Mental Health: Research and Practice

ISSN: 2980-8502 2022, Vol. 1, No. 1, 1-13
DOI: 10.22034/MHRP.2022.154060

Multi-Method Diagnosis of ADHD in Children: Psychometric and ROC Properties of Persian ADHD Diagnostics

Saeed Azami✉ and Hadi Safari
Semnan University, Semnan, Iran

Hadis Aghaei
Department of Education, Karaj, Iran

Fatemeh Hemmati
Tarbiat Modares University, Tehran, Iran

Mahdieh Mirmohammad
Tarbiat Modares University, Tehran, Iran

Ameneh Ahmadi
University of Science and Culture, Tehran, Iran

There has been little research on the use of multi-method assessment approaches for differential diagnosis of ADHD in non-English speaking youth samples. A multi-method assessment approach is suggested to be used with clinically referred Iranian children. 8-12 years old Iranian boys (N= 80: 40 diagnosed with ADHD and 40 diagnosed with a non-ADHD diagnosis) who were referred from the clinics of the Shahriar district of Karaj were assessed via a multi-method assessment approach comprised of: a semi-structured interview, the child symptom inventory-4, the 27-item Conners' parent rating scale, the Swanson, Nolan and Pelham-IV scale all had been previously translated and adapted into Persian, as well as some computerized executive functions tests. The time frame for data collection was October 22, 2016, to January 29, 2017. A 15-minute behavioral observation also was used. Inter-rater reliability across trained diagnosticians on the battery was high (Kappa=.825). Observed internal consistencies and test-retest reliabilities of the rating scales were good to excellent. Both the rating scales and the EF tests could significantly differentiate children with ADHD from those with other non-ADHD disorders (all p's ≤ 0.05). The sensitivity analyses indicated that the rating scales differentiated the groups with excellent sensitivities and specificities. The diagnostic sensitivity of the EF tests, however, showed considerably more variability. The adapted, Persian multi-method assessment approach is psychometrically sound and sensitive diagnostically. Parent ratings, alongside other direct measures, are useful diagnostic tools in ADHD assessment. However, caution should be exercised when using EF tests in isolation.

Keywords: ADHD, Multi-Method Assessment, Multi-Cultural Considerations

Although Attention-Deficit Hyperactivity Disorder (ADHD) has been recognized as a psychiatric disorder since the second edition of the DSM (APA, 1968), in which it was termed 'hyperkinetic impulse disorder,' it remains among the most controversial disorders (Kollins, 2007; Mayes et al., 2008). In 2011, over five million American children aged 3-17 years were diagnosed with ADHD (Bloom et al., 2012). Also, 3 to 6% of school-aged children have diagnosed with ADHD in Tehran (Khushabi et al., 2006). Controversy over ADHD revolves primarily around the issues of possible over-diagnosis. Over-diagnosis (i.e., a 'false positive') of ADHD is more likely to happen when relying upon a single source (for example, only parents) or a single diagnostic indicator (Hersen, 2011). Children diagnosed with ADHD may show diminished hyperactivity in doctors' office, as children are typically better-behaved during visits to the doctor than they are at home with their parents (Barkley, 2015; Sleator & Ullmann, 1981). As such,

✉ Correspondence should be addressed to Dr. Saeed Azami, PhD. in Psychology, Department of Psychology, Semnan University, Semnan, Iran; Email: sa.azamy@gmail.com.

physicians may rely heavily, sometimes perhaps exclusively, on either their intuition (Bruchmuller et al., 2012) or on parent reports (Sciutto & Eisenberg, 2007) when diagnosing ADHD, which can contribute to a poor cross-clinician diagnostic agreement (Diller, 2015). Concerns have been raised that some behaviors in normally developing children such as motor activity, distractibility, or impulsivity, are misinterpreted as ADHD (Vitiello, 2001). Finally, comorbidity may further complicate the diagnosis of ADHD, given that some disorders (e.g., autism, anxiety disorders, specific learning disorders) can produce symptoms similar to ADHD (Centers for Disease Control and Prevention, 2013). This is a notable concern given upwards of 50% of children diagnosed with ADHD have co-occurring psychiatric diagnoses (Jensen & Steinhausen, 2015).

Given the concerns above and the fact that there exists no single definitive psychological or biological test for the diagnosis of ADHD, assessment should involve a multi-method evaluation that relies on information from several sources through several procedures (National Collaborating Centre for Mental Health, 2009). Barkley (2015) suggested a multi-method assessment comprised of clinical interviews with parents, teachers and the child, a medical examination, rating scales (from the parent, teacher, and child), and direct observations of the child in academic situations. Additional, optional assessments that can help inform final diagnosis include cognitive, neuropsychological, and developmental functioning measures. Each of the components of such a multi-method assessment should have sound psychometric properties to optimize diagnostic accuracy.

Behavioral rating scales are useful in an ADHD assessment. They have been among the most efficient methods for obtaining information from children’s caregivers (Wolraich et al., 2003). The Revised 27-item Conners Parent Rating Scale (CPRS: R-S) (Conners, 1997), Child Symptom Inventory-4 [CSI-4] (Gadow & Sprafkin, 2002), and MTA version of Swanson, Nolan and Pelham-IV questionnaire [MTA SNAP-IV] (Swanson et al., 2001) are among the most widely used rating scales. There are several methodologically rigorous studies supporting the psychometric properties of these rating scales (Bussing et al., 2008; Conners, 1997; Gadow & Sprafkin, 2002). Although nearly all of this research has been done with the original English language versions of the scales, the clinical validity and psychometric robustness of the Persian translations has begun to be evaluated. For example, Sadrossadat et al. (2007) investigated the validity and reliability of a Persian translation of the MTA SNAP-IV scale and documented its clinical utility for Iranians. Mohammad-Esmail (2007) studied the sensitivity, discriminant validity, and internal consistency of a Persian translation of CSI-4 and reported good sensitivity and specificity as well as proper discriminant validity for most of its subscales. The clinical utility, including psychometric properties and diagnostic sensitivity, of the recommended multi-method assessment battery (rating scales, interviews, observations), when translated into Persian and used in an Iranian sample, have yet to be examined.

