Development and Aging

Incremental clinical utility of continuous performance tests in childhood ADHD – an evidence-based assessment approach

PIA TALLBERG,[1] MARIA RÅSTAM,[1] LENA WENHOV,[2] GLEN ELIASSON[3] and PEIK GUSTAFSSON[1]

[1]*Department of Clinical Sciences Lund, Child and Adolescent Psychiatry, Skane University Hospital, Lund University, Lund, Sweden*
[2]*Skåne University Hospital, Lund, Sweden*
[3]*Malmö Municipality, Sweden*

Tallberg, P., Råstam, M., Wenhov, L., Eliasson, G. & Gustafsson, P. (2019). Incremental clinical utility of continuous performance tests in childhood ADHD – an evidence-based assessment approach. *Scandinavian Journal of Psychology, 60,* 26–35.

Despite extensive research on attention deficit hyperactivity disorder (ADHD), there are still uncertainties regarding the clinical utility of different ADHD assessment methods. This study aimed to examine the incremental clinical utility of Conners' continuous performance test (CPT) II and QbTest in diagnostic assessments and treatment monitoring of attention deficit hyperactivity disorder (ADHD). Retrospective data from child and adolescent psychiatric records of two populations were studied. The diagnostic clinical utility of Conners' CPT II and QbTest was analysed using receiver operator characteristics (ROC) and post-test probability in 80 children with and 38 without ADHD. Dose titrations of central stimulants in 56 children with ADHD were evaluated using QbTest and the Swanson, Nolan, Pelham, version IV (SNAP-IV) scale. Conners' CPT II, but not QbTest, had incremental clinical utility in diagnostic assessment of children with ADHD when teacher and parent ratings were inconclusive. QbTest proved useful in titration of central stimulant treatment when parent ratings were inconclusive. Continuous performance tests were found to be clinically useful when rating scales were inconclusive.

Key words: ADHD, child psychiatry, evidence-based, assessment, CPT, central stimulants.

Pia Tallberg, *Department of Clinical Sciences Lund, Child and Adolescent Psychiatry, Faculty of Medicine, Lund University*, *Baravägen 1C*, *Box 188*, *SE-221 85 Lund*, *Sweden*. E-mail: pia.tallberg@med.lu.se

INTRODUCTION

Despite extensive research on attention deficit hyperactivity disorder (ADHD), and despite existing guidelines on the condition, there are still uncertainties regarding the clinical utility of different ADHD assessment methods and how a clinician should incorporate multiple informants, multiple methods, co-occurring symptoms and functional impairment in the diagnostic process (Jarrett, Van Meter, Youngstrom, Hilton & Ollendick, 2016). "Clinical utility" means that the assessment methods provide valuable information to the practitioners, improving the clinical work, and that they are cost-effective and give specific information not provided by any other method (incremental validity) (Hunsley & Mash, 2007; Mash & Hunsley, 2005; Pelham, William, Fabiano & Massetti, 2005). A systematic review of international guidelines on ADHD, by Seixas, Weiss and Muller (2012), revealed that all guidelines recommended that the diagnostic assessments of ADHD in children should include a family interview pertaining to the developmental history of the child, current symptoms, behaviors and everyday functioning. Adding some standardized validated behavior rating scales improves the accuracy of the diagnosis (Seixas *et al*., 2012) and has been proposed to be the most cost-effective approach to performing ADHD diagnostic assessments (Pelham *et al*., 2005). Generally in child psychiatry, the agreement between different informants varies, and is often only low to moderate (De Los Reyes *et al*., 2011; Youngstrom, Pabon, Youngstrom, Feeny & Findling, 2011; Martel, Markon & Smith, 2017). Consistent information from different informants suggests that symptoms are pervasive, while inconsistent information may reflect different behaviors in the child in different contexts (De Los Reyes, Augenstein, Wang *et al*., 2015; Martel *et al*., 2017). More research on the construct validity of multi-informant assessments is needed (De Los Reyes *et al*., 2015).

According to the Seixas *et al*. (2012) review, most guidelines neither recommend nor advise against neuropsychological testing for diagnosis of ADHD (Seixas *et al*., 2012). Many neuropsychological tests have been used in an effort to measure the core symptoms of ADHD (Willcutt, Doyle, Nigg, Faraone & Pennington, 2005), but none is considered reliable and sufficiently valid to be used as a single measure for diagnosing ADHD (Hall, Valentine, Groom *et al*., 2016; Willcutt *et al*., 2005).

Continuous performance tests in diagnostic assessments

When it comes to differentiating between children with and children without ADHD, the strongest and most consistent results have been obtained by using continuous performance tests (CPTs) (Frazier, Demaree & Youngstrom, 2004; Huang-Pollock, Karalunas, Tam & Moore, 2012; Losier, McGrath & Klein, 1996). The CPTs are a group of computerized tests that measure attentiveness, impulsivity and vigilance through analysis of hits, omissions, commissions, reaction time (RT) and reaction time variability (RTV) (Schatz, Ballantyne & Trauner, 2001). A meta-analytic review of 319 studies, by Kofler, Rapport, Sarver and others (2013), indicates that RTV is the parameter that has the largest effect size with regard to how children and adolescents with ADHD differ from non-clinical groups (Kofler *et al*., 2013). However, the review also shows that not everyone with ADHD has deficits in executive functions including RTV. In addition,

individuals with other neurodevelopmental diagnoses, such as dyslexia and acquired brain injury, seem to have similar difficulties with executive functions and RTV. Consequently, the discriminant validity of CPTs has been questioned, particularly with regard to diagnostic utility (Lipszyc & Schachar, 2010; Nichols & Waschbusch, 2004; Preston, Fennell & Bussing, 2005; Riccio, Reynolds, Lowe & Moore, 2002; Schatz *et al*., 2001). Hall *et al*. (2016) have reviewed the current evidence base for the use of CPTs in the diagnostic assessments and medication management of children with ADHD. They suggest that CPTs can objectively assess attention and impulsivity but that the results are inconclusive for diagnostic decision making and medication monitoring. They recommend further studies, especially randomized controlled trials (Hall *et al*., 2016). Other objections to CPTs are that different CPTs differ in terms of construction and normative data samples (Huang-Pollock *et al*., 2012) and that they are time-consuming and expensive.

