in anonymous way at kindergarten. In order to avoid any bias related to the fact that caregivers of the clinical groups were subjected to different diagnostic interviews, CBCL was administered before the beginning of the clinical assessment.

Data analysis

The CBCL scales were examined for normality using skewness tests and Kolmogorov–Smirnov testing. None of the scales had significant departures from normality.

Chi-square test was used to compare categorical variables among the three groups. One-way analysis of variance (ANOVA) with *post hoc* S-N-K was performed to test differences on age and the CBCL scales among ASD, OPD and TD groups, multivariate analysis of covariance (MANCOVA) was used to evaluate differences among the three groups on CBCL scales, regardless of gender and age.

Logistic regression analysis with odds ratios (ORs) was performed to identify CBCL scales discriminating among the three groups. We used separate logistic regression models to compare ASD *v.* TD and ASD *v.* OPD. In Model 1, the independent variable was CBCL total score; in Model 2, the independent variables were internalizing and externalizing scores; in Model 3 the independent variables were syndrome scales; in Model 4 the independent variables were the five different DSM-oriented scales.

CBCL scales that were identified as predictors of an ASD diagnosis in the logistic regression analysis at $p < 0.001$ were used in a receiver operating characteristic (ROC) analysis, in order to determine their optimal cut-offs to differentiate children with ASD from children with TD or OPD.

In the ROC analysis, sensitivity and specificity were plotted over the range of cut-off points. The area under the curve (AUC) represents the accuracy of the instrument in predicting children who will have or will not have ASD. The interpretation of the AUC values is traditionally as follows: an AUC < 0.7 suggests ‘low’ diagnostic accuracy; an AUC from 0.7 to 0.9 suggests ‘moderate’ diagnostic accuracy and an AUC ≥ 0.9 suggests ‘high’ diagnostic accuracy (Sweet & Pickett, 1982).

Analyses were carried out using SPSS version 15.0 for Windows (SPSS Inc. Chicago, IL, USA).

Results

Preliminary analyses

Overall, 313 subjects were recruited (61, 6% males and 38, 4% females, mean age 43.8 ± 12.5 months).

Chi-square analysis revealed a significant difference on gender distribution among ASD, OPD and TD groups (chi-square=36.32, $p < 0.001$): the percentage of females was significantly lower in the ASD group when compared with the other two groups.

Chi-square analysis showed no significant differences for socio-economic status (chi-square=0.27, $p = 0.965$): children belong mostly to middle-class families according to the Hollingshead & Redlich criteria (Hollingshead & Redlich, 1958).

One-way ANOVA indicated a significant difference on age among the groups ($F[2,310] = 5.20$, $p = 0.006$); S-N-K *post hoc* test revealed that the difference was due to the younger age of the OPD compared to TD group ($p = 0.005$); no differences were found between ASD and TD and between ASD and OPD.

Clinical characteristics

One-way ANOVA comparing the three groups (Table 1) revealed that ASD and OPD groups had significantly higher scores in all CBCL scales than TD group, except for somatic complaints and sleep problems.

Moreover, the ASD group presented higher scores than OPD group on Withdrawn and Attention Problems scales of the CBCL syndrome profile (Fig. 1a) and on PDP scale of the CBCL DSM-oriented (Fig. 1b). OPD group had higher scores compared to ASD on Anxious/Depressed, Somatic Complaints, Sleep Problems and Aggressive Behaviour scales of the CBCL syndrome profile and on Anxiety Problems and Oppositional Defiant Problems of the CBCL DSM-oriented scales.

MANCOVA showed that the results were not significantly different after controlling for age and gender.

In the logistic regression analysis with OR (Table 1), CBCL scales predicting the presence of an ASD towards both TD and OPD were the Withdrawn scale, the Attention Problems scale and the PDP scale. CBCL Total and internalizing scores were predictors of the presence of an ASD, when comparing ASD with TD, while they did not distinguish between ASD and OPD.