Some researchers have recommended the use of computerized neuropsychological tests designed to to assess specific aspects of ADHD. Examples of such computerized tests include the Gordon Diagnostic System (Gordon et al., 1996), (Coners & MSH Staff, 2000). However, Brock and Clinton (2007) noted that using neuropsychological tests as part of ADHD diagnostic evaluation is controversial, because their sensitivity and specificity, when used in isolation for diagnostic purposes, has not been established (Barkley, 2014; Doyle et al., 2000; McGee et al., 2000). Nevertheless, groups with ADHD and non-ADHD disorders have shown significant mean differences on such neuropsychological tests (Dunn & Kronenberger, 2003).

The present study was undertaken to investigate the inter-rater agreement among diagnosticians using a multi-method assessment procedure in driving ADHD diagnosis. The discriminant validities, sensitivities, specificities, internal consistencies and test-retest reliabilities of parent rating scales, and computerized neuropsychological tests (EF tests) were also investigated.

Method

Participants

Participants consisted of 80 Iranian male students (aged between 8-12 years) from 10 primary schools of the Shahriar area of Karaj, a county located immediately west of Tehran with a population of 516,022 (Statistical Center of Iran, 2006). Students between 8 to 12 years were selected because, in Iran, this is the usual age of receiving ADHD diagnosis for the first time. The participants had been referred to clinics because of their behavioral and/or emotional problems. Half the sample diagnosed with ADHD (n = 40) and another half diagnosed with other non-ADHD disorders (n = 40). Table 1 shows the demographic characteristics of the participants.

Measurement Instruments

Persian translation of Parent Checklist of the Child Symptom Inventory-4 [CSI-4]: An adapted CSI-4 (Gadow & Sprafkin, 2002), a 97-item rating scale that is used to screen for 15 different categories of DSM-IV emotional and behavioral disorders in children (Checkmateplus, 2015), was used in the current study to obtain parent ratings of their children's behavior. As previously noted, Mohammad-Esmail (2007) reported good psychometric properties for a Persian translation of the parent checklist of CSI-4.

Persian translation of The MTA version of Swanson, Nolan, and Pelham-4 [MTA SNAP-IV]: The first 18 items of MTA SNAP-IV (Swanson et al., 2001) were used to obtain parent ratings of behavioral symptoms of ADHD in children (the remaining eight-items assess the symptoms of oppositional defiant disorder). The authors translated this rating scale into Persian, with some language adaptations for the Persian speakers.

Persian continuous performance test [Persian CPT] (Khodadadi et al., 2009a): Persian CPT is a visual CPT test, similar to the original CPT (Beck et al., 1956), with visual symbols or digits as stimuli. It presents a total of 150 stimuli with 1000 milliseconds [MS] interval between every two stimuli (ISI), and 250 MS stimulus display time [DT]. Persian CPT was developed by Khodadadi et al. (2009a) in Iran to measure the sustained attention and response inhibition (correct responses were used as an indicator of sustained attention, omission errors as an indicator of problems in sustained attention, and commission errors as an indicator of problems in response inhibition). Nazifi et al. (2011) found that the Persian CPT could significantly differentiate between children with and without ADHD.

Persian Tower of London Test [Persian TOL]: The Persian TOL is a computerized version of the Tower of London Test (TOL) (Krikorian et al., 1994) which was developed in Iran by Khodadadi et al. (2009b) to measure the planning skills. Mashhadi et al. (2010) found that the Persian TOL could significantly differentiate between children with and without ADHD.

Persian Stroop color-word test [Persian Stroop]: Developed by Khodadadi et al. (2009c) to assess executive interference control, Persian Stroop presents a total of 96 Persian colored words including 48 congruent Persian colored words (e.g., the word **Blue** in blue ink) and 48 incongruent (e.g., the word **Blue** in red ink). The DTs are 2000 MS, and the inter-stimulus intervals (ISI) are 800 MS. The interference effect is calculated using the number of correct congruent color namings minus the number of incongruent color namings. The higher the interference effect the greater the child's problems in controlling interference (Khodadadi et al., 2009c). Mashhadi et al. (2009) found that Persian Stroop could significantly differentiate between children with and without ADHD.

Forward/reverse digit span from the WISC-R testing battery: Forward and reverse digit span (Wechsler, 1974) were used to measure short-term and verbal working memory, respectively. Shahim (1991) translated WISC-R into Persian, adapted it for Iranians, and studied its psychometric properties and norms in a community sample in Shiraz. She reported good psychometric properties for the Persian translation of WISC-R (Shahim, 1991).

Span board task: The span board task (Lumos-Labs, 2012) from the Lumosity.com internet-based software is an online, computerized version of the original span board task from the WAIS-RNI battery (Kaplan et al., 1991) which was used to measure visuospatial working memory (WM). In this task, the child is asked to remember the locations of cards he or she has just seen in either forward or reverse order.

Procedures

After indicating interest in the study, parents and their children were given ample time to review the informed consent and assent documents, and all participants consented before participating. In the consent form, it was emphasized that “parents and their children agree that they will not ask researchers any feedback about the results until the end of the study where researchers themselves would inform them about their children’s results” in order to prevent introduction of new treatments or educational accommodations (i.e., any services that could influence assessment findings at Time 2 evaluation) during the one-month interval between assessments.

Two graduate students in psychology were trained to administer the full battery for diagnostic decisions with children who were referred to a the clinics in Shahriar. Each of our main diagnosticians was helped by another trained graduate student in psychology administering the assessments. As such, children were assessed by separate clinician-dyads at different times with a one-month interval between the two assessments.

All participants were children clinically referred for some kinds of emotional or behavioral problems or disturbances that prompted evaluation at a general outpatient clinic. The first team of trained diagnosticians administered the multi-method assessment battery with 157 of the referred children and their parents until 40 children diagnosed with ADHD and 40 with a disorder other than ADHD. Child participants in the group with ADHD also met symptoms of other comorbid disorders such as oppositional defiant disorder, specific learning disorder, tic, sleep disorders, and anxiety.