In 2016, Jarrett and colleagues examined diagnostic efficiency using an evidence-based assessment approach, calculating diagnostic likelihood ratios with different threshold values for the Child Behavior Checklist (CBCL) and Conners' CPT (Jarrett *et al*., 2016). The parameter hit reaction time standard error (HRT-SE) showed diagnostic utility and was recommended as supplemental diagnostic assessment in unclear cases of ADHD (Jarrett *et al*., 2016). "Evidence-based assessment" refers to the use of research and theory to guide clinicians in how to assess different conditions (Hunsley & Mash, 2007). It provides a systematic way of integrating different sources of information (from multiple informants and multiple instruments) using Bayesian statistics (Frazier, Youngstrom & Hamilton, 2006; Hunsley & Mash, 2007; Jarrett *et al*., 2016; Mash & Hunsley, 2005). Bayesian methods provide a statistical and conceptual framework for taking research data and translating them into answers to practical clinical assessment questions by calculating pretest (or prior) and post-test probability (Akobeng, 2007; Ashby, 2006; Jarrett *et al*., 2016; Youngstrom, 2013). "The base rate provides an estimate of the prior probability of a diagnosis (in other words, a 'best guess' before gathering additional assessment data), and then combine it with the change in risk attached to a particular assessment finding, estimating the updated posterior probability" (Youngstrom, Choukas-Bradley, Calhoun & Jensen-Doss, 2015, p. 21). It is possible to use this procedure repeatedly when performing several tests. Then, the post-test probability of test number one is used as pretest probability of test number two (Akobeng, 2007; Ashby, 2006; Jarrett *et al*., 2016; Youngstrom, 2013). Another way of examining the clinical utility of different assessment methods is to analyze the Receiver operator characteristics (ROC) curve. The ROC curve is a graphical plot that illustrates the diagnostic ability of a binary outcome as its discrimination threshold is varied (Youngstrom, 2014).

Continuous performance tests in treatment evaluation

Another important assessment is the evaluation of the medical treatment of ADHD. In children this is usually based on subjective descriptions by parents, teachers and professionals. However, symptoms and the behavior of the child can be influenced by factors that have nothing to do with the effect of the drug itself, such as changes in school demands, teacher collaboration and family factors (Fernández-Jaén, Fernández-Mayoralas, Pardos, Calleja-Pérez & Jareño, 2009). In the titration of central stimulants, different doses are tested and evaluated based on reports from parents and teachers, often using behavior rating scales such as SNAP-IV (Bussing, Fernandez, Harwood *et al*., 2008). Some studies have shown that central stimulants improve the performance in CPTs, which may therefore be used in dose titration (Fernández-Jaén *et al*., 2009; Losier *et al*., 1996; Silberstein, Pipingas, Farrow, Levy & Stough, 2016; Spencer, Hawk, Richards, Shiels, Pelham & Waxmonsky, 2009). In a recent study on changes in attention before and after treatment with central stimulants (Ramtvedt & Sundet, 2014), both the teachers' assessment and the results of the CPT showed improvements after medication. However, these two variables do not always correlate. In one study, children who showed improvements in CPTs had generally improved clinically (Fernández-Jaén *et al*., 2009), while in another, parents who perceived their children's inattentive symptoms as mild were more likely to deny any subjective improvement from medication, although the children had improved their performance on CPTs (Park, Kim, Cho *et al*., 2013). Further, whereas certain doses of central stimulants may be optimal for improving focused and sustained attention, other doses may be optimal for improved inhibition and set-shifting (Konrad, Günther, Hanisch & Herpertz-Dahlmann, 2004). The question is whether treatment of children with ADHD could be optimized by a titration procedure based on behavior rating scales or CPTs or both (Konrad *et al*., 2004).

Aims

The first aim of the present study was to investigate the incremental clinical diagnostic utility of the Conners' CPT II and the QbTest (QbTech Ltd, www.qbtech.com) using Bayesian statistics and ROC analysis. The second aim was to evaluate the clinical utility of using the QbTest as a supplement to SNAP-IV in dose titration of stimulant medication.

We wanted to test the following hypotheses:

(1) The Conners' CPT II and the QbTest have incremental clinical value in the diagnostic assessment of ADHD.
(2) The QbTest may be useful in dose titration to find the optimal medication for the patient.

METHODS

Study group 1 (diagnostic)

To investigate the clinical diagnostic utility of CPT, clinical retrospective data from ADHD assessments were collected from Child and Adolescent Psychiatry (CAP) clinical records. During the period 1 November 2009 to 31 December 2010, a total of 118 children who screened positive for ADHD were referred for further assessments in a CAP clinic in the south of Sweden. In southern Sweden, parents with concerns about their child's mental

wellbeing are advised to contact a specific intake unit where specially trained nurses will interview the parent(s) using the Brief Child and Family Phone Interview (BCFPI) to screen for child psychiatric symptoms.