ROC analyses

Because Withdrawn, Attention Problem and PDP scales have been identified as the best predictors of the probable presence of ASD in the logistic regression analysis, we used ROC analyses to estimate the best cut-offs for these scales (Fig. 2). In Table 2, sensitivity, specificity, NPV and PPV and AUC at the optimal cut-

offs for the three scales in discriminating ASD from TD and OPD are reported.

ASD v. TD

ROC analysis indicated that in discriminating ASD from TD group the optimal compromise between sensitivity and specificity was achieved at a score of 65 both on the PDP scale (AUC=0.947; 95% CI 0.920–0.975) and the Withdrawn scale (AUC=0.945; 95% CI 0.914–0.977).

For the PDP scale, the sensitivity, that is the proportion of actual ASD subjects who were correctly identified, was 0.85, and the specificity, that is the proportion of actual TD subjects who were correctly identified, was 0.90. The score of 65 yielded a positive predictive value of 0.88 (i.e. the proportion of individuals with a score of 65 or more who were diagnosed in the ASD group) and a negative predictive value of 0.87 (i.e. the proportion of individuals with a score of less than 65 who were diagnosed in the TD group).

For the Withdrawn scale, sensitivity (0.89), specificity (0.92), PPV (0.90) and NPV (0.90) were even slightly better than in case of the PDP scale.

For the Attention Problems scale, the best cut-off discriminating ASD from TD was 55 but all values (sensitivity, specificity, PPV and NPV) were lower and the AUC (0.850; 95% CI 0.799–0.902) showed a moderate diagnostic accuracy.

ASD v. OPD

In order to discriminate ASD from OPD, as far as PDP scale is regarded the optimal cut-off was 65 (AUC=0.813; 95% CI 0.753–0.873). Using this cut-off, the proportion of subjects with ASD who were correctly diagnosed was 0.85 (sensitivity) and the proportion of cases with OPD who were correctly diagnosed was 0.60 (specificity).

For the Withdrawn scale, the optimal compromise between sensitivity (that was 0.89) and specificity (that was 0.65) was achieved at a score of 62 (AUC=0.850; 95% CI 0.794–0.905). For the Attention Problems scale, the optimal cut-off was 55 (AUC=0.704; 95% CI 0.632–0.776) with a sensitivity of 0.72 and a specificity of 0.55. Following Sweet & Picket criteria (Sweet & Picket, 1982), the AUC have to be considered moderate when comparing ASD and OPD for all the three scales (PDP, Withdrawn and Attention Problems).

Discussion

Brief and validated tools should be used in primary care settings to detect children at risk of autism and

Table 1. One-way ANOVA and logistic regression on CBCL T-scores (means and standard deviations) for ASD, OPD and TD groups