All 80 children again were then referred to the second team of trained diagnosticians one month later. The second team had not seen these children before and did not have any information about their diagnoses. They administered the multi-method assessment with these children and their parents again and diagnosed children with or without ADHD, based on their evaluation results. The multi-method assessment procedure included all the measurement instruments described above. In addition, both clinician teams conducted a brief, 15-minute, unstructured observations with the children while they were engaged in a writing task. During the observations, whenever the child moved from his or her seat, the observer would have instructed the child to come back to the seat. These observations were conducted to provide first-hand observational data about each child in a naturalistic setting emulating the school environment, to inform the final diagnostic decision.

Data Analysis

Interrater agreement between the diagnostic decisions made by the two independent diagnostician teams was assessed via inter-rater reliability (Kappa). Also, discriminant validity was studied by running MANOVAs. The time 1 categorical diagnoses (or the group membership) were independent variables, and the time 2 data obtained from the ADHD diagnostic instruments were dependent variables in these MANOVAs. Partial eta-squared (η_p^2) was reported as effect size estimates for each MANOVA. However, for those instruments which provided only one measure, independent *t*-tests were used. Post-hoc univariate ANOVAs were used after each MANOVA, with Bonferroni corrections as an adjustment for Type I error (Rom, 1990). Cohen’s *d* was also provided for each comparison as an effect size estimate (Cohen, 1988). With a sample size of $N = 80$ and alpha set at .01, the statistical power was .70 for detecting a medium effect size. The diagnostic sensitivity of the parent ratings and EF tests also investigated using Receiver Operating Characteristics analysis [ROC]. Selected cut-off scores were scores that produced the largest sensitivity and specificity values simultaneously. Finally, the internal consistency of the rating scales and the test-retest reliability of both the rating scales and the EF tests with one-month interval between time 1 and time 2 assessments were investigated.

Results

Table 1 shows the demographic characteristics of the participants. Group differences were only significant for age and school grade, with higher age and school grades in the group with ADHD compared to non-ADHD. It should be noted that the first clinician dyad was instructed to select as much as possible children with non-ADHD disorders who were demographically matched with the participants in the group with ADHD. The inter-rater reliability between the diagnostic teams who used our multi-method assessment was found to be strong, Kappa = .825 ($p < .001$).

Table 1
Demographic Characteristics of the Participants

Characteristics	With ADHD[a]	Without ADHD[b]	*t or* χ^2	*p*
M[c] Age (SD)	9.82 (1.15)	9.2 (1.13)	2.442	.017
Sex, *n* (%)				
Male	40(100)	40(100)		
Female	0(0)	0(0)		
School Grade, *n* (%)				
Second	12(30)	18(45)		
Third	14(35)	8(20)	10.473	.015
Fourth	8(20)	14(35)		
Fifth	6(15)	0(0)		
Ethnicity				
Persian	30	30		
Turk	9	10	1.053	.591
Kurd	1	0		
Arab	0	0		
Parental Education				
Diploma	30	26		
Bachelor's Degree	10	12	2.468	.291
Master's/Doctoral Degree	0	2		
Marital Status of Parents				
Single (Never Married, Divorced, Separated, Deceased Partner)	8	5	.827	.363
Married (2 Cohabitating Parents)	32	35		
Annual Family Income				
<80,000,000[d]	30	24gvb		
160,000,000 to 480.000.000[d]	8	12	2.133	.344
≥480,000,000[d]	2	4		

Note. a: *n*=40; b: *n*=40; c: M=Mean; d: the monetary unite is Iranian Rails.

Table 2 shows means and standard deviations of the diagnostic measures for groups diagnosed with ADHD and non-ADHD as well as the results of between subject comparisons. The MANOVA test for sustained attention and response inhibition (CPT variables) was significant, F (4, 75) = 5.341, $p< .001$, and $\eta_p^2= .222$. The MANOVA for planning (the total score of the Persian TOL and the total false responses), was significant, F (3, 76) = 5.196, $p < .01$, and $\eta_p^2= .181$. The MANOVA for short-term memory (forward/backward digit span and span board tests) was significant, F (3, 76) = 9.812, $p < .001$, and $\eta_p^2= .27$. the MANOVA for interference control (Persian Stroop interference, and consistent/ inconsistent errors) was significant, F (3, 76) = 3.052, $p < .05$, and $\eta_p^2= .108$. And finally, the MANOVA for parent-rated ADHD severity (MTA SNAP-IV) was significant, F (2, 77)= 74.255, $p< .0001$, and $\eta_p^2= .659$.

Table 3 displays the results of ROC analyses for ADHD diagnostic measures. The ADHD-C scale of SNAP-IV exhibited the largest AUC of .948 ($p <$ .001). Optimum cut-off for SNAP-IV's ADHD-C scale in differentiating children with and without ADHD was 18.5. The sensitivity or the true positive rate (TPR) at this cut-off point was 89% and the specificity or the true negative rate (TNR) was 90%. These findings indicated excellent diagnostic precision for the Persian SNAP-IV's ADHD-C scale with respect to very low false positive rate (FPR= 11%) as well as very low false negative rate (FNR= 10%) (Greiner et al., 2000; Zou et al., 2007). ROC curve analyses on the remaining ADHD diagnostic measures were then conducted to investigate each measure's diagnostic power in differentiating children with and without ADHD (see Table 3).

Table 2

Means, Standard Deviations, and Between Group Comparisons for ADHD Diagnostic Measures