All children were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) (American Psychiatric Association, 1994). Functional level was determined during the family and teacher interview, using the BCFPI and SNAP-IV rating scale. The assessments were performed by a multidisciplinary team and consensus diagnoses were assigned by the team. The assessments consisted of a child psychiatric examination including neurological status, conducted by an experienced child psychiatrist, a clinical semi-structured interview with the parents and teachers, an interview with the child, and a neuropsychological assessment of the child including cognitive level and CPTs. The age of onset was established from the child's developmental history. In the study, children who were diagnosed with ADHD (the ADHD group, n = 80) were compared with children who were deemed not to have ADHD (the non-ADHD group, n = 38) (see flow chart, Fig. 1. For participants' background information, screening results of the BCFPI and results from neuropsychological assessments, see Table 1). In the non-ADHD group, altogether 24 children were diagnosed with autism spectrum disorders (n = 5), tic disorders (n = 3), language impairments or learning disorders (n = 12), and internalized problems such as mood disorder or anxiety disorder (n = 12). Fourteen children did not fulfill any diagnostic criteria. There were no significant differences between the ADHD group and the non-ADHD group regarding age, gender and intellectual ability. (Wechsler Intelligence Scale for Children, Intelligence Quotient (IQ) for the ADHD group, mean 87.15 (confidence interval (CI) 74.58–99.72) and the non-ADHD group, mean 91.86 (CI 78.59–105.13). Two cases had full scale IQ just below 70, but with uneven cognitive profiles).

Study group 2 (medication)

To examine the clinical utility of the QbTest in medical titration, a group of 186 ADHD patients from another CAP clinic in the south of Sweden were selected. The clinical ADHD diagnoses according to DSM-IV criteria were based on consensus team meetings with experienced CAP clinicians. The assessments were performed by a multidisciplinary team and were concluded with a consensus team discussion regarding the diagnosis, in the same way as described above for the clinical sample used for the evaluation of diagnostic assessments. A total of 186 children with an ADHD diagnosis were assessed with the SNAP-IV and the QbTest by the team nurses between January 2007 and June 2011 before starting treatment with methylphenidate (see flow chart, Fig. 2). They were all followed up and evaluated 1 year later. In 130 children the data were incomplete, leaving 56 patients (45 boys and eleven girls, aged 7.1–17.8 years, mean age 12.29, standard deviation (SD) 2.45 years) for evaluation and follow-up. Titration of the optimal dose of methylphenidate in the treatment of ADHD was performed following the general principles for dose titration described in the National Institute for Health and Care Excellence (NICE) guidelines: https://www.nice.org.uk/guidance/CG72/chapter/Recommendations#treatment-for-children-and-young-people.

Dosage titration started at a low dose, of 18 mg or 20 mg, and the dose was titrated in steps of 10 mg for Ritalin, Equasym and Medikinet and 18 mg for Concerta up to a maximal dose of 60 mg, or less in case of marked side effects. At each dose, the parents and teachers filled in the SNAP-IV scale and the child was tested using the QbTest. An improvement in SNAP-IV symptom scores of about 0.4 SD was considered clinically significant in the MTA study (Swanson, Kraemer, Hinshaw *et al*., 2001). In the present medication study, a decrease in SNAP-IV scores of >0.2 (0.4 SD lies between the score of 0.2 and 0.3) was regarded as a clinically significant improvement. In the same way a decrease in QbTest scores of >0.4 SD was regarded as a clinically significant improvement. If none of the doses led to a clinically significant improvement the titration was regarded as negative. If one or more of the doses led to a clinically significant improvement the titration was regarded as positive and the dose with the best results (optimal dose) was used. A good outcome was defined as being on the optimal dose one year after titration.

Instruments

The *BCFPI* is a structured telephone interview used as a triage assessment method, administered to the parents of 3–18-year-old children and adolescents and standardized to T- scores (m = 50, SD = 10) based on Canadian community and clinical samples (Cunningham, Boyle, Hong, Pettingill & Bohaychuk, 2009). The BCFPI is a screening method with good validity and diagnostic accuracy (Boyle *et al*., 2009). The BCFPI includes two mental

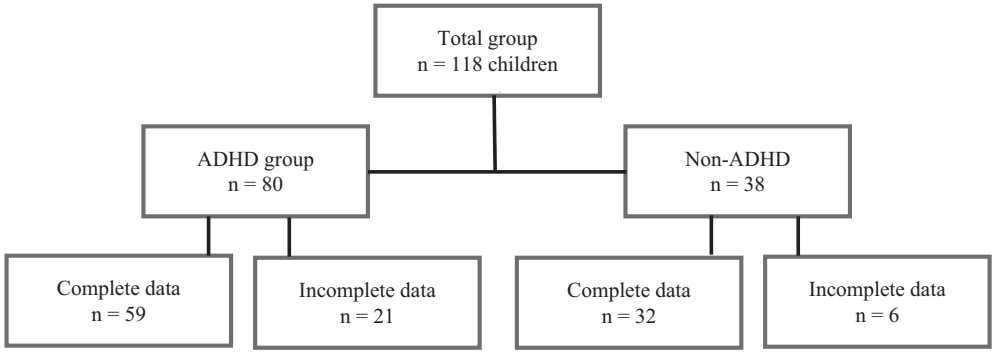

Fig. 1. Flow chart Study group 1 (diagnostic). ADHD = attention deficit hyperactivity disorder.

Table 1. *Participants' background information, screening results of the Brief Child and Family Phone Interview (BCFPI) and results of neuropsychological assessments in Study group 1 (diagnostic)*

	ADHD n = 80 (68%)	Non-ADHD n = 38 (32%)	
Gender: Boys, n (%)	57 (71)	24 (63)	
Age, yrs, median (1st–3rd quartiles)	12.5 (9.6–14.4)	11.2 (9.6–13.0)	
ADHD subtype			
ADHD-C, n (%)	56 (70)		
ADHD-I, n (%)	22 (28)		
ADHD-H, n (%)	2 (2)		
Brief Child and Family Phone Interview (BCFPI)			
Mental health subscales	Mean T-score[a] (SD)	Mean T-score (SD)	p-value
Regulating attention	73.66 (65.18–82.14)	70.21 (60.66–79.77)	ns
Regulating impulsivity & activity	66.41 (55.79–77.03)	58.76 (46.74–70.81)	0.002
Regulating attention, impulsivity & activity	71.78 (62.86–80.70)	65.62 (55.48–75.77)	0.003
Cooperativeness	63.57 (50.41–76.73)	60.52 (45.64–75.41)	ns
Conduct	64.14 (39.29–89.00)	60.29 (37.60–82.98)	ns
*Externalizing behaviour	70.28 (58.12–82.44)	65.04 (52.07–78.00)	ns
Separating from parents	52.81 (41.64–63.97)	54.12 (42.18–66.06)	ns
Managing anxiety	56.95 (43.89–70.00)	55.82 (43.22–68.42)	ns
Managing mood	55.75 (40.83–70.67)	61.47 (43.05–79.88)	ns
Self-harming	56.51 (41.77–71.25)	62.85 (44.37–81.33)	ns
*Internalizing behaviour	56.74 (43.95–69.52)	58.68 (45.48–71.88)	ns
*Total externalized & internalized problems	65.50 (53.85–77.16)	63.58 (51.77–75.39)	ns
Child functioning scales			
Social participation	60.70 (46.35–75.05)	61.12 (45.38–76.86)	ns
Quality of social relations	59.32 (44.62–74.01)	55.69 (40.73–70.65)	ns
School participation and achievement	61.21 (46.42–76.00)	59.28 (45.13–73.43)	ns
*Global functioning	62.27 (49.38–75.16)	60.37 (48.01–72.74)	ns
Family functioning scales			
Family activities	63.63 (36.32–90.94)	59.61 (36.40–82.83)	ns
Family comfort	67.96 (55.90–80.01)	63.62 (51.67–75.57)	ns
*Family impact	68.60 (51.24–85.96)	63.14 (49.13–77.16)	ns
Risk factors			
Informant mood	52.22 (40.64–63.81)	54.44 (39.07–69.80)	ns
Family functioning	47.02 (36.39–57.65)	53.86 (38.30–69.43)	0.017
SNAP–IV	Mean score (SD)	Mean score (SD)	p-value
Parent inattention	2.01 (1.89–2.14)	1.81 (1.54–2.08)	ns
Parent hyperactivity/impulsivity	1.54 (1.32–1.77)	1.13 (0.83–1.43)	0.035
Teacher inattention	2.02 (1.86–2.18)	1.60 (1.35–1.85)	0.005
Teacher hyperactivity/impulsivity	1.51 (1.29–1.73)	0.92 (0.63–1.21)	0.002
Conners' CPT II	Mean T-score (SD)	Mean T-score (SD)	p-value
Hit reaction time	52.80 (49.77–55.84)	47.25 (44.68–49.83)	0.008
Hit reaction time standard error	54.59 (52.02–57.16)	46.86 (44.54–49.17)	<0.001
Confidence Index (%)	57.61 (52.93–62.29)	45.23 (40.61–49.85)	0.001
QbTest	Mean Q-score[b] (SD)	Mean Q-score (SD)	p-value
Qb Activity	0.60 (0.26–0.93)	0.74 (0.20–1.28)	ns
Qb Impulsivity	0.96 (0.56–1.36)	0.29 (-0.16–0.75)	0.045
Qb Inattention	0.70 (0.43–0.98)	0.40 (-0.06–0.74)	ns

Notes: ADHD = attention deficit hyperactivity disorder; ADHD-C = ADHD combined type; ADHD-H = ADHD predominantly hyperactive-impulsive type; ADHD-I = ADHD predominantly inattentive type; CPT = continuous performance test; IQ = intelligence quotient; ns = non-significant; p = probability value; SD = standard deviation; SNAP-IV = Swanson, Nolan and Pelham, version IV, scale.

*Composite scale.

[a]Results are standardized into T-scores (m = 50; SD 10); [b]results are standardized into Q scores, i.e. a statistical model to transform skewed statistical distributions into normally distributed z-scores (m = 0; SD 1).

health composite scales of externalizing and internalizing behavior, one composite scale of children's global functioning and one composite scale of family functioning. The composite scale score of children's global functioning was used to estimate global functioning in the present study. The composite scales of mental health consist of subscales measuring behaviors associated

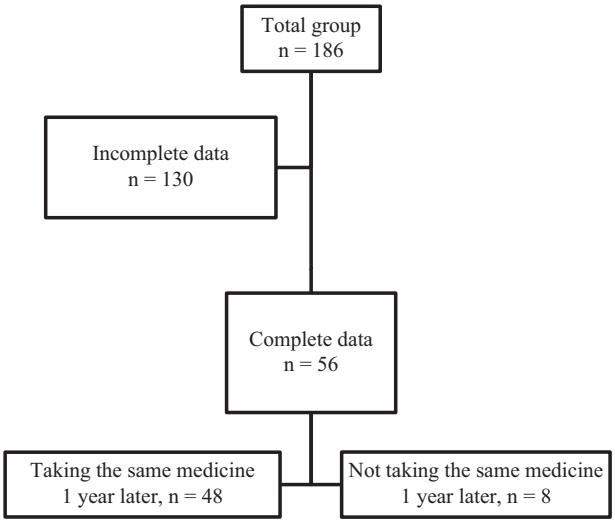

Fig. 2. Flow chart Study group 2 (medication).

with ADHD, oppositional defiant disorder (ODD), conduct disorder (CD), separation anxiety disorder (SAD), generalized anxiety disorder (GAD) and major depressive disorder (MDD) (Cunningham *et al*., 2009).

The *SNAP-IV* is a DSM-IV-based ADHD rating scale for parents and teachers (Bussing *et al*., 2008). There are different versions of the SNAP-IV with different numbers of questions. In assessments and in treatment evaluations for ADHD patients, the short version containing 30 questions is often used. The questions are based on the DSM-IV criteria for ADHD (18 questions) supplemented with eight questions concerning ODD symptoms and four supplementary questions regarding ODD (two questions) and ADHD (two questions). Average ratings are calculated for inattentive ADHD, hyperactive/impulsive ADHD, combined-type ADHD, and ODD subscales (Swanson *et al*., 2001). In the MTA study, scores above the 95th percentile are labeled as clinically relevant (Swanson *et al*., 2001; Bussing *et al*., 2008). The cut-off values were based on a normative sample of 2,744 children who participated in a larger, school-based mental health service program during 1993–1994, provided to low-income communities in several Texan cities. The ethnic composition was 76% Hispanic, 16% African American and 8% Caucasian (Gaub & Carlson, 1997). In a new evaluation, the psychometric properties of the SNAP-IV scale were found to be acceptable (Bussing *et al*., 2008; Posserud, Ullebo, Plessen, Stormark, Gillberg & Lundervold, 2014). Cronbach's alpha for overall parent ratings was 0.94 and for teacher ratings, 0.97. The interrater reliability for teachers and parents was moderate (0.49 for inattention and 0.43 for hyperactivity/impulsivity) (Bussing *et al*., 2008).

The *Conners' CPT II* is a computerized CPT used in individuals 6 years of age and older. The test has evolved from previous versions (Conners, Epstein, Angold & Klaric, 2003; Conners & MHS staff, 2000, 2004). The individual is instructed to press the spacebar whenever a letter appears on the screen (Go trials), except when the letter "X" appears (NoGo trials). The task includes a total of 360 letters presented consecutively on the screen within 250 milliseconds (ms). The trials are divided into six blocks and the blocks are randomly divided into three parts, one for each stimulus frequency, at interstimulus intervals (ISIs) of 1, 2 and 4 seconds. Ten per cent of the stimuli are NoGo trials ("X"). The measures in Conners' CPT II are divided into three subdomains: attention, impulsivity and vigilance, and have been standardized according to age and gender (Conners & MHS staff, 2000, 2004; Homack & Riccio, 2006). The Conners' CPT II manual (Conners & MHS staff, 2000) provides expanded normative data. Previous normative data included 1,190 participants and the database for the updated Conners' CPT II (Conners & MHS staff, 2000) consisted of 2,521 participants. Normative groups included 1,920 non-referred, healthy individuals and 378 diagnosed ADHD cases. Split-half reliability (hit RT (HRT) = 0.95; HRT-SE = 0.87) and test-retest correlation (HRT = 0.55; HRT-SE = 0.65; confidence index ADHD = 0.89) support the psychometric soundness of the Conners' CPT II (Conners & MHS staff, 2004). The confidence index indicates the degree of fit to the profile of clinical respondents. The higher the confidence index the greater the fit to the clinical profile. More extreme values (>60%) offer stronger evidence for clinical classification. Another way of classifying Conners' CPT is by examining the overall summary. If two or more T-scores are >60 (1 SD) this indicates a clinical attention problem (Conners & MHS staff, 2004).

The *QbTest* is a computerized CPT and activity test (Ulberstad, 2012). The QbTest has been approved by the US Food and Drug Administration (FDA; ref: K133382) to supplement other methods in evaluating the effect of stimulant medication for ADHD. Unlike other CPT tests, the QbTest measures hyperactivity by registering head movements during the test. The individual being tested is asked to press a key on the computer keyboard each time a target stimulus is shown, but not when a non-target figure is shown. The target stimulus for 6–12-year-old children is a grey circle without a cross and the non-target stimulus is a circle with a cross on it. The stimuli for individuals 13–55 years of age are four different figures, a red circle, a blue circle, a red square and a blue square. The target stimulus is a figure that is exactly the same (color and shape) as the previous one.

Head movements are registered by an infrared sender and camera, sending infrared light to a reflector placed on the forehead of the individual being tested and then receiving and registering the reflected light. For individuals 6–12 years of age, the stimulus is shown on the screen for 100 ms; for 13–55-year-olds, the stimulus is shown for 200 ms. The inter-stimulus interval is 2 seconds. Head movements are calculated by measuring the X and Y coordinate of the reflected light 50 times per second. Parameters used in the present study are Qb Activity, Qb Impulsivity and Qb Inattention. These results are expressed as standard deviations from normal values (Brocki, Tillman & Bohlin, 2010; Ulberstad, 2012) and are calculated from the first-line parameters, namely, omission errors, RT, RTV, commission errors and normalized commission errors for Qb Inattention. Qb Impulsivity also includes anticipatory responses. Qb Activity is calculated from time active, distance, area and micro events. The test-retest reliability has been found to be high, with r = 0.87 for Qb Inattention and r = 0.88 for Qb Activity (Ramtvedt & Sundet, 2014).

Statistical analysis

To investigate the clinical diagnostic utility of the Conners' CPT II and QbTest in Study group 1 (diagnostic), we first analyzed the

area under the ROC curve (AUC) to examine the diagnostic accuracy of the overall assessment measures. The outcome measure was the clinical diagnosis of ADHD. Tested variables were continuous values of the confidence index in the Conners' CPT II and Qb Activity, Qb Impulsivity and Qb Inattention in the QbTest. The significant level was set to $p = 5\%$ in all analyses. Only assessment methods with a statistically significant AUC were further analyzed concerning clinical utility. Clinical utility was defined by the degree of increment or decrease in post-test probability by stepwise adding different assessment methods. The post-test probabilities were calculated according to Bayes' theorem, which can be formulated as follows:

Probability (of a true outcome when the test is positive)
= probability (of the test being positive when having a true outcome)
× probability (of having a true outcome)/probability
(of the test being positive)

The judgement of clinical utility is a cost benefit decision with no fixed limits. It may be sufficient for some clinicians to start treatment when there is 80% chance of ADHD, while others would like to do more assessments. When patients want a high degree of certainty before accepting an ADHD diagnosis, 90% chance of having ADHD may be an appropriate limit (Youngstrom *et al*., 2015). In the present study, clinical utility was defined as a post-test probability of 0.85 (85%), since 85% chance of ADHD is reasonably sufficient for most clinicians to start treatment. Eight possible combinations of test results from dichotomized values of Conners' CPT II confidence index (cut-off = 50) and parent and teacher SNAP-IV scores for combined ADHD (MTA cut-off, for parent ratings cut-off = 1.67 and for teacher ratings cut-off = 2) were analyzed (see Table 2). Every test result could either be ≥ cut-off or it could be < cut-off. Each combination consisted of three steps. In step one, the post-test probability in the parents' assessment using SNAP-IV was calculated from the base rate of ADHD in the clinical setting, by using the positive predictive value (PPV) as post-test probability if the test result was ≥ cut-off and 1-negative predictive value (NPV) (the probability of having a diagnosis when the test is negative) if the rating was < cut-off. In step two, the post-test probability in step one was used as pre-test probability in the calculation of a new post-test probability (PPV for positive and 1-NPV for negative results) in the teachers' SNAP-IV assessment. In step three, this was repeated for the Conners' CPT II to give the final post-test probability. Spearman correlations between the assessment methods were calculated, since a prerequisite for this kind of serial computation is that the correlations are low.

To examine the clinical utility of QbTest in medical titration (Study group 2 (medication)), we performed Spearman correlations between cases with optimal dose and cases with good outcome. A good outcome was defined as being on the optimal dose one year after titration. An optimal dose was defined as a decrease in SNAP-IV symptom scores or QbTest scores of >0.4 SD. Qb Activity, Qb Impulsivity and Qb Inattention were analyzed in the correlations. Qb parameters with a statistically significant correlation were further analyzed. Univariable logistic regression analyses were performed for variables with a *p*-value

Table 2. *Post-test probabilities for eight possible clinical combinations, Study group 1 (diagnostic)*

	Pre-test probability	1st test: parent SNAP-IV combined ADHD Cut-off=1.67	2nd test: teacher SNAP-IV combined ADHD Cut-off=2	3rd test: Conners' CPT II Confidence Index Cut-off=50	ADHD n = 59; non-ADHD n = 32
Situation 1	0.68	≥Cut-off	≥Cut-off	<Cut-off	ADHD n = 7
Post-test probability		0.76	0.89	0.83	non-ADHD n = 2
Situation 2	0.68	≥Cut-off	<Cut-off	≥Cut-off	ADHD n = 9
Post-test probability		0.76	0.70	0.89	non-ADHD n = 1
Situation 3	0.68	<Cut-off	≥Cut-off	≥Cut off	ADHD n = 3
Post-test probability		0.53	0.74	0.91	non-ADHD n = 0
Situation 4	0.68	≥Cut-off	<Cut-off	<Cut-off	ADHD n = 10
Post-test probability		0.76	0.70	0.60	non-ADHD n = 8
Situation 5	0.68	<Cut-off	≥Cut-off	<Cut-off	ADHD n = 4
Post-test probability		0.53	0.74	0.65	non-ADHD n = 3
Situation 6	0.68	<Cut-off	<Cut-off	≥Cut-off	ADHD n = 5
Post-test probability		0.53	0.45	0.74	non-ADHD n = 3
Situation 7	0.68	<Cut-off	<Cut-off	<Cut-off	ADHD n = 12
Post-test probability		0.53	0.45	0.35	non-ADHD n = 15
Situation 8	0.68	≥Cut-off	≥Cut-off	≥Cut-off	ADHD n = 9
Post-test probability		0.76	0.89	0.97	non-ADHD n = 0

Notes: Eight possible combinations of test results from dichotomized values of Conners' CPT II confidence index (cut-off = 50) and parent and teacher SNAP-IV scores for combined ADHD (MTA cut-off, for parent ratings cut-off = 1.67 and for teacher ratings cut-off = 2) were analysed. Every test result could either be ≥ cut-off or it could be < cut-off. Each combination consisted of three steps. In step one, the post-test probability in the parents' assessment using SNAP-IV was calculated from the base rate of ADHD in the clinical setting, by using the positive predictive value (PPV) as post-test probability if the test result was ≥ cut-off and 1-negative predictive value (NPV) (the probability of having a diagnosis when the test is negative) if the rating was < cut-off. In step two, the post-test probability in step one was used as pre-test probability in the calculation of a new post-test probability (PPV for positive and 1-NPV for negative results) in the teachers' SNAP-IV assessment. In step three, this was repeated for the Conners' CPT II to give the final post-test probability.

ADHD = attention deficit hyperactivity disorder; CPT = continuous performance test; SNAP-IV = Swanson, Nolan and Pelham, version IV, scale.

of <0.20 in the correlation analyses. Sensitivity and specificity were calculated for the SNAP-IV parameters, the different QbTest scores and their stepwise combinations. In the stepwise analyses SNAP-IV ratings by parents were analyzed first and if the results were inconclusive (i.e., no optimal dose could be identified), Qb Inattention was analyzed. If results still were inconclusive an analysis of Qb Activity was made. Qb Impulsivity was omitted from the analyses because this variable had no significant correlation with treatment results one year later.

Ethical considerations

The study protocol followed the tenets of the Declaration of Helsinki and was approved by the Research Ethics Committee at Lund University, Lund, Sweden (Reg. No. 2012/88).

RESULTS

Analysis to achieve the first aim, Study group 1 (diagnostic)

The ROC analysis yielded a statistically significant AUC for Conners' CPT II confidence index (AUC = 0.73; $p < 0.001$). QbTest measures (Attention, Activity, Impulsivity) were not statistically significant. Only statistically significant measures were further analyzed. The Spearman correlations were 0.35 ($p = 0.001$) between continuous values of the parent and teacher SNAP-IV scores for combined ADHD, 0.17 ($p = 0.11$) between parent SNAP-IV scores for combined ADHD and Conners' CPT II confidence index, and 0.15 ($p = 0.15$) between teacher SNAP-IV scores for combined ADHD and Conners' CPT II confidence index. All three measures were judged to be only weakly associated, which made it possible to perform post-test probability analyses. Post-test probabilities were calculated for the eight possible combinations of the outcome measures (see Table 2).

Analysis to achieve the second aim, Study group 2 (medication)

When Spearman correlations were analyzed, only Qb Inattention had a significant correlation with finding the optimal dose one year later (rho = 0.289, $p = 0.013$). Qb Impulsivity gave very insignificant results ($p > 0.20$) in predicting the treatment response one year later and was omitted from further analyses. Logistic univariable analyses with good outcome (being on the optimal dose one year later) were significant for Qb Inattention as dependent variable (odds ratio (OR) 2.641; 95% confidence interval (CI) 1.139–6.124) but not for the SNAP-IV inattention parent rating scale.

The probability of predicting a good outcome was calculated for the SNAP-IV inattention parent rating scale, Qb Inattention and Qb Activity, as well as for the combinations of SNAP-IV inattention parent rating scale and the two QbTest variables (see Table 3 for results). Of the SNAP-IV variables, only SNAP-IV inattention was analyzed since all other parent and teacher reported SNAP-IV scores showed no significance when compared with the treatment effects one year later.

The assessments by parents using the SNAP-IV inattention rating scale were analyzed and the children without an optimal dose, according to SNAP-IV, were identified. For these children we further analyzed the Qb Inattention results to find an optimal dose which lead to an increase of the sensitivity to a high level (see Table 3). For individuals without a clear result for inattention, either using the SNAP-IV parent rating scale or the Qb Inattention, we also analyzed the Qb Activity to find an optimal dose, and then almost all individuals (47 out of 48) with a good outcome were identified (sensitivity 0.98), but the specificity was low (see Table 3).

Drop-out analysis

Out of 118 children in Study group 1 (diagnostic), complete data were obtained for 91 children, leaving an attrition rate of 23%. In the ADHD group, data were incomplete for 21 out of 101 children, while in the non-ADHD group, six out of 44 children had missing data. The 27 children with missing data were compared with the 91 with complete data with regard to existing common variables. We found no significant differences, according to the Mann-Whitney U-test, for any of the analyzed variables (SNAP-IV parent and teacher rating scale, QbTest and Conners' CPT-II results, and age and gender).

Table 3. *Titration of the optimal dose, Study group 2 (medication)*

	Sensitivity	Specificity	True positive cases	True negative cases	False positive cases	False negative cases
Parent SNAP-IV inattention (n = 56)	0.56	0.75	27	6	2	21
Qb Inattention (n = 60)	0.82	0.60	41	6	4	9
Qb Activity (n = 60)	0.76	0.40	38	4	6	12
Parent SNAP-IV inattention + Qb Inattention (n = 56)	0.94	0.62	45	5	3	3
Parent SNAP-IV inattention + Qb Inattention + Qb Activity (n = 56)	0.98	0.25	47	2	6	1

Notes: Qb-test performed with calculation of the parameters Qb Inattention and Qb Activity. Outcome was defined as being on the optimal dose one year after titration. An optimal dose was defined as a decrease in SNAP-IV symptom scores or QbTest scores of >0.4 SD. Qb Activity, Qb Impulsivity and Qb Inattion. Sensitivity and specificity were calculated for the SNAP-IV parameters, the different QbTest scores and their stepwise combinations. In the stepwise analyses SNAP-IV ratings by parents were analysed first and if the results were inconclusive (i.e. no optimal dose could be identified), Qb Inattention was analysed. If results still were inconclusive an analysis of QB Activity was made. Qb Impulsivity was omitted from the analyses because this variable had no significant correlation with treatment results one year later.
Parent SNAP-IV inattention = parent ratings of the Swanson, Nolan and Pelham, version IV (SNAP-IV), scale for inattention.

In Study group 2 (medication), QbTest data were incomplete for altogether 130 out of 186 cases. Therefore, only 56 children had QbTest results from assessments of different doses of central stimulants, making titration possible. When comparing the different variables in the group with complete data and individuals with missing data, no significant differences based on the Mann-Whitney U-test were found concerning number of individuals on the same medication and indication of the optimal dose one year later, parent and teacher SNAP-IV scores, QbTest results or gender. A difference was found regarding age, with the group with complete data having a mean age of 12.3 (SD 2.5) years compared with 13.5 (SD 3.0) years for the group with missing data ($p = 0.003$).

DISCUSSION

In the current study we used Bayesian statistics, ROC analyses, sensitivity, specificity and likelihood ratios to investigate the incremental clinical utility of Conners' CPT II and QbTest in assessment for diagnostic and treatment monitoring purposes.

The first aim was to investigate the incremental clinical diagnostic utility of the Conners' CPT II and the QbTest using Bayesian statistics and ROC analyses in Study group 1 (diagnostic). Conners' CPT II used as stand- alone test did not show clinical utility defined as post-test probability greater than 85% chance of ADHD. When parent ratings and teacher ratings were inconclusive, the post-test probability was approximately equal to base-rate probability. Adding Conners' CPT II either markedly increased (if positive) or decreased (if negative) the post-test probability. In cases where both parent and teacher SNAP scores for combined-type ADHD were above the cut-off, the post-test probability was sufficiently high, 89% chance of ADHD. It seemed unnecessary to add more information in those cases. Therefore, our study confirms the suggestion of Jarrett *et al*. (2016) that the Conners' CPT II can add incremental diagnostic information in cases where the ADHD diagnosis is uncertain.

The second aim was to evaluate the clinical utility of using the QbTest as a supplement to SNAP-IV in dose titration of stimulant medication in Study group 2 (medication). Using the QbTest in titration of the optimal dose when parent SNAP inattention scores showed inconclusive results gave a prediction of a good outcome (being on the optimal dose one year later) with higher sensitivity but poorer specificity. When the parents' assessment was inconclusive, good sensitivity and specificity were achieved by supplementing parent SNAP-IV inattention scores with Qb Inattention scores. Studies on methylphenidate treatment titration have rarely been performed previously, especially not in realistic clinical situations. This study was performed retrospectively in patients with naturalistic titration and treatment follow-up. The drop-outs had similar symptoms and treatment, except that they were somewhat older in the titration study. Previous studies have shown that, compared with pre-pubertal children, adolescents more often discontinue their treatment within one year (McCarthy, Wilton, Murray, Hodgkins, Asherson & Wong, 2012). Our results show that the teacher SNAP-IV assessments and the parent SNAP-IV scores for hyperactivity/impulsivity and combined ADHD were not useful in finding an optimal dose with good outcome. It seems that the parental scores for inattention are the only reliable predictor and that teachers are not good at registering treatment effects. Somewhat surprisingly, the parent scores on hyperactivity/impulsivity were not good predictors of drug effects. A possible explanation is that the subjects were mostly pre-pubertal or adolescents and at this age hyperactivity usually declines. Parents' hyperactivity/impulsivity scores may be of more obvious prognostic value in assessment of younger children.

Combining parent SNAP-IV inattention scores with Qb Inattention scores and, where inconclusive, supplementing these with Qb Activity ratings, gave some false positives but the lowest number of false negatives (only one child). This may be useful in a first screening to determine the optimal dose, since children with inconclusive titration results can then be regarded as non-responders. The rate of children diagnosed with ADHD and getting medicine is low at the clinic in Malmö compared with other parts of Sweden, indicating a problem of underdiagnosing and undermedication. Children diagnosed at the Malmö clinic probably have a real need for effective treatment including medical treatment and false negative cases should be avoided. False positive cases can be identified during the follow-up. On the other hand, having many false negatives regarded as non-responders would lead to many children not getting the chance to benefit from a medicine.

Limitations

One important limitation is that the sample sizes were rather small with the obvious risk of type two error. Limitations according to analyses concerning Study group 1 (diagnostic): our approach of using Bayes' formula in analyzing combinations of tests may lead to an overestimation of the posterior probabilities as a result of correlations between the variables. However, the correlations between parent ratings and CPT results have been found to be low. There is also a risk of "criteria contamination," that is to say that the assessment method in the analyses was part of the diagnostic process (Youngstrom, 2014). Despite negative results for Conners' CPT II confidence index, 33 out of 59 children were diagnosed with ADHD, because other kinds of information, such as derived from a thorough child psychiatric examination including a clinical semi-structured interview, were important for the diagnosis. A problem in a clinical setting is that many children coming for assessment will be diagnosed as not having ADHD. They might still have "subclinical" ADHD or other diagnoses such as depression or anxiety and these children can have increased Conners' CPT II and SNAP-IV scores, making the discriminative power of the tests only moderate. ADHD in turn is very likely to be comorbid with other psychiatric diagnoses. Nevertheless, there was no statistical difference between the ADHD group and the non-ADHD group concerning the BCFPI screening of externalization and internalising comorbidity (see Table 1). Consequently, comorbidity should not have affected the results. Another limitation is that there was some missing data because of the naturalistic design of the study. To avoid bias, attrition analyses were performed with Mann Whitney U-test regarding SNAP-IV parent and teacher rating scale, QbTest and Conners' CPT-II results, age and gender. The drop-outs did not

differ from the rest of the patients. Further, the clinical data were not derived from a standardized structured diagnostic interview. This may have caused limited generalizability. Instead data from the standardized structured screening phone interview BCFPI were included in the study to report symptoms, functional level and risk factors.

Limitations according to Study group 2 (medication): obvious limitations are that there was no available information on comorbidity and cognