Available online at www.sciencedirect.com
SCIENCE DIRECT
Comprehensive Psychiatry 47 (2006) 116–122
Comprehensive PSYCHIATRY
www.elsevier.com/locate/comppsych

The relations between DISC-IV *DSM* diagnoses of ADHD and multi-informant CBCL-AP syndrome scores☆

Eske M. Derks[a,]*, Jim J. Hudziak[b,c], Conor V. Dolan[d], Robert F. Ferdinand[e], Dorret I. Boomsma[a]

[a]*Department of Biological Psychology, Vrije Universiteit, 1081 BT Amsterdam, The Netherlands*
[b]*Department of Psychiatry and Medicine, Division of Human Genetics, Center for Children, Youth and Families, Burlington, VT 05401, USA*
[c]*College of Medicine, University of Vermont, Burlington, VT 05401, USA*
[d]*Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands*
[e]*Department of Child and Adolescent Psychiatry, Erasmus Medical Center/Sophia Children's Hospital, 3015 GJ Rotterdam, The Netherlands*

Abstract

Background: Previous studies have examined the relation between attention problems (APs) obtained with the Child Behavior Checklist (CBCL) and attention deficit hyperactivity disorder (ADHD) assessed with the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*). We will examine this relation across sex using multi-informant data.
Methods: Parents of 12 538 twins, aged 7, 10, and 12 years, and teachers of twins, aged 10 years, completed the questionnaires. The mothers of a sample of 283 boys and 291 girls who scored either low or high on longitudinal maternal CBCL-AP were interviewed.
Results: Children with a low AP score obtained a negative ADHD diagnosis in 96% of cases. Children with a high AP score obtained a positive diagnosis in 36% (girls) and 59% (boys) of cases. The association between paternal and maternal AP ratings and ADHD was the same, whereas the association between teacher AP ratings and ADHD was low.
Conclusions: The association between AP and ADHD is higher in boys than girls, possibly because of a bias toward the male manifestation of ADHD.

1. Introduction

Attention deficit hyperactivity disorder (ADHD) is defined by impaired attention skills, impulsivity, and hyperactivity across the life span [1]. It is one of the most commonly diagnosed behavioral disorders in children: the prevalence rates usually lie between 3% and 5% [2]. ADHD is diagnosed in a variety of ways, including both checklists and interviews. Interview methods are time-consuming and costly and are therefore difficult to use in large-scale studies. For this reason, epidemiological studies often use behavior checklists.

Several studies have shown that behavior checklists and *DSM* interviews measure overlapping constructs [3-10]. Correlations between attention problem (AP) scores and the number of *DSM* symptoms are moderate to high [11]. As expected, the correlations are higher when both measures are obtained from the same informant (the parents) than when measures are obtained from different informants (parents vs children) [12]. In addition, checklist scores discriminate between children with and without a *DSM* diagnosis of ADHD. Steinhausen et al [10] studied a population sample of 6- to 17-year-old children and adolescents and compared the mean Child Behavior Checklist (CBCL)-AP scores of 272 control children, and 35 children diagnosed with ADHD. Children with ADHD scored higher on parental AP than controls.

Other measures of association are the positive and negative predictive power (PPP and NPP, respectively), sensitivity, and specificity. PPP refers to the proportion of children with a high CBCL score who obtain a positive *DSM* diagnosis (ie, diagnosed as affected), and NPP refers to the proportion of children with a low CBCL score who obtain a negative *DSM* diagnosis (ie, unaffected). Sensitivity and specificity refer to the proportion of children with a

☆ This work was supported by NWO grants 575-25-006, 575-25-012, and 904-57-94 (DI Boomsma, principal investigator), NIMH grant MH58799 (JJ Hudziak, principal investigator), and by the Center for Neurogenomics and Cognitive Research.

* Corresponding author. Tel.: +31 20 598 8743; fax: +31 20 598 8832.
E-mail address: em.derks@psy.vu.nl (E.M. Derks).

0010-440X/$ – see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.comppsych.2005.05.006

positive *DSM* diagnosis, who score high on the CBCL, and the proportion of children with a negative *DSM* diagnosis, who score low on the CBCL.

Hudziak et al [6] examined the sensitivity and specificity of ADHD at different CBCL-AP cut-points in 101 male and 82 female siblings of referred children. They found that a *T* score of 55 minimized the number of false-positive and false-negative cases. The PPP and NPP were both high: 80% and 90%, respectively.