| CBCL scales | ASD (N = 101) | OPD (N = 95) | TD (N = 117) | Logistic regression with OR and 95% CI | | | | | | | |
|-------------------------------|-----------------------------|------------------------------|---------------|--|--------|--------------|------|--------------|--------------|------|--------------|
| | | | | ANOVA | | ASD v. OPD | | | ASD v. TD | | |
| | | | | F | p | p | OR | 95% CI | p | OR | 95% CI |
| Total score | 59.97 (8.20) ^a | 59.22 (10.70) ^c | 47.70 (9.11) | 58.56 | <0.001 | 0.447 | 1.01 | 0.98 to 1.04 | 0.000 | 1.17 | 1.12 to 1.23 |
| Internalizing | 62.11 (7.48) ^a | 59.67 (10.28) ^c | 48.84 (10.56) | 59.84 | <0.001 | 0.014 | 1.05 | 1.01 to 1.10 | 0.000 | 1.13 | 1.07 to 1.18 |
| Externalizing | 56.01 (7.60) ^a | 56.47 (10.56) ^c | 46.70 (8.47) | 42.01 | <0.001 | 0.116 | 0.96 | 0.92 to 1 | 0.015 | 1.07 | 1.01 to 1.13 |
| Emotionally reactive | 59.12 (8.21) ^a | 59.07 (9.13) ^c | 53.31 (5.32) | 21.29 | <0.001 | 0.063 | 1.10 | 0.99 to 1.22 | 0.047 | 1.10 | 1 to 1.21 |
| Anxious/depressed | 56.44 (6.51) | 60.98 (10.08) ^{c,d} | 53.97 (5.70) | 23.02 | <0.001 | 0.001 | 0.85 | 0.78 to 0.94 | 0.099 | 0.91 | 0.82 to 1.01 |
| Somatic complaints | 55.95 (6.84) | 59.02 (7.72) ^{c,d} | 54.15 (5.69) | 13.85 | <0.001 | 0.024 | 0.87 | 0.78 to 0.98 | 0.132 | 0.92 | 0.83 to 1.02 |
| Withdrawn | 71.77 (8.40) ^{a,b} | 59.07 (8.67) ^c | 54.03 (5.89) | 151.18 | <0.001 | 0.000 | 1.32 | 1.2 to 1.45 | 0.000 | 1.29 | 1.19 to 1.39 |
| Sleep problems | 55.29 (6.86) | 59.81 (11.64) ^{c,d} | 54.15 (5.51) | 13.41 | <0.001 | 0.723 | 0.98 | 0.90 to 1.06 | 0.145 | 0.92 | 0.82 to 1.02 |
| Attention problems | 62 (8.15) ^{a,b} | 56.62 (7.11) ^c | 52.72 (4.41) | 53.09 | <0.001 | 0.001 | 1.17 | 1.06 to 1.29 | 0.000 | 1.18 | 1.07 to 1.30 |
| Aggressive behaviour | 55.49 (6.11) ^a | 58.94 (8.93) ^{c,d} | 52.07 (4.07) | 29.18 | <0.001 | 0.000 | 0.75 | 0.66 to 0.86 | 0.303 | 0.92 | 0.70 to 1.07 |
| Affective problems | 58.85 (8.09) ^a | 60.55 (10.47) ^c | 54.03 (5.40) | 18.96 | <0.001 | 0.216 | 0.96 | 0.90 to 1.02 | 0.334 | 0.95 | 0.87 to 1.04 |
| Anxiety problems | 57.22 (7.54) ^a | 60.97 (9.72) ^{c,d} | 53.53 (5.58) | 24.77 | <0.001 | 0.000 | 0.87 | 0.81 to 0.94 | 0.035 | 0.91 | 0.83 to 0.99 |
| PDP | 71.59 (7.30) ^{a,b} | 60.82 (9.40) ^c | 54.33 (6.26) | 138.77 | <0.001 | 0.000 | 1.27 | 1.18 to 1.37 | 0.000 | 1.34 | 1.23 to 1.45 |
| ADHD | 58.50 (7.02) ^a | 57.71 (7.93) ^c | 52.97 (4.52) | 23.04 | <0.001 | 0.970 | 1 | 0.93 to 1.07 | 0.099 | 1.08 | 0.98 to 1.19 |
| Oppositional defiant problems | 54.98 (5.63) ^a | 56.98 (7.56) ^c | 51.69 (3.87) | 22.93 | <0.001 | 0.110 | 0.93 | 0.85 to 1.10 | 0.885 | 0.99 | 0.86 to 1.13 |

ASD, autism spectrum disorder; OPD, other psychiatric disorders; TD, typical development; CBCL, child behaviour checklist; OR, odds ratio; CI, confidence interval.

Superscripts stand for S-N-K *post hoc* tests: ^aASD significantly higher v. TD ($p < 0.01$); ^bASD significantly higher v. OPD ($p < 0.01$); ^cOPD significantly higher v. TD ($p < 0.01$); ^dOPD significantly higher v. ASD.

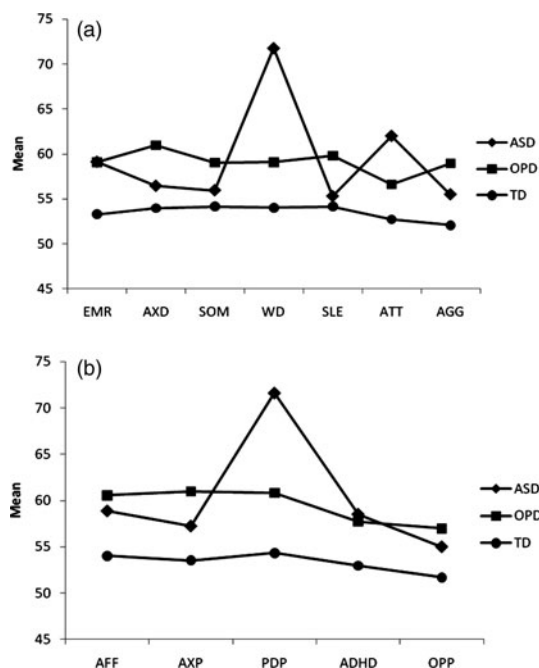


Fig. 1. (a) Means of CBCL T-scores for syndrome scales. ASD, autism spectrum disorder; OPD, other psychiatric disorders; TD, typical development; EMR, emotionally reactive; AXD, anxious/depressed; SOM, somatic complaints; WD, Withdrawn; SLE, sleep problems; ATT, attention problems; AGG, aggressive behaviour. (b) Means of CBCL T-scores for DSM-oriented scales. AFF, affective problems; AXP, anxiety problems; ODP, oppositional defiant problems; PDP, pervasive developmental problems; ADHD, attention deficit and hyperactivity disorder.

to redirect families for an in-depth examination by professionals experienced in developmental disabilities; the final aim of such tools is to reduce the significant time lag between age at first parental worry and age at first ASD diagnosis that is still consistently reported in the literature (Barbaro & Dissanayake, 2009). Our results on the ability of CBCL scales in

discriminating ASD preschoolers suggest the following considerations.

First, we confirm the validity of the PDP scale in differentiating preschoolers with ASD from those with TD. According to Sweet & Pickett (1982) interpretation of the AUC, the diagnostic accuracy of the PDP scale is high. Moreover, when comparing ASD and TD, sensitivity (85%) and specificity (90%) of the PDP scale are both very high and broadly above 80%, which is the recommended cut-off for first-level instruments (Meisels, 1989). Sensitivity is considered the most important measure of a good Level 1 tool as it indicates the probability that a positive test reflects the underlying pathological (ASD in our case) condition; in fact, a high sensitivity corresponds to a low percentage of false negatives, so that the possibility of being affected by an ASD and not being properly diagnosed is reduced. The low rate of false negatives indicates that the PDP scale is able to identify the majority of preschoolers at risk for ASD to refer them to appropriate services with minimal delay. For this reason, our results provide support for the PDP scale as an effective Level 1 tool in individuating children at risk of ASD in the general population. Moreover, the PDP scale shows an even higher specificity that means low rates of false positives. This is the second reason to support its use as a screening tool (Sikora *et al.* 2008) because it limits to families of healthy preschoolers an unnecessary, time-consuming and emotionally exhausting referral to specialty clinics.

As a complementary finding, the Withdrawn scale has shown a power of discrimination that is even slightly higher than the PDP scale (sensitivity: 89%; specificity: 92%). The presence of Withdrawn as a discriminative scale for ASD is consistent with Sikora findings (Sikora, 2008) and with the more recent paper on older children conducted in Singapore (Ooi *et al.* 2010). In both these papers, elevation on this scale is reported as a specific behavioural pattern

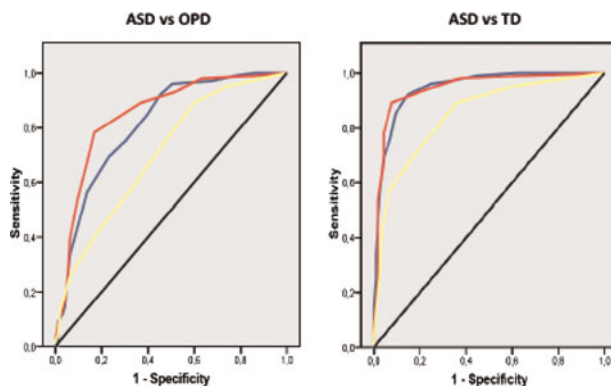


Fig. 2. ROC for Withdrawn (red), attention problems (yellow) and PDP (blue) scales. (A colour version of this figure is available online at <http://journals.cambridge.org/eps>)

Table 2. Sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) at the best cut-off points on Withdrawn, Attention problems and PDP scales when comparing autism spectrum disorder (ASD) v. typical developing children (TD) or v. other psychiatric disorders (OPD)

| | ASD v. TD | | | ASD v. OPD | | |
|---|-----------------------------|--------------------------------------|-----------------------|-----------------------------|--------------------------------------|-----------------------|
| | Withdrawn (cut-off = 65) | Attention problems (cut-off = 55) | PDP (cut-off = 65) | Withdrawn (cut-off = 62) | Attention problems (cut-off = 55) | PDP (cut-off = 65) |
| Sensitivity | 89% | 72% | 85% | 89% | 72% | 85% |
| Specificity | 92% | 80% | 90% | 65% | 55% | 60% |
| PPV | 90% | 76% | 88% | 72% | 63% | 69% |
| Predictive Negative Value (PNV) | 90% | 77% | 87% | 87% | 55% | 79% |
| AUC | 0.945 | 0.850 | 0.947 | 0.850 | 0.704 | 0.813 |
| Sweet and Pickett criteria for AUC interpretation | High | Moderate | High | Moderate | Moderate | Moderate |

ASD, autism spectrum disorder; TD, typical development; OPD, other psychiatric disorders; PDP, pervasive developmental problems; PPV, positive predictive values; PNV, Predictive negative value; AUC, area under the curve; PDP, pervasive developmental problems.

indicative of autism. It may be pointed out that the Withdrawn cluster of items has to be considered in future research and practice not only indicative of an affective or mood disorder, as usually it is intended, but also the expression of social difficulties specific to ASD. Findings from the present research support that high value on both the Withdrawn scale (which is derived from factor analysis) and the PDP scale (which is derived by expert judgment) do an excellent job of screening for ASD.

Third, the present study adds a contribution to the literature on differentiating ASD from OPD. When autism group is compared to OPD, the ASD group is marked by higher scores on Withdrawn and PDP scales and these scales obtained a high sensitivity, nevertheless these two scales showed lower specificity and PPV. It means that these scales contain the risk of over-identify ASD in children with OPD. The fact that the discrimination between ASD and OPD is not as strong as that between ASD and TD was not totally unexpected, even if we have not included in OPD group subjects with boundary conditions such ADHD, MSDD and RD. We can sustain that the moderate diagnostic accuracy could be linked to the prevalence in our OPD group of young children with internalizing disorders; usually these children have high scores in the Withdrawn scale and the fact that five out of the eight items in this scale are shared with the PDP scale could be the reason of elevations on the PDP scale in internalizing disorders. A second explanation concerns the frequent co-occurring emotional problems such as depression in autistic conditions (Pandolfi *et al.* 2009). Recent meta-analyses have reported that up to

84% of ASD experience anxiety (White *et al.* 2009) and up to 34% experience depression (Stewart *et al.* 2006). Thus, we can hypothesize that parents of ASD children answer positively to some item considering the internalizing traits of their autistic child. For these reasons, the specificity of the PDP and Withdrawn scales decreases in differentiating ASD v. OPD; we can also imagine a further decrease when ADHD, MSDD or RD are considered, because these conditions have overlapping symptoms with ASD disorder. Nevertheless, we think that this is not a major drawback for using the CBCL as a screening tool in general population settings. First of all, because CBCL differentiates extremely well between children likely to have a disorder and typically developing children. Second, because the low specificity is due to the elevated scores on most scales in both ASD and OPD (see Figs. 1 and 2). In a primary screening programme, the low specificity towards children with OPD can be tolerated; indeed it allows the identification of children who show behavioural and developmental problems that, even if generally not severe such as in ASD, could benefit from a more specialized diagnostic assessment than the standard mental health evaluation. Thus, high scores in PDP and Withdrawn scales could represent a criterion for the paediatrician to decide whether the child needs a deeper ASD evaluation, with the knowledge that children who are shy, withdrawn (and perhaps language delayed) may turn out truly not to have ASD. Nevertheless, some other diagnosis may be made, so that these children are apparently 'false' positives because they are not ASD but they are not really false positives from a clinical perspective.

Finally, mean scores elevation on the Attention Problem scale in ASD compared to TD or OPD confirms, also in preschoolers, the frequently observed co-existence of attention problems in older children with autism (Bölte et al. 1999; Sinzig et al. 2009) and the overlap between ASD and ADHD is often reported in the literature (de Bruin et al. 2007). Nevertheless, the Attention Problems scale does a much worse job of predicting ASD than do PDP and Withdrawn scales, regardless of which control group is used. In fact, according to the AUC, diagnostic accuracy is only moderate when ASD are compared to both TD and OPD groups. Thus, while we confirm a frequent early clinical association of ASD with attention problems, we do not suggest the utility of this scale in screening autism.

Conclusions

Our study confirms how CBCL, filled out by parents in paediatric primary care, can successfully meet the needs of a broad-scale rating instrument (Tancredi et al. 2002) and shows its ability to evaluate also the risk of ASD in preschoolers. The high sensitivity and specificity of PDP and Withdrawn scales indicate these scales as a useful screening tool that can complement direct paediatric observation of the child (e.g. response to name task, joint attention task and declarative pointing), maximizing the role of the parents.

Nevertheless, there are some limitations associated with the current study. First, CBCL 1.5–5 PDP scale is able to differentiate already-diagnosed patients with ASD from TD children, but remains to investigate if the high sensitivity and specificity is maintained in a screening survey. Moreover, we have to consider the low specificity towards OPD and that the risk to over-identify ASD in children with other OPD could increase when ADHD, MSDD and RD will be considered. Second, we did not present data on the mental age of clinical samples so that present findings prevent us from claiming our data as specific of an ASD and not of a more general developmental delay. Third, in line with the epidemiological data of a strong male preponderance in ASD and a mild female prevalence in OPD (Center for Disease Control and Prevention, 2007), our sample shows significant sex differences in the three groups. However, this gender discrepancy should not interfere with our results, since statistical analysis carried out to analyse significant differences among the three groups on CBCL scales, was controlled for gender. Fourth, CBCL results are generated from primary caregivers and parental bias in interpreting the questions and quantifying the behaviours must be taken into account (Ozonoff et al. 2009). For example, parents with anxiety or mood disorder could over-estimate

maladaptive behaviours of their own kid, while others may be reluctant to acknowledge their child's problems. Nevertheless, the literature has established that parents know their child very well and are for the most part reliable informants about their child's development (Glascoe & Dworkin, 1995; Glascoe, 1999).

Notwithstanding these limits, the distinctive PDP and Withdrawn profiles have shown excellent sensitivity and specificity that are better than in other well-known tools for screening autism as M-CHAT; thus, these scales, which represent two separate and distinct templates for grouping CBCL item, could become a brief, rapid, easy, specific tool for screening of ASD in primary settings. Some problems persist for its specificity particularly towards internalizing disorders so that it could be expected to develop more research on items that have a better power of discrimination between ASD and other emotional or developmental problems.

Declaration of Interest

The author(s) declare that they have not received economic support and they have no conflict of interest related to the present paper.

Acknowledgements

The authors thank Dr. Sandra Maestro, Chief of Infant Psychiatry Unit and two anonymous manuscript reviewers for their valuable discussions and comments.

References

- Achenbach TM, Rescorla L (2001). *Manual for the ASEBA School-Age Forms and Profiles*. University of Vermont, Research Center for Children, Youth and Families: Burlington, VT.
- Achenbach TM, Rescorla L (2000). *Manual for the ASEBA Preschool Forms and Profiles: An Integrated System of Multi-Informant Assessment*. University of Vermont, Department of Psychiatry: Burlington, VT.
- APA (2000). *Diagnostic and Statistical Manual of Mental Disorders, 4th edn text rev. (DSM-IV-TR)*. APA: Washington, DC.
- Barbaro J, Dissanayake C (2009). Autism spectrum disorders in infancy and toddlerhood: a review of the evidence on early signs, early identification tools, and early diagnosis. *Journal of Developmental and Behavioral Pediatrics* 30, 447–459.
- Bölte S, Dickhut H, Poustka F (1999). Patterns of parent-reported problems indicative in autism. *Psychopathology* 32, 93–97.
- Briggs-Gowan MJ, Carter AS (2006). *BITSEA Brief Infant-Toddler Social and Emotional Assessment Examiner's Manual*. Harcourt Assessment: San Antonio, TX.
- Center for Disease Control and Prevention (2007). Prevalence of autism spectrum disorder—Autism and

- developmental disability monitoring network, 14 sites, United States, 2002. *Surveillance Summary* 9/56 (SS01), 12–28.
- Chawarska K, Paul R, Klin A, Hannigen S, Dichtel LE, Volkmar F** (2007). Parental recognition of developmental problems in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 37, 62–72.
- De Bruin EI, Ferdinand RF, Meester S, de Nijs PE, Verheij F** (2007). High rates of psychiatric co-morbidity in PDD-NOS. *Journal of Autism and Developmental Disorders* 37, 877–886.
- Duarte CS, Bordin IAS, De Oliveira A, Bird H** (2003). The CBCL and the identification of children with autism and related conditions in Brazil: pilot findings. *Journal of Autism and Developmental Disorders* 33, 703–707.
- Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH Jr, Dawson G, Gordon B, Gravel JS, Johnson CP, Kallen RJ, Levy SE, Minshew NJ, Ozonoff S, Prizant BM, Rapin I, Rogers SJ, Stone WL, Teplin SW, Tuchman RF, Volkmar FR** (2000). Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. *Neurology* 55, 468–479.
- Frigerio A, Cozzi P, Pastore V, Molteni M, Borgatti R, Montiroso R** (2006). La valutazione dei problemi emotivo comportamentali in un campione italiano di bambini in età prescolare attraverso la CBCL. *Infanzia e Adolescenza* 12, 26–33.
- Glascoe FP, Dworkin PH** (1995). The role of parents in the detection of developmental and behavioral problems. *Pediatrics* 95, 829–836.
- Glascoe FP** (1999). The value of parents' concerns to detect and address developmental and behavioral problems. *Journal of Pediatrics and Child Health* 35, 1–8.
- Hartley SL, Sikora DM, McCoy R** (2008). Prevalence and risk factors of maladaptive behaviour in young children with autistic disorder. *Journal of Intellectual Disability Research* 52, 819–829.
- Hartley SL, Sikora DM** (2009). Sex differences in autism spectrum disorder: an examination of developmental functioning, autistic symptoms, and coexisting behaviour problems in toddlers. *Journal of Autism and Developmental Disorders* 39, 1715–1722.
- Hollingshead AB, Redlich FC** (1958). *Social Class and Mental Illness*. Wiley: New York, NY.
- Krol NP, De Bruyn EF, Coolen JC, van Aarle EJ** (2006). From CBCL to DSM: a comparison of two methods to screen for DSM-IV diagnoses using CBCL data. *Journal of Clinical Child and Adolescent Psychology* 35, 127–135.
- Maestro S, Casella C, Milone A, Muratori F, Palacio-Espasa F** (1999). Study of the onset of autism through home-movies. *Psychopathology* 32, 292–300.
- Matson JL, Fodstad JC, Mahan S, Rojahn J** (2010). Cut-offs, norms and patterns of problem behaviours in children with developmental disabilities on the Baby and Infant Screen for Children with autism Traits (BISCUIT-Part 3). *Developmental Neurorehabilitation* 13, 3–9.
- Meisels SJ** (1989). Can developmental screening tests identify children who are developmentally at risk? *Pediatrics* 83, 578–585.
- Muratori F, Apicella F, Muratori P, Maestro S** (2010). Intersubjective disruptions and caregiver-infant interaction in early autistic disorder. *Research in Autism Spectrum Disorders* 5, n1, 408–417.
- Ooi YP, Rescorla L, Ang RP, Woo B, Fung DSS** (2010). Identification of autism spectrum disorders using the child behaviour checklist in Singapore. *Journal of Autism and Developmental Disorders*. [Epub ahead of print Apr. 20]. DOI 10.1007/s10803-010-1015-x
- Ozonoff S, Young GS, Steinfeld MB, Hill MM, Cook I, Hutman T, Macari S, Rogers SJ, Sigman M** (2009). How early do parent concerns predict later autism diagnosis? *Journal of Developmental and Behavioral Pediatrics* 30, 367–375.
- Pandolfi V, Magyar CI, Dill CA** (2009). Confirmatory factor analysis of the child behaviour checklist 1.5–5 in a sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 39, 986–995.
- Pinto-Martin JA, Young LM, Mandell DS, Poghosyan L, Giarelli E, Levy SE** (2008). Screening strategies for autism spectrum disorders in pediatric primary care. *Journal of Developmental and Behavioral Pediatrics* 29, 345–350.
- Rescorla LA** (1988). Cluster Analytic Identification of Autistic Preschoolers. *Journal of Autism and Developmental Disorders* 18, 475–492.
- Rescorla LA** (2005). Assessment of young children using the Achenbach System of Empirically Based Assessment (ASEBA). *Mental Retardation and Developmental Disabilities Research Review* 11, 226–237.
- Ricciardi A, McAllister V, Dazzan P** (2008). Is early intervention in psychosis effective? *Epidemiologia e Psichiatria Sociale* 17, 227–235.
- Ruggeri M, Tansella M** (2009). The interaction between genetics and epidemiology: the puzzle and its pieces. *Epidemiologia e Psichiatria Sociale* 18, 77–80.
- Rutter M, Bailey A, Lord C** (2003). *Social Communication Questionnaire*. Western Psychological Services: Los Angeles, CA.
- 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, 440–448.
- Sinzig J, Walter D, Doepfner M** (2009). Attention deficit/hyperactivity disorder in children and adolescent with autism spectrum disorder. Symptom or Syndrome? *Journal of Attention Disorders* 20, 1–10.
- Stewart ME, Barnard L, Pearson J, Hasan R, O'Brien G** (2006). Presentation of depression in autism and asperger syndrome: a review. *Autism* 10, 103–116.
- Sweet JA, Pickett RM** (1982). *Evaluation of Diagnostic Systems: methods from Signal Detection Theory*. Academic Press: New York.
- Tancredi R, Picchi L, Floriani C, Seppia P, Muratori F** (2002). Behavioral problems in pediatric primary care. *Italian Journal of Pediatrics* 28, 377–382.
- Tosato S, Lasalvia A** (2009). The contribution of epidemiology to defining the most appropriate approach to genetic research on schizophrenia. *Epidemiologia e Psichiatria Sociale* 18, 81–90.

- Vismara LA, Rogers SJ** (2010). Behavioral treatments in autism spectrum disorder: what do we know? *Annual Review of Clinical Psychology* **6**, 447–468.
- White SW, Oswald D, Ollendick T, Scahill L** (2009). Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review* **29**, 216–229.

- Wiggins LD, 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 and Behavioral Pediatrics* **27**, 79–87.
- Zero to Three** (2005). *Diagnostic classification of mental health and developmental disorders of infancy and early childhood, revised (DC:0–3R)*. Zero to Three Press: Washington, DC.

Appendix. List of the items composing the ‘Pervasive Developmental Problems’ scale

| Item number in CBCL | Item description |
|---------------------|---|
| 3. | Afraid to try new things |
| 4. | Avoids looking others in the eye |
| 7. | Can't stand having things out of place |
| 21. | Disturbed by any change in routine |
| 23. | Doesn't answer when people talk to him/her |
| 25. | Doesn't get along with other children |
| 63. | Repeatedly rocks head or body |
| 67. | Seems unresponsive to affection |
| 70. | Shows little affection toward people |
| 76. | Speech problem |
| 80. | Strange behaviour |
| 92. | Upset by new people or situations |
| 98. | Withdrawn, doesn't get involved with others |