Utility of the Japanese version of the Checklist for Autism in Toddlers for predicting pervasive developmental disorders at age 2

Tomonori Koyama, PhD,* Eiko Inokuchi, MD, Naoko Inada, MA, Miho Kuroda, MA, Aiko Moriwaki, MA, Masatoshi Katagiri, MA, Madoka Noriuchi, PhD and Yoko Kamio, MD, PhD

Department of Child and Adolescent Mental Health, National Institute of Mental Health, National Center of Neurology and Psychiatry, Tokyo, Japan

We evaluated the utility of the Japanese version of the Checklist for Autism in Toddlers for predicting pervasive developmental disorders (PDD) among 2-year-old children in clinical settings. Confirmed diagnosis revealed that the pass rate on four items (social interest, proto-imperative pointing, proto-declarative pointing and joint-attention) was significantly lower in 52 PDD children than in 48 non-PDD children, and if abnormal development was reported in two or more items, the sensitivity, specificity, and positive/negative predictive values for PDD diagnosis were 0.85, 0.73, and 0.77/0.81, respectively. This simple screening tool can provide valuable information to clinicians when diagnosing PDD.

Key words: checklist for autism in toddlers, early detection, pervasive developmental disorders, sensitivity, specificity.

EARLY INTERVENTION IS very important for children with pervasive developmental disorders (PDD) as it facilitates their social adaptation and reduces difficulties in parenting.1–3 Although acceptable diagnostic stability has been reported for PDD diagnoses made at age 2,4 some clinicians may hesitate to make a diagnosis at such a young age.

The Checklist for Autism in Toddlers (CHAT) is an epidemiological screening tool that was developed in England for detecting PDD at 18 months of age.5 It consists of nine parental interview questions (Section A) and five behavior observations (Section B). The Japanese version of the CHAT (CHAT-J) was developed by the first author and his colleagues, and its internal consistency (high Cronbach’s alpha), criterion validity (significant correlation with the Childhood Autism Rating Scale Tokyo Version6), and construct validity (lower pass rate in PDD than in control) have been confirmed.7

The questions in Section A are applicable for older infants in clinical settings; however, the questions used for epidemiological screening include dummy items in which the authors assumed that some aspects of the development of autistic children, such as rough and tumble play (item 1) and motor development (item 3), were the same as normal children.5 In clinical settings, assessment using selected items may be more practical in comparison to using all items. In this study, we evaluated the efficacy of using a simplified version of the CHAT-J in clinical settings as a tool for predicting PDD.

METHODS

Procedures and subjects

The subjects in this study were 100 2-year-old children who were consecutively referred to an experienced child psychiatrist at a regional clinic in Tokyo specializing in developmental disorders between
January 2003 and December 2007. The subjects were followed up until after 3 years of age (range, 36–94 months; mean length of follow up, 24.6 months) and had a confirmed diagnosis.

At each child’s initial visit, an experienced psychologist retrospectively asked about the development of the child before 2 years of age using Section A of the CHAT-J. Cognitive evaluation with the Japanese version of the Stanford–Binet Intelligence Scale or the Kyoto Scale of Psychological Development (K-test) was also performed depending on the developmental level of the child. The developmental quotient of the K-test is considered almost equivalent to IQ.

In the clinic, a clinical team comprised of three to five experienced professionals (i.e. child psychiatrists, clinical psychologists, psychiatric social workers, and pediatric neurologists) made DSM-IV diagnoses by consensus using all available information regarding the children, such as detailed parental interviews about developmental history, comprehensive clinical evaluations of the children, and behavior assessments by nursery teachers who participated in remedial therapy groups, but excluding the CHAT-J data. Of the 100 children, 52 children were diagnosed with PDD (mean age at CHAT-J administration, 30.1 ± 3.4 months; range, 24–35 months; 47 boys and five girls) and 48 children were not (non-PDD: mean age, 31.0 ± 3.0 months; range, 25–35 months; 38 boys and 10 girls). PDD subcategories were five with autistic disorder, three with Asperger’s disorder, and 44 with PDD not otherwise specified (PDD-NOS). The predominance of PDD-NOS observed in this study is consistent with the findings of a recent review of PDD prevalence. No significant differences in mean age or sex ratio were observed between the two groups, although the mean IQ was significantly lower in PDD children (71.2 ± 18.4; range, 35–137) than non-PDD children (81.8 ± 17.7; range, 49–129).

The protocol of the present study was approved by the ethics committee of the Tokyo University Graduate School of Medicine when the first author was affiliated with the university.

**Statistical analysis**

First, the pass rate for each item was compared between the two groups using Fisher’s exact test with Bonferroni’s correction. For items with statistical significance, we counted the number of failed items for each child and searched for an appropriate cut-off value for differentiating PDD from non-PDD children by examining the sensitivity, specificity, and positive/negative predictive values for each cut-off value.

All tests were two-tailed and statistical significance was set at $P < 0.01$. All statistical analyses were performed using SPSS 17.0J for Windows (SPSS Japan, Tokyo, Japan).

**Results**

The pass rate was significantly different between the two groups for item 2 (social interest: PDD vs non-PDD, 38.5% vs 79.2%), item 6 (proto-imperative pointing: 25.0% vs 75.0%), item 7 (proto-declarative pointing: 23.1% vs 79.2%), and item 9 (joint-attention: 26.9% vs 68.8%). No significant differences were observed for item 1 (rough and tumble play: 96.2% vs 100.0%), item 3 (motor development: 88.5% vs 83.3%), item 4 (social play: 75.0% vs 85.4%), item 5 (pretend play: 13.5% vs 35.4%), and item 8 (functional play: 73.1% vs 95.8%).

Based on these results, we concluded that four items (items 2, 6, 7, and 9) were useful for predicting PDD (Cronbach’s alpha, 0.77). Subsequently, the number of failed items (0–4) was counted for each child. Table 1 shows the cut-off value and related measures for predicting pervasive developmental disorders.

<table>
<thead>
<tr>
<th>Failed items (n)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.37</td>
<td>0.94</td>
<td>0.86</td>
<td>0.58</td>
</tr>
<tr>
<td>≥3</td>
<td>0.67</td>
<td>0.81</td>
<td>0.80</td>
<td>0.70</td>
</tr>
<tr>
<td>≥2</td>
<td>0.85</td>
<td>0.73</td>
<td>0.77</td>
<td>0.81</td>
</tr>
<tr>
<td>≥1</td>
<td>0.98</td>
<td>0.54</td>
<td>0.70</td>
<td>0.96</td>
</tr>
</tbody>
</table>


© 2010 The Authors

Journal compilation © 2010 Japanese Society of Psychiatry and Neurology
measures for differentiating PDD from non-PDD. A PDD diagnosis was best predicted if the cut-off value was set at 2 or more failed items.

The same analyses were conducted for IQ-balanced subgroups of children (IQ ≥ 70; 27 PDD and 36 non-PDD children), and similar results were obtained: sensitivity (0.74), specificity (0.78), and positive/negative predictive values (0.71/0.80).

**DISCUSSION**

The present study showed that satisfactory prediction of a PDD diagnosis could be made using only four CHAT-J items (social interest, proto-imperative pointing, proto-declarative pointing and joint-attention). Although proto-imperative pointing was originally considered to be normal in autism, a lack of such pointing was found to be an important predictor of PDD. In contrast, pretend play, which was originally considered to be a key item, was not included because a low pass rate was also observed for non-PDD children. As these results are based on a clinical sample, the findings have low generalizability; therefore, it is necessary to conduct replication studies in Japan and other countries.

In the present study, a limited number of children were incorrectly classified according to the cut-off value; therefore, it is important that clinicians not make a diagnosis based solely on a single score, which may be affected by caregiver recall bias. However, as these four items are quick to ask and easy to score, they can provide valuable information to clinicians when diagnosing PDD at age 2. Future studies should examine the applicability of these items for younger or older children.

**ACKNOWLEDGMENTS**

This study was supported in part by a Research Grant for Nervous and Mental Disorders (20B-5) from the Japanese Ministry of Health, Labour and Welfare. We thank all the staff members at the Nerima Welfare Center for Handicapped Persons, especially Dr Hiroshi Kurita.