A drawback of the diagnostic efficiency measures (DES) is their dependence on the baseline prevalence of the disorder. This makes it difficult to compare the results of studies with different sampling procedures. Therefore, we summarized the DES statistics and the baseline prevalences of 6 studies that examined the convergence of CBCL-AP and ADHD in Table 1. In some studies, the DES statistics were not reported directly, and the statistics were derived from the reported number of children who scored below or above the CBCL cut-point by diagnosis. More details about these calculations may be obtained from the first author.

As can be seen in Table 1, Gould et al [5] found the lowest PPP. This may, in part, be because of a lower baseline prevalence in her sample. The chance of being diagnosed positive for ADHD was 23% in the total sample and 36% in children with a high CBCL score. The studies of Eiraldi et al [13] and Sprafkin et al [14] have been conducted on clinically referred samples and therefore report high baseline prevalences of ADHD (83% and 71%, respectively). They observed a PPP of 0.93 and 0.78, respectively. The high prevalence of ADHD decreases the chance of a false-positive case, but the specificity of the *DSM* diagnosis is much lower than in the Gould et al study [5]. These comparisons emphasize the fact that the value of a single statistic does not provide information about the association between a test and a diagnosis, whereas the combination of all 4 statistics does.

A low sensitivity was seen in the study of Lengua et al [15]. An examination of the 2 × 2 contingency table revealed that the logistic regression analyses predicted a positive *DSM* diagnosis in only 2 cases. The study of Lengua et al [15] illustrates that CBCL scores may not improve the prediction of ADHD if the PPP is lower than 50%. In summary, the convergence between AP and ADHD is moderate to high, depending on the kind of sample that is used. Because of differences in baseline prevalences, the PPP is higher in referred than in nonreferred samples.

Most studies that report on the convergence between checklists and *DSM* diagnoses have used the CBCL [11]. However, AP scales of the Devereux Scale of Mental Disorders and the Child Symptom Inventory–4 have also been shown to be good predictors of ADHD [13,14].

Although the convergence between AP and ADHD has been the focus of previous research, several questions remain unanswered. First, is the association between behavior checklists and ADHD similar in boys and girls? Chen et al [4] found a higher PPP in boys than girls, but the significance of this difference was not tested formally. Second, most studies have used clinically referred samples. It is not clear to what extent the results of these studies generalize to population samples. The study of Gould et al [5] is the only study that selected children from a general population. The association in this study was lower than the association in studies that examined referred samples. Finally, none of the studies included maternal, paternal, and teacher ratings. The inclusion of different informants might improve the prediction of *DSM* diagnoses, because each informant observes the child in a slightly different situation and may have unique interactions with the child.

The purpose of this study is to further investigate the association between CBCL-AP and *DSM-IV* ADHD. Equal numbers of boys and girls were selected from a general population sample on the basis of longitudinal maternal CBCL scores. Half of the subjects scored low on AP (*T* score below 55, controls), and the other half of the subjects scored high on AP (*T* score above 65, probands). The selection of probands diminishes the problem of underrepresentation of affected children. Data on maternal-, paternal-, and teacher-rated AP were collected when the children were 10 years old. Mothers were interviewed when the child was around the age of 12 years. Shortly after this diagnostic interview, mothers completed the CBCL. We addressed the following issues. First, is the cross-sectional association between AP and ADHD different in boys and girls? Second, can the prediction of ADHD be improved if AP scores from multiple informants are available?

Table 1
Diagnostic efficiency statistics of studies that examined the association between behavior checklist scores and ADHD

Study	Sample	*n* (boys/girls)	Cut-point	Prevalence (%)	PPP	NPP	SE	SP
Gould et al [5]	NR	157	$T > 65$	23	0.36	0.96	0.46	0.95
Chen et al [4]	SR	111/108	$T \geq 65$	16 (boys), 8 (girls)	1.00 (boys), 0.67 (girls)	0.86 (boys), 0.93 (girls)	0.17 (boys), 0.22 (girls)	1.00 (boys), 0.99 (girls)
Eiraldi et al [13]	R	192/50	$T \geq 65$	83	0.93	0.37	0.78	0.69
Lengua et al [15]	R	203	Based on regression	29	0.50	0.71	0.02	0.99
Sprafkin et al [14]	R	247/0	$T \geq 60$	71	0.78	0.83	0.97	0.33
Hudziak et al (in press)	SR	101/82	$T \geq 65$	36	0.97	0.76	0.47	0.99

R indicates clinically referred sample; NR, nonreferred sample; SR, siblings of referred children; SE, sensitivity; SP, specificity.