Diagnostic Measures	With ADHD	Without ADHD	Comparisons		Between Group ES
	Mean (SD)	Mean (SD)	F	*p*	[95% CI]
CPT's Total Correct Responses	140.52(12.48)	145.85 (5.82)	5.982	.017*	.554 [-2.661, 1.553]
CPT's Omission Errors	4.17 (3.63)	1.325 (2.97)	14.83	.001**	.871 [.155, 1.588]
CPT's Commission Errors	5.25 (11.4)	2.82 (4.12)	1.6	.21	.287 [-1.57, 2.141]
CPT's Reaction Time (RT)	655.55 (156.6)	595.02 (90.75)	4.474	.038*	.479 [-27.21, 28.167]
TOL's Total Scores	26.64 (3.107)	27.8 (3.196)	2.66	.1	.373 [-17.22, 47.27]
TOL's Total False Responses	17.55 (7.581)	20.84 (7.53)	3.75	.05*	.441 [-.04, 6.64]
Forward Digit Span	4.37 (1.314)	6.7 (2.398)	28.9	.001***	1.22 [-3.18, -1.46]
Reversed Digit Span	3.25 (1.171)	4.57 (1.865)	14.47	.001***	.858 [-2.01, -.63]
Span Board	5.25 (1.705)	6.22 (1.609)	6.91	.01**	.593 [-1.71, -.23]
Stroop's Interference	2.85 (4.276)	3.44 (4.494)	0.475	.4	.136 [-2.62, 1.27]
Stroop's Consistent Errors	1.57 (1.985)	0.5 (.83)	6.177	.015*	.714 [.17, 1.62]
Stroop's Inconsistent Errors	2.67 (3.661)	1.68 (2)	2.056	.1	.340 [-.36, 2.26]
SNAP-IV's ADHD-PI	17.1 (4.55)	5.0 (4.48)	143.68	.001***	2.714 [1.737, 3.691]
SNAP-IV's ADHD-PHI	14.1 (6.299)	4.9 (4.93)	52.917	.001***	1. 647 [.432, 2.87]
SNAP-IV's ADHD-C	31.2 (8.486)	9.9 (8.63)	123.86	.001***	2.521 [.669, 4.372]
			t	*p*	
Raven's IQ	22.6 (4.242)	27.12 (4.936)	4.317	.001***	.977 [-6.611, -2.438]
CSI-IV's ADHD Scale	28.15 (9.55)	8.1 (5.22)	11.64	.001***	2.638 [16, 23]

Note. *p*= significance level; ES= effect sizes (Cohen's *d*); CI= confidence interval; ADHD-PI= ADHD Predominantly Inattentive; ADHD-PHI= ADHD Predominantly Hyperactive/Impulsive; ADHD-C= ADHD Combined.

Table 3

The results of ROC analysis for ADHD Diagnostic Measures

	Cut off	*AUC*	*Sensitivity*	*Specificity*	*PPP*	*NPP*
SNAP-IV's ADHD-PI	9.5	.943	.94	.928	.928	.939
SNAP-IV's ADHD-PHI	9.5	.856	.74	.902	.883	.776
SNAP-IV's ADHD-C	18.5	.948	.89	.902	.90	.891
CSI-IV's ADHD Scale	17.5	.946	.89	.902	.90	.891
Forward Digit Span	5.5	.801	.718	.70	.705	.712
Backward Digit Span	3.5	.706	.59	.692	.657	.627
Span Board	6.5	.640	.795	.45	.591	.687
Raven's IQ	23.5	.769	.615	.829	.782	.682
CPT's Total Correct Responses	141.5	.625	.410	.732	.604	.553
CPT's Omission Errors	1.5	.636	.622	.60	.608	.613
CPT's Commission Errors	5.5	.591	.282	.805	.591	.528
CPT's Reaction Time (RT)	620.5	.686	.641	.683	.669	.655
TOL's Scores	25.5	.615	.410	.805	.677	.577
TOL's False Responses	19.5	.669	.595	.675	.646	.625
Stroop's Interference	1.5	.464	.579	.425	.501	.502
Stroop's Consistent Errors	.5	.677	.725	.63	.662	.696
Stroop's Inconsistent Errors	2.5	.551	.350	.75	.583	.535

Note. *AUC*=Area under curve; *PPP*=Positive predictive power; *NPP*=Negative predictive power.

Table 4 shows the internal consistencies of parent rating scales used in the current study. As shown in table 4, all the Cronbach Alphas were above .936. Table 5 shows the test-retest reliabilities of the ADHD diagnostic measures. The parent ratings showed strong test-retest reliabilities. However, there were differential results with respect to the EF tests with some showing acceptable test-retest reliabilities while others were less stable over a one-month interval.

Table 4
Internal Consistencies of Parent Rating Scales

	Cronbach Alpha	*Number of Items*
SNAP-IV's ADHD-PI	.953	9
SNAP-IV's ADHD-PHI	.936	9
SNAP-IV's ADHD-C	.956	18
CSI-IV's ADHD Scale	.953	18

Table 5
Test-retest Reliabilities of ADHD Diagnostic Measures

	r	*p*	*95% CI*
SNAP-IV's ADHD-PI	.90	.001***	.826, .952
SNAP-IV's ADHD-PHI	.86	.001***	.636, .953
SNAP-IV's ADHD-C	.89	.001***	.728, .967
CSI-IV's ADHD Scale	.93	.001***	.906, .972
Forward Digit Span	.67	.001***	.425, .847
Reverse Digit Span	.44	.001***	.308, .680
Span Board	.68	.001***	.534, .836
Raven's IQ	.75	.001***	.609, .868
CPT's Total Correct Responses	.75	.001***	.634, .891
CPT's Omission Errors	.68	.001***	.478, .899
CPT's Commission Errors	.39	.001***	.343, .730
CPT's Reaction Time (RT)	.54	.001***	.224, .880
TOL's Scores	-.15	.1	-.333, .583
TOL's False Responses	.34	.002**	.167, .577
Stroop's Interference	.33	.002**	-.066, .402
Stroop's Consistent Errors	.387	.001***	.065, .604
Stroop's Inconsistent Errors	.184	.1	.026, .500

Note. *r* = test-retest reliability; *p*= significance level

Discussion

A common agreement among researchers is that the assessment of childhood disorders should make use of multi-method and multi-informant measurements (Johnston & Murray, 2003; Mash & Barkley, 2007). This is especially important for ADHD assessment because symptoms vary across different situations (McConaughy et al., 2010). Diagnosticians, who used our multi-method assessment, showed excellent inter-rater reliability of Kappa= .825. This coefficient could possibly be increased if more objective behavioral observations, as well as teacher ratings, had been used in the diagnostic battery. Likewise, such strong inter-rater agreement may not have been achieved if clinicians were to rely on less information or data from a single source. Since different informants (e.g., parents, children and clinicians) have their own unique perspectives, often arising from different situations or expectations, agreement with respect to the nature of the problems or symptoms tends to be low (Achenbach et al., 1987; De Los Reyes & Kazdin, 2005; Kazdin, 2005). The richness and complexity of such a multi-informant, multi-method evaluation underscores its diagnostic importance.

The observed between-group differences provide evidence for discriminant validity of the different instruments used in this study. Parent rating scales developed according to DSM conceptualization of ADHD are among vigorous discriminators of children with ADHD from other treatment-seeking children, with pooled Cohen's effect sizes ranging from 1.647 to 2.714. These findings are consistent with several review studies on the psychometric properties of ADHD rating scales, which collectively support the clinical validity of parent-report behavioral scales (Alda & Serrano-Troncoso, 2013; Collett et al., 2003; Pelham et al., 2005).

Besides, parent rating scales showed excellent sensitivity and specificity in diagnosing children with ADHD, with minimal over-identification (false positives), and AUCs ranging from .856 to .948, sensitivities ranging from .74 to .94, and specificities ranging from .902 to .928. Alda and Serrano-Troncoso (2013) have reached similar conclusions using the SNAP-IV scale. They reported an acceptable sensitivity and specificity of 82.3% and 82.4%, respectively, when SNAP-IV ratings were validated against categorical diagnoses of ADHD by pediatricians. Grañana et al. (2011) also found acceptable sensitivities and specificities of 72.7% and 65.3% respectively for inattention subscale and 86.4% and 73.5% for hyperactivity/impulsivity subscale of SNAP-IV against categorical diagnoses. Our findings were also consistent with those of Collett et al. (2003). The internal consistencies and test-retest reliabilities of these parent rating scales were also excellent.

In contrast to the support for the clinical validity of the parent-report scales, the present study's results were variable for the tests of executive functions. Correct responses, omission errors, and RTs from the Persian CPT, false responses from the Persian TOL, forward and reverse digit span from the WISC-R battrey, span board test, and finally consistent errors from Persian Stroop could significantly differentiate children with ADHD from other treatment-seeking children with pooled Cohen's effect sizes ranging from .441 to 1.22. However, commission errors from the Persian CPT, total score from the Persian TOL, and interference and inconsistent errors from the Persian Stroop could not significantly differentiate children with ADHD in this sample.

According to a literature review on CPT tasks, Riccio et al. (2001) noted that commission errors differentiated between groups with and without ADHD more robustly than did omission errors, and that commission errors could be the best discriminators of ADHD. However, in the present study, neither omission nor commission errors were useful in distinguishing groups. The authors believe these results may be a reflection of the task variables in the Persian CPT. It seems that our Persian CPT, with only 150 stimuli (a relatively short time on task), 1000 milliseconds ISI, 200 milliseconds DT, and only 20% of target stimulus, may not be completely successful in distinguishing children with ADHD and it may require some modifications in these task variables. Corkum and Siegel (1993) noted that CPT tasks might best differentiate children with and without ADHD when they pose a heavy attentional demand on the child. For example, CPT tasks with short DT, relatively short ISI, and higher percentages of target variables are better in distinguishing ADHD (Corkum & Siegel, 1993). Although omission errors and reaction times (RTs) of Persian CPT did differentiate children with and without ADHD in group comparisons with Cohen's *d* of .871 and .479 respectively, the categorical discrimination was not acceptable (see table 3). The authors believe that 20% of target variables are relatively infrequent, and it may decrease the discriminative power of commission errors, because of the diminished attentional demand on the child. Hence, the authors suggest using more frequent target variables as well as more time on task in CPT tests to improve the discrimination ability of commission errors.

The total score of the Persian TOL did not show appropriate discriminant validity in our study, and it could not differentiate children with ADHD. This was, to some extent, inconsistent with review studies on the tower tasks' abilities in differentiating children with ADHD (Klorman et al., 1999; Nigg et al., 2002; Pennington & Ozonoff, 1996; Willcutt et al., 2005). The additional demands of the computerized Persian TOL beyond the common tower tasks may have been responsible for these inconsistent results. Specifically, because computerized versions of TOL require children to use the computer mouse to respond, they impose the demand of using fine motor skills (Wei et al., 2014) beside the planning skills in order to perform successfully in solving the problems. Further, temporarily inappropriate working of the computer mouse due to possible hardware problems or the problem of inadequate space for moving the mouse (because of using a mouse pad), may have affected children's performance in our study. Nevertheless, the Persian TOL's false responses did significantly differentiate children with and without ADHD but the effect size was relatively small (pooled Cohen's *d* = .441). Test-retest stability was also clearly low (see table 5). The authors think these unimpressive psychometric properties may reflect the additional demands involved in the computerized version of the TOL, which may obfuscate the more direct assessment of children's ability to plan as assessed with the traditional, manual version of the TOL.

Forward and reverse digit span tests could adequately discriminate children with ADHD. These digit span tests were also successful in categorical discrimination, findings that are consistent with those of prior studies (Babikian et al., 2006; Iverson & Tulsky, 2003; Mathias et al., 2002). Wechsler's digit span is a short-term memory task, and it can impose some attentional and processing demands on children, which seem to be especially impaired in children with ADHD. Forward digit span showed acceptable test-retest reliability of .67 (see table 5). Greiffenstein et al. (1994) similarly found that the forward digit span has proper test-retest reliability. Span board test, a visuospatial working memory (WM) test, also could significantly differentiate children with ADHD. Span board showed adequate discriminative power in diagnosing ADHD and its test-retest reliability was good. These findings were consistent with Westerberg et al. (2004) in that visuospatial WM tasks could effectively discriminate children with ADHD.

Raven's colored progressive matrices also could significantly differentiate children with ADHD with pooled Cohen's *d* of .977. ROC analysis also showed that Raven's IQ score could diagnose ADHD with some acceptable sensitivity and specificity rates. Mahone et al. (2002) similarly found that among children without intellectual disability, IQ could differentiate children with and without ADHD, but in the higher than average or gifted levels, there were no IQ differences between children with and without ADHD. Frazier et al. (2004), in their meta-analytic study of intellectual and neuropsychological test performance in diagnosing ADHD, concluded that children with ADHD have lower assessed IQ scores than those without ADHD. However, a person can be diagnosed with ADHD in the presence of a high IQ (Antshel et al., 2007). Among those with very high ability, IQ cannot be seen as a useful indicator to differentiate the presence of ADHD. However, Raven's test is arguably more a measure of executive function (EF) than actual intelligence, which is largely believed to be much more multifaceted, based on both fluid thinking abilities and crystallized learning. Raven's problems require skills such as spatial reasoning, matching shapes with a whole design, and concentration. Hence, our study showed that, as an EF measure, Raven task could be a good index for discriminating ADHD. In addition, it can be hypothesized that the Raven task maybe superior to traditional IQ tests, in distinguishing ADHD in children with high IQ and it can be a direction for further research.

The current study has some significant limitations that should be noted. Systematic observational methods or teacher ratings have not been included in this study, and the authors were unable to compare these important methods with other diagnostic measures. There are also a large number of EF tests that the authors failed to include in the battery, due to time constraints. Also, the authors cannot determine the degree to which observational methods might significantly increase the diagnostic power of the multi-method assessment. Therefore, more research remains to be done to properly answer these questions.

Conclusion

Our findings further support the importance of multi-method, multi-informant assessments for the diagnosis of ADHD in clinically referred children. The variability seen in the discriminative power of specific tests, when used in isolation, suggests a heightened risk for over-identification, under-identification, or misdiagnosis when relying on a single source to make diagnostic decisions. The prominent Kappa statistic of .825 between diagnosticians using our multi-method assessment showed that the use of such a battery would enhance the agreement among clinicians. Our findings also can help to select appropriate and sensitive diagnostic measures for identifying ADHD symptoms and impairments.

Finally, the fact that the study was conducted in Iran with Persian-speaking families should be underscored, because it required some adaptations in traditional ADHD assessments, which were originally developed for use with English-speaking, Western European families. Our findings extend the inter-cultural generalizability of the traditional English ADHD assessments, as well as lay the groundwork for subsequent adaptation and development of multi-method assessments for ADHD, as well as other disorders, in non-English speaking countries.

Author Note:

The authors thank all child participants and their parents who patiently participated in our study. The authors also appreciate the cooperation of the staff of the counseling and guidance center of the department of education of the Shahriar district of Karaj, Iran.

Statements:

There is no conflict of interest. This study was approved by the scientific and ethical committee of the department of psychology of Allameh

Tabatabaei University. All the child participants and their parents read and approved the informed consent forms.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychol Bull*, *101*(2), 213-232.

Alda, J. A., & Serrano-Troncoso, E. (2013). Attention-deficit hyperactivity disorder: agreement between clinical impression and the SNAP-IV screening tool. *Actas espanolas de psiquiatria*, *41*(2), 76-83.

Antshel, K. M., Faraone, S. V., Stallone, K., Nave, A., Kaufmann, F. A., Doyle, A., Fried, R., Seidman, L., & Biederman, J. (2007). Is attention deficit hyperactivity disorder a valid diagnosis in the presence of high IQ? Results from the MGH Longitudinal Family Studies of ADHD. *Journal of child psychology and psychiatry*, *48*(7), 687-694. https://doi.org/10.1111/j.1469-7610.2007.01735.x

APA. (1968). *Diagnostic and statistical manual of mental disorders (2ne ed.)* Author.

Babikian, T., Boone, K. B., Lu, P., & Arnold, G. (2006). Sensitivity and specificity of various digit span scores in the detection of suspect effort. *The Clinical Neuropsychologist*, *20*(1), 145-159. https://doi.org/10.1080/13854040590947362

Barkley, R. A. (2014). *Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment*. Guilford Publications.

Barkley, R. A. (2015). *ADHD in Children: Diagnosis and Assessment* http://www.continuingedcourses.net/active/courses/course004.php

Beck, L. H., Bransome, E. D., Jr., Mirsky, A. F., Rosvold, H. E., & Sarason, I. (1956). A continuous performance test of brain damage. *J Consult Psychol*, *20*(5), 343-350.

Bloom, B., Cohen, R. A., & Freeman, G. (2012). Summary health statistics for U.S. children: National Health Interview Survey, 2011. National Center for Health Statistics. *Vital Health Stat*, *10*(254).

Brock, S., & Clinton, A. (2007). Diagnosis of Attention-Deficit/Hyperactivity Disorder (AD/HD) in Childhood: A Review of the Literature. *The California School Psychologist*, *12*(1), 73-91. https://doi.org/10.1007/BF03340933

Bruchmuller, K., Margraf, J., & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *J Consult Clin Psychol*, *80*(1), 128-138. https://doi.org/10.1037/a0026582

Bussing, R., Fernandez, M., Harwood, M., Wei, H., Garvan, C. W., Eyberg, S. M., & Swanson, J. M. (2008). Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: psychometric properties and normative ratings from a school district sample. *Assessment*, *15*(3), 317-328. https://doi.org/10.1177/1073191107313888

Centers for Disease Control and Prevention, C. D. C. (2013). *Attention-Deficit/ Hyperactivity Disorder (ADHD): Symptoms and Diagnosis* http://www.cdc.gov/ncbddd/adhd/diagnosis.html

Checkmateplus. (2015). *Child Symptom Inventory* http://www.checkmateplus.com/product/csi-4.htm

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Edition)*. Routledge.

Collett, B. R., Ohan, J. L., & Myers, K. M. (2003). Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*, *42*(9), 1015-1037. https://doi.org/10.1097/01.chi.0000070245.24125.b6

Coners, C., & MSH Staff. (2000). *Conners' Continuous Performance Test II: Computer program for Windows technical guide and software manual*. Multi-Health Systems.

Conners, C. K. (1997). *Conners' Rating Scales-Revised: Technical Manual*. Multi-Health Systems. https://books.google.com/books?id=8B2cGQAACAAJ

Corkum, P. V., & Siegel, L. S. (1993). Is the Continuous Performance Task a valuable research tool for use with children with Attention-Deficit-Hyperactivity Disorder? *Journal of child psychology and psychiatry*, *34*(7), 1217-1239.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological bulletin*, *131*(4), 483-509. https://doi.org/10.1037/0033-2909.131.4.483

Diller, L. (2015). *Interview by PBS Frontline [online Archive]. Medicating Kids* http://www.pbs.org/wgbh/pages/frontline/shows/medicating/interviews/diller.html

Doyle, A. E., Biederman, J., Seidman, L. J., Weber, W., & Faraone, S. V. (2000). Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without attention deficit-hyperactivity disorder. *J Consult Clin Psychol*, *68*(3), 477-488.

Dunn, D. W., & Kronenberger, W. G. (2003). Attention-deficit/hyperactivity disorder in children and adolescents. *Neurologic clinics*, *21*(4), 933-940.

Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology*, *18*(3), 543-555. https://doi.org/10.1037/0894-4105.18.3.543

Gadow, K. D., & Sprafkin, J. N. (2002). *Child Symptom Inventory 4: Screening and Norms Manual*. Checkmate Plus. https://books.google.com/books?id=Zcy2tgAACAAJ

[Record #20 is using a reference type undefined in this output style.]

Grañana, N., Richaudeau, A., Robles Gorriti, C., O'Flaherty, M., Scotti, M. E., Sixto, L., Allegri, R., & Fejerman, N. (2011). Evaluación de déficit de atención con hiperactividad: la escala SNAP IV adaptada a la Argentina. *Revista Panamericana de Salud Pública*, *29*(5), 344-349.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218.

Greiner, M., Pfeiffer, D., & Smith, R. D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Prev Vet Med*, *45*(1-2), 23-41.

Hersen, M. (2011). *Clinician's handbook of child behavioral assessment*. Academic Press.

Iverson, G. L., & Tulsky, D. S. (2003). Detecting malingering on the WAIS-III. Unusual Digit Span performance patterns in the normal population and in clinical groups. *Archives of clinical neuropsychology*, *18*(1), 1-9.

Jensen, C. M., & Steinhausen, H. C. (2015). Comorbid mental disorders in children and adolescents with attention-deficit/hyperactivity disorder in a large nationwide study. *Atten Defic Hyperact Disord*, *7*(1), 27-38. https://doi.org/10.1007/s12402-014-0142-1

Johnston, C., & Murray, C. (2003). Incremental validity in the psychological assessment of children and adolescents. *Psychol Assess*, *15*(4), 496-507. https://doi.org/10.1037/1040-3590.15.4.496

Kaplan, E., Fein, D., Morris, R., & Delis, D. (1991). *WAIS-R as a neuropsychological instrument*. The Psychological Corporation.

Kazdin, A. E. (2005). Evidence-based assessment for children and adolescents: issues in measurement development and clinical application. *Journal of clinical child and adolescent psychology*, *34*(3), 548-558. https://doi.org/10.1207/s15374424jccp3403_10

Khodadadi, M., Mashhadi, A., & Amani, H. (2009a). *Persian Continuous Performance Test [Software]: Iran Registration No.: 11590*. Ravantajhiz e Sina.

Khodadadi, M., Mashhadi, A., & Amani, H. (2009b). *Persian Tower of London Test [Software]: Iran Registration No.: 11590*. Ravantajhiz e Sina.

Khodadadi, M., Mashhadi, A., & Amani, H. (2009c). *Persian Stroop Color-Word Test [software]: Iran Registration No.: 11590*. Ravantajhiz e Sina.

Khushabi, K., Pour-Etemad, H., Mohammadi, M., & Mohammad Khani, P. (2006). The prevalence of ADHD in primary school students in Tehran. *Medical Journal of the Islamic Republic of Iran*, *20*(3), 147-150.

Klorman, R., Hazel-Fernandez, L. A., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J. M., Stuebing, K. K., & Shaywitz, B. A. (1999). Executive functioning deficits in attention-deficit/hyperactivity disorder are independent of oppositional defiant or reading disorder. *J Am Acad Child Adolesc Psychiatry*, *38*(9), 1148-1155. https://doi.org/10.1097/00004583-199909000-00020

Kollins, S. H. (2007). Abuse liability of medications used to treat attention-deficit/hyperactivity disorder (ADHD). *Am J Addict*, *16 Suppl 1*, 35-42; quiz 43-34. https://doi.org/10.1080/10550490601082775

Krikorian, R., Bartok, J., & Gay, N. (1994). Tower of London procedure: a standard method and developmental data. *J Clin Exp Neuropsychol*, *16*(6), 840-850. https://doi.org/10.1080/01688639408402697

Lumos-Labs. (2012). *Brain Games & Brain Training-Lumosity* [Internet based cognitive training program]. San Francisco

Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pelletier, S. F., Rawlins, C., Singer, H. S., & Denckla, M. B. (2002). Effects of IQ on executive function measures in children with ADHD. *Child neuropsychology*, *8*(1), 52-65. https://doi.org/10.1076/chin.8.1.52.8719

Mash, E. J., & Barkley, R. A. (2007). *Assessment of childhood disorders*. Guilford Press.

Mashhadi, A., Rasoolzadeh-Tabatabaei, S. K., Azad-Fallah, P., & Soltanifar, A. (2009). [مقایسه بازداری پاسخ و کنترل تداخل در کودکان مبتلا به نارسایی توجه فزون کنشی و کودکان بهنجار] Comparison of response inhibition and interference control in children with ADHD and normal children. مجله روان شناسی بالینی, ۱(۲), ۳۷-۵۰.

Mashhadi, A., Rasoul-Zadeh Tabatabie, K., Azad-Fallah, P., & Soltanifar, A. (2010). [توانایی برنامه ریزی و سازمان دهی در کودکان مبتلا به نارسایی توجه/ فزون کنشی] The ability of planning and organization in children with attention deficit/hyperactivity disorder. مطالعات تربیتی و روانشناسی, ۱۱(۱), ۱۵۱-۱۷۰. http://www.sid.ir/fa/VEWSSID/J_pdf/71013890108.pdf

Mathias, C. W., Greve, K. W., Bianchini, K. J., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment*, *9*(3), 301-308.

Mayes, R., Bagwell, C., & Erkulwater, J. (2008). ADHD and the rise in stimulant use among children. *Harv Rev Psychiatry*, *16*(3), 151-166. https://doi.org/10.1080/10673220802167782

McConaughy, S. H., Harder, V. S., Antshel, K. M., Gordon, M., Eiraldi, R., & Dumenci, L. (2010). Incremental Validity of Test Session and Classroom Observations in a Multimethod Assessment of Attention Deficit/Hyperactivity Disorder. *Journal of clinical child and adolescent psychology*, *39*(5), 650-666. https://doi.org/10.1080/15374416.2010.501287

McGee, R. A., Clark, S. E., & Symons, D. K. (2000). Does the Conners' Continuous Performance Test aid in ADHD diagnosis? *J Abnorm Child Psychol*, *28*(5), 415-424.

Mohammad-Esmail, E. (2007). [انطباق و هنجاریابی سیاهه نشانه های مرضی کودک ویرایش چهارم] Adaptation and Standardization of child symptom inventory-4 (CSI-4). پژوهش در حیطه کودکان استثنایی, ۷(۱), ۷۹-۹٦.

National Collaborating Centre for Mental Health. (2009). National Institute for Health and Clinical Excellence: Guidance. In *Attention Deficit Hyperactivity Disorder: Diagnosis and Management of ADHD in Children, Young People and Adults*. British Psychological Society (UK) The British Psychological Society & The Royal College of Psychiatrists.

Nazifi, M., Rasoolzadeh-Tabatabaei, K., Azad-Fallah, P., & Moradi, A. (2011). [توجه پایدار و بازداری پاسخ در کودکان مبتلا به نارسایی توجه/ فزون کنشی در مقایسه با کودکان عادی] Sustained Attention and response inhibition in children with attention deficit/ hyperactivity disorder compared to normal children. مجله روان شناسی بالینی, ۳(۲), ۵۵-٦٤.

Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD subtypes. *J Am Acad Child Adolesc Psychiatry*, *41*(1), 59-66. https://doi.org/10.1097/00004583-200201000-00012

Pelham, W. E., Jr., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of clinical child and adolescent psychology*, *34*(3), 449-476. https://doi.org/10.1207/s15374424jccp3403_5

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of child psychology and psychiatry*, *37*(1), 51-87.

Riccio, C. A., Reynolds, C. R., & Lowe, P. A. (2001). *Clinical applications of continuous performance tests: Measuring attention and impulsive responding in children and adults*. John Wiley & Sons Inc.

Rom, D. M. (1990). A Sequentially Rejective Test Procedure Based on a Modified Bonferroni Inequality. *Biometrika*, *77*(3), 663-665. https://doi.org/10.2307/2337008

Sadrossadat, S., Houshyari, Z., Zamani, R., & Sadrossadat, L. (2007). [تعیین مشخصات روانسنجی مقیاس اسنپ ٤: اجرای والدین] Investigation of psychometric properties of SNAP-IV rating scale: Parents' ratings. توانبخشی, ۸(۳), ٥٩-٦٥.

Sciutto, M. J., & Eisenberg, M. (2007). Evaluating the evidence for and against the overdiagnosis of ADHD. *Journal of Attention Disorders*, *11*(2), 106-113.

Shahim, S. (1991). [هنجاریابی مقیاس هوشی وکسلر کودکان در شیراز] Norm finding for the Wechsler's Intelligence Scales for Children in Shiraz. مجله علوم اجتماعی و انسانی دانشگاه شیراز, ۷(۱), ۱۲۳-۱٥٤. http://www.ensani.ir/storage/Files/20110225131530-%D9%87%D9%86%D8%AC%D8%A7%D8%B1%D9%8A%D8%A7%D8%A8%D9%8A%20%D9%85%D9%82%D9%8A%D8%A7%D8%B3.pdf

Sleator, E. K., & Ullmann, R. K. (1981). Can the physician diagnose hyperactivity in the office? *Pediatrics*, *67*(1), 13-17.

Statistical Center of Iran. (2006). *[سرشماري عمومي نفوس و مسکن ۱۳۸۵ بدون احتساب خانوارهاي موسسه اي و غيرساکن] National Population and Housing Census 1385 (2006) (without taking into account institutionalized and non-resident families)* http://wayback.archive.org/web/20111114233307/http://www.amar.org.ir/DesktopModules/FTPManager/upload/upload2360/newjkh/newjkh/23.xls

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., Clevenger, W., Davies, M., Elliott, G. R., Greenhill, L. L., Hechtman, L., Hoza, B., Jensen, P. S., March, J. S., Newcorn, J. H., Owens, E. B., Pelham, W. E., Schiller, E., Severe, J. B., Simpson, S., Vitiello, B., Wells, K., Wigal, T., & Wu, M. (2001). Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *J Am Acad Child Adolesc Psychiatry*, *40*(2), 168-179. https://doi.org/10.1097/00004583-200102000-00011

Vitiello, B. (2001). Psychopharmacology for young children: clinical needs and research opportunities. *Pediatrics*, *108*(4), 983-989.

Wechsler, D. (1974). *Manual for the Wechsler intelligence scale for children, revised*. Psychological Corporation.

Wei, K., Yan, X., Kong, G., Yin, C., Zhang, F., Wang, Q., & Kording, K. P. (2014). Computer use changes generalization of movement learning. *Curr Biol*, *24*(1), 82-85. https://doi.org/10.1016/j.cub.2013.11.012

Westerberg, H., Hirvikoski, T., Forssberg, H., & Klingberg, T. (2004). Visuo-spatial working memory span: a sensitive measure of cognitive deficits in children with ADHD. *Child neuropsychology*, *10*(3), 155-161. https://doi.org/10.1080/09297040409609806

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological psychiatry*, *57*(11), 1336-1346. https://doi.org/10.1016/j.biopsych.2005.02.006

Wolraich, M. L., Lambert, W., Doffing, M. A., Bickman, L., Simmons, T., & Worley, K. (2003). Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *J Pediatr Psychol*, *28*(8), 559-567.

Zou, K. H., O'Malley, A. J., & Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, *115*(5), 654-657. https://doi.org/10.1161/circulationaha.105.594929

Received	January 4, 2022
Revision received	March 13, 2022
Accepted	April 1, 2022