2. Methods

2.1. Subjects

The subjects are Dutch twins whose parents voluntarily registered with the Netherlands Twin Registry when the twins were born [16,17]. Twelve thousand five hundred thirty-eight twins are currently participating in a longitudinal study, in which surveys are sent to their parents and teachers. In this longitudinal study, mothers and fathers are asked to complete the CBCL. For the present study, we analyzed parental data of 10-year-old twins from cohorts 1986 to 1993. In addition, we included teacher data of 10-year-old twins from cohorts 1989 to 1991.

In the present study, 356 families from the cohorts 1989 to 1992 were selected based on the maternal AP scores obtained at age 7, 10, and 12 years. These families received a letter in which the mother of the twins was asked to participate with a structured interview study. Of these 356 families, 287 agreed to participate (80.6%), 64 families refused (18.0%), and 5 families did not respond and could not be contacted by telephone (1.4%). The sample of 287 twin pairs consisted of 283 boys and 291 girls. At the time of the interview, the twins were 10 to 13 years old, with a mean of 11.99 years.

2.2. Selection

Subjects were selected on the basis of their standardized maternal CBCL ratings (*T* scores) at the ages 7, 10, and 12 years. Subjects were excluded if maternal ratings were available only at one time, or if they suffered from a severe handicap, which disrupts daily functioning. Twin pairs were selected if at least one of the twins scored high on AP (probands) or if both twins scored low on AP (controls). A high score was defined as a *T* score above 60 at all available time points (ages 7, 10, and 12 years) and a *T* score above 65 at least once. A low score was defined as a *T* score below 55 at all available time points. The control twins were matched with proband twins on the basis of sex, cohort, maternal age, and social economic status. The criteria resulted in the selection of 3 types of children: children who scored low (controls), children who scored high (probands), and children who obtained an intermediate score (intermediate group). *T* scores were computed in boys and girls separately. In other words, girls were selected if they scored low or high compared with other girls, and boys were selected if they scored low or high compared with other boys. This procedure resulted in the selection of an equal number of boys and girls (283 and 291, respectively). The mean AP scores before and after selection are reported in Table 2.

2.3. Procedure

The selected families received a letter inviting them to participate. Mothers of twins who agreed to participate and who returned the informed consent form by mail were contacted by telephone to schedule the interview. At the agreed date and time, the mother was interviewed by telephone. Within 4 months after the interview, the mother completed a CBCL, which she received by mail.

2.4. Measures

The CBCL [18,19] is a standardized questionnaire for parents to report the frequency and intensity of behavioral and emotional problems exhibited by their child in the past 6 months. The CBCL 4/18 was completed by both parents when the children were 7, 10, and 12 years old and by the mother within 4 months after the interview. The AP scale (11 items) was used as a predictor for ADHD [20].

Teachers completed the Teacher's Report Form [21] when the children were 10 years old. The Dutch AP scale (20 items) was used in the present study. Ten items of the CBCL-AP scale and the Teacher's Report Form–AP scale overlap.

The Diagnostic Interview Schedule for Children–Version 4 (DISC-IV) [22] is a structured diagnostic interview. It can be used to assess the presence of *DSM-IV* diagnoses, including ADHD. The Dutch translation was obtained from Ferdinand and van der Ende [23]. The mothers of twins were interviewed by 2 experienced research assistants to determine which symptoms of ADHD were displayed by the twins during the last year. A child was diagnosed positive for ADHD if he or she met type-A criteria of the *DSM-IV*.

Table 2
Mean maternal AP scores before and after selection

		n	Mean maternal AP score (SD)		
			Aged 7 years	Aged 10 years	Aged 12 years
Boys					
Before selection	Total	6191	3.40 (3.12)	3.50 (3.32)	3.09 (3.16)
After selection	Proband group	56	10.58 (2.98)	11.20 (2.69)	12.00 (4.77)
	Intermediate group	53	6.73 (2.84)	6.71 (2.14)	4.11 (2.64)
	Control group	174	1.71 (1.41)	1.90 (1.56)	2.27 (2.20)
Girls					
Before selection	Total	6347	2.49 (2.66)	2.46 (2.66)	2.06 (2.36)
After selection	Proband group	84	8.40 (2.22)	8.35 (2.90)	8.00 (2.27)
	Intermediate group	46	4.83 (1.70)	5.11 (2.49)	3.33 (1.21)
	Control group	161	1.18 (1.11)	1.05 (.99)	1.35 (1.56)

No distinction was made between ADHD of the primarily inattentive type and ADHD of the primarily hyperactive/impulsive type. Two hundred forty-eight interviews were audiotaped. To assess the quality of the interviews, a research assistant, who was blind to the results of the interview, listened to 40 of these interviews. It was established that the interviewers had made no mistakes that could have altered the results of the interview, and the number of ADHD symptoms as scored by the interviewer and the research assistant were the same for all subjects.

2.5. Statistical analyses

All statistical analyses were performed on sum scores in SPSS/Windows 11.0 (SPSS, Chicago, IL) [24]. The total sample of 287 twin pairs was divided into 2 samples to avoid dependency inherent in twin and sibling data. To this end, the 2 members of a twin pair were randomly assigned to a different sample. The descriptive statistics are reported for the total sample, but the statistical tests were performed in each sample separately. This enabled a cross-validation of the results, although the cross-validation sample does not provide a completely independent replication, because of the correlation between the twins.

PPP, NPP, sensitivity, and specificity were computed to examine the cross-sectional association between the maternal CBCL AP score that was obtained shortly after the interview and ADHD. Children who obtained an intermediate score on the CBCL were excluded from the sensitivity analyses. They were included in all remaining analyses. To examine the discriminative power of the CBCL, mean AP scores of children with and without ADHD were computed. A 2-way analysis of variance was used to test for effects of diagnosis and sex. The type I error rate was corrected for multiple testing in 2 ways. First, the α (type I error probability) of each test was set at .01, which is more stringent than the usual α of .05. Second, an effect was only assumed present if it was significant, given that α is .01 in both random samples of twins. In addition, correlations between AP scores and number of ADHD symptoms were calculated. These correlations were corrected for restriction of range because of selection with Pearson-Lawley selection rules [25].

Linear stepwise regression analyses were performed to examine whether the inclusion of paternal or teacher data improved the prediction of the number of ADHD symptoms. An advantage of regression analysis is the possibility of including covariates such as sex and of investigating associated interaction effects (eg, between sex and CBCL scores). A significant interaction term suggests that the predictive value of AP is different in boys and girls.

Table 3
Number of children who obtained a diagnosis for ADHD by CBCL-high ($T > 65$) and CBCL-low ($T < 55$)

	Boys			Girls		
	No ADHD	ADHD	Total sample	No ADHD	ADHD	Total sample
CBCL low	151	7	158	155	5	160
CBCL high	14	20	34	36	20	56
Total sample	165	27	192	191	25	216

Table 4
PPP, NPP, specificity, and sensitivity of the CBCL

	PPP	NPP	Specificity	Sensitivity
Boys	59	96	92	74
Girls	36	97	81	80

3. Results

3.1. Predictive power, specificity, and sensitivity

The cross-sectional association between CBCL-AP and *DSM*-ADHD was examined by calculating the predictive power of AP and its specificity and sensitivity. The number of children by diagnosis and by CBCL-high vs CBCL-low is reported for each sex in Table 3.

Table 4 provides the diagnostic efficiency statistics. The high NPP in boys and girls suggests that a low score on the CBCL almost perfectly predicts a negative diagnosis for ADHD. In contrary, a high score on the CBCL correctly predicts the presence of ADHD in 59% of the boys and 36% of the girls. Log-linear tests showed that the percentages of boys and girls in the 4 categories were significantly different ($P < .05$ in both samples). Boys with a high AP score are more often diagnosed positive for ADHD than girls with a high AP score.

3.2. Discriminative power of the CBCL

To examine if children with ADHD obtain different AP scores than children without ADHD, mean AP scores by informant were computed. The results are reported in Table 5. Main effects of sex, diagnosis, and their interaction were examined with a 2-way analysis of variance.

Children with ADHD scored significantly higher than children without ADHD on both maternal and paternal AP scales. The AP scores obtained from teachers did not discriminate significantly between groups of children with and without ADHD, although the mean scores of boys with ADHD were almost twice as large as the mean scores of boys without ADHD.

3.3. Regression analyses

To examine the predictive value of AP, linear stepwise regression analyses were performed. The dependent variable was the total number of ADHD symptoms. We included maternal AP scores collected after the interview and paternal and teacher AP scores collected at age 10 years. The best fitting model in the first random sample was a model that includes both maternal and paternal AP scores plus an interaction effect between sex and maternal AP score. This model explained 55% of the variance in the number of ADHD symptoms. The inclusion of teacher data

Table 5
Mean AP scores of the total sample of boys and girls by diagnosis and correlations between AP and the number of ADHD symptoms

AP score	*n*				Mean AP score (SD)				Correlations	
	No ADHD		ADHD		No ADHD		ADHD			
	M	F	M	F	M	F	M	F	M	F
Maternal, aged 10 years	234	254	44	36	3.66 (3.44)	3.18 (3.37)	9.75 (4.24)	8.06 (3.83)	0.63	0.62
Paternal, aged 10 years	176	189	30	23	2.77 (2.74)	2.23 (2.71)	8.63 (4.26)	6.91 (3.55)	0.64	0.62
Teacher, aged 10 years	133	136	25	16	7.10 (7.19)	4.69 (6.04)	13.92 (7.69)	6.50 (6.45)	0.32	0.19
Maternal, after interview	196	220	34	30	2.88 (2.93)	2.61 (2.81)	8.91 (4.62)	7.53 (4.22)	0.69	0.68
Total number	238	255	45	36						

M indicates male; F, female.

did not improve the prediction of ADHD. The standardized regression weight of the maternal AP score was 0.47, of the paternal AP score 0.41, and of the interaction effect between sex and maternal AP −0.13. The negative regression weight of the interaction effect implies that for the same values of AP, girls are expected to have a lower number of ADHD symptoms than boys. The results of the same regression analyses in the second random sample were very similar. However, no significant interaction effect between sex and AP was found.

4. Discussion

We examined the relations between parental and teacher ratings on APs, and the interview-based DSM-ADHD diagnoses in a general population sample of boys and girls. Children were selected on the basis of their maternal CBCL scores at ages 7, 10, and 12 years. The mothers of these children were interviewed with a standardized *DSM* interview to verify the presence of ADHD.

4.1. The association between AP scores and DSM-ADHD

We examined the cross-sectional association between *DSM*-ADHD and maternal CBCL-AP and found that CBCL-AP is an excellent screening instrument for the absence of ADHD in a population sample. However, screening for the presence of ADHD is associated with a high proportion of false-positive cases. Of the total sample of children with a high maternal CBCL score, 59% of the boys and 36% of the girls were diagnosed positive for ADHD. The remaining children were diagnosed negative for ADHD. This suggests that the CBCL can be used as a screening instrument for ADHD and children who score high on the CBCL have to be examined with additional methods to verify if they indeed have ADHD.

The PPP was low compared with the PPP in most other studies (compare Tables 1 and 4). Because the PPP depends on the prevalence of the disorder, this can probably be explained by the fact that the prevalence of ADHD was much lower in our sample than the prevalence in clinically referred samples. The baseline prevalence of ADHD was 14% in boys and 12% in girls. In children with a high CBCL score, these percentages increased to 59% and 36%. In contrast, the baseline prevalence in studies of referred samples ranged from 29% to 83%. Because of these higher baseline rates, the PPP will be higher even if the association between the checklist and the interview is similar. The results of Gould et al [5], who studied a nonreferred sample in Puerto Rico, are in agreement with our results. The prevalence of ADHD in their sample was 23%; this percentage increased to 36% conditional on a high CBCL score. This supports the hypothesis that the different levels of PPP are caused by different baseline prevalences. Concluding, selecting children with ADHD from a general population sample is associated with more false-positive cases than selecting children with ADHD from a referred sample.

4.2. Sex differences in the association between AP and ADHD

Sex differences were found in most association measures. First, the PPP was higher in boys than in girls. This means that a boy with a high CBCL score has a higher chance of obtaining a positive diagnosis for ADHD than a girl with a high CBCL score. This is in agreement with the results of Chen et al [4] who also found higher PPP in boys than girls. Second, the correlations between AP and the number of ADHD symptoms were higher in boys than girls, regardless of informant. Third, regression analyses revealed significant negative interaction effects between AP and sex. The negative interaction effect implies that for a certain value of AP, the predicted number of ADHD symptoms is lower in girls than it is in boys. This is in agreement with the lower predictive power in girls than boys. A possible explanation for the higher association between AP and ADHD in boys than girls is that symptoms that are displayed by boys with ADHD are more likely to be included in the *DSM* diagnosis than symptoms that are displayed by girls with ADHD. This could be caused by the fact that the development of ADHD was based on data from studies with more boys than girls or because ADHD has a higher prevalence in boys than girls [26]. In other words, the diagnosis of ADHD may be biased toward the male manifestation of the syndrome. If this hypothesis is correct, this may result in the underidentification of girls with ADHD, as has already been suggested by Hudziak [27].

4.3. Multi-informant checklist data

The correlations between paternal AP scores and ADHD were just as high as the correlations between maternal AP scores and ADHD. Although fathers report specific and unique aspects of their children's behavior, we observed overall agreement between fathers' reports on a checklist and mothers' report during interview on ADHD behaviors.

An unexpected finding was that children with and without ADHD did not differ significantly with respect to their teacher ratings at 10 years old. The lower association between teachers' reports and *DSM* interviews may be because of a variety of factors, most prominently the fact that the *DSM* interview was collected from mothers and not teachers. Other factors include the different context of the children's behavior: teachers report on children's behavior in the classroom, whereas parents report on children's behavior elsewhere. The finding of higher cross-correlations between paternal and maternal reports than between parental and teacher reports [11] is in line with this observation. The low discriminative power of teacher data was confirmed by low correlations between teacher ratings and the number of *DSM* symptoms. The higher association between parental AP ratings and ADHD than between teacher AP ratings and ADHD may also reflect a better understanding of the child's behavior by the parents; the teachers had known the child only for an average of 12 months, whereas the parents had known the child since birth.

Regression analyses showed that the inclusion of paternal ratings slightly improves the prediction of ADHD. This suggests that a clinician should, whenever possible, obtain information from both parents when screening for the presence of behavior problems.

4.4. Limitations

This study concentrated on the association between behavior checklist scores and interview-based diagnoses. A limitation of this study is that the diagnosis for ADHD was only based on a *DSM* interview with the mother. The results could have been different if the father, teacher, or child was interviewed as well. However, the correlations between paternal AP scores and ADHD were just as high as correlations between maternal AP scores and ADHD.

Furthermore, we limited ourselves to the examination of the concurrent validity between 2 measures of ADHD and did not examine the external validity of these measures. Previous research has shown that a combination of CBCL scores and *DSM* diagnosis provides a better prediction of outcome measures (eg, disciplinary problems in school and receiving inpatient or outpatient treatment) than either one of these instruments alone [28]. Finally, we made the choice to examine children in the age range of 10 to 13 years. Because of the exclusion of children younger than 10 years or older than 13 years, we cannot tell if our results generalize to children within these age groups.

References

[1] American Psychiatric Association. Task force on DSM-IV. Diagnostic and statistical manual of mental disorder: DSM-IV. 4th ed. Washington (DC): American Psychiatric Association; 1994.

[2] Bradley JDD, Golden CJ. Biological contributions to the presentation and understanding of attention-deficit/hyperactivity disorder: a review. Clin Psychol Rev 2001;6:907-29.

[3] Biederman J, Faraone SV, Doyle A, Krifcher Lehman B, Kraus I, Perrin J, et al. Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. J Child Psychol Psychiatry 1993; 34:1241-51.

[4] Chen WJ, Faraone SV, Biederman J, Tsuang MT. Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit hyperactivity disorder: a receiver-operating characteristic analysis. J Consult Clin Psychol 1994;62:1017-25.

[5] Gould MS, Bird H, Staghezza Jaramillo B. Correspondence between statistically derived behavior problem syndromes and child psychiatric diagnoses in a community sample. J Abnorm Child Psychol 1993; 21:287-313.

[6] Hudziak JJ, Copeland W, Stanger C, Wadsworth M. Screening for DSM-IV externalizing disorders with the Child Behavior Checklist: a receiver-operating characteristic analysis. J Child Psychol Psychiatry 2004;45:1299-307.

[7] Jensen PS, Watanabe H. Sherlock Holmes and child psychopathology assessment approaches. The case of the false-positive. J Am Acad Child Adolesc Psychiatry 1998;38:138-46.

[8] Kasius MC, Ferdinand RF, van den Berg H, Verhulst FC. Associations between different diagnostic approaches for child and adolescent psychopathology. J Child Psychol Psychiatry 1997;38: 625-32.

[9] Kazdin AE, Heidish IE. Convergence of clinically derived diagnoses and parent checklists among inpatient children. J Abnorm Child Psychol 1984;12:421-36.

[10] Steinhausen HC, Winkler Metzke C, Meier M, Kannenberg R. Behavioral and emotional problems reported by parents for ages 6 to 17 in a Swiss epidemiological study. Eur Child Adolesc Psychiatry 1997;6:136-41.

[11] Achenbach TM, Rescorla LA. Manual for the ASEBA school-age forms and profiles. Burlington (Vt): University of Vermont, Research Center for Children, Youth and Families; 2001.

[12] De la Osa N, Ezpeleta L, Domenech JM, Navarro JB, Losilla JM. Convergent and discriminant validity of the structured Diagnostic Interview for Children and Adolescents (DICA-R). Psychol Spain 1997;1:37-44.

[13] Eiraldi RB, Power TJ, Karustis JL, Goldstein SG. Assessing ADHD and comorbid disorders in children: the Child Behavior Checklist and the Devereux Scales of Mental Disorders. J Clin Child Psychol 2000;29:3-16.

[14] Sprafkin J, Gadow KD, Salisbury H, Schneider J, Loney J. Further evidence of reliability and validity of the Child Symptom Inventory–4: parent checklist in clinically referred boys. J Clin Child Adolesc Psychol 2002;31:513-24.

[15] Lengua LJ, Sadowski CA, Friedrich WN, Fisher J. Rationally and empirically derived dimensions of children's symptomatology: expert ratings and confirmatory factor analyses of the CBCL. J Consult Clin Psychol 2001;69:683-98.

[16] Boomsma DI, Vink JM, Beijsterveldt van CEM, Geus EJC, Beem AL, Mulder EJCM, et al. Netherlands Twin Register: a focus on longitudinal research. Twin Res 2002;5:1-6.

[17] Boomsma DI. Twin registers in Europe: an overview. Twin Res 1998;1:34-51.

[18] Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington (Vt): University of Vermont, Department of Psychiatry; 1991.

[19] Achenbach TM. Manual for the Child Behavior Checklist/2-3 and 1992 profile. Burlington (Vt): University of Vermont, Department of Psychiatry; 1992.
[20] Verhulst FC, Ende van der J, Koot HM. Handleiding voor de CBCL/ 4-18. Rotterdam: Afdeling Kinder-en jeugdpsychiatrie. Rotterdam: Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam/Erasmus Universiteit; 1996.
[21] Achenbach TM. Manual for the teacher's report form and 1991 profile. Burlington (Vt): University of Vermont, Department of Psychiatry; 1991.
[22] Shaffer D, Schwab-Stone M, Fisher P, Cohen P, Piacentini J, Davies M, et al. The Diagnostic Interview Schedule for Children–Revised Version (DISC-R): I. Preparation, field testing, interrater reliability, and acceptability. J Am Acad Child Adolesc Psychiatry 1993;32:643-50.
[23] Ferdinand RF, van der Ende J. Diagnostic interview schedule for children IV parent version. Rotterdam: Erasmus University Rotterdam, Department of Child and Adolescent Psychiatry; 1998.
[24] SPSS. SPSS for Windows, Rel 11.0.1. Chicago (Ill): SPSS; 2001.
[25] Lawley DN. A note on Karl Pearson's selection formulae. Proc R Soc Edinb Sect A 1943;62:28-30.
[26] Gaub M, Carlson CL. Sex differences in ADHD: a meta-analysis and critical review. J Am Acad Child Adolesc Psychiatry 1997;36: 1036-45.
[27] Hudziak JJ. Genetics of attention-deficit/hyperactivity disorder. In: Brown TE, editor. Attention-deficit disorders and comorbidities in children, adolescents, and adults. Washington (DC): American Psychiatric Press; 2000. p. 57-78.
[28] Ferdinand RF, Heijmens Visser J, Hoogerheide KN, Ende J, Kasius MC, Koot HM, et al. Improving estimation of the prognosis of childhood psychopathology; combination of *DSM-III-R*/DISC diagnoses and CBCL scores. J Child Psychol Psychiatry 2003;44:1-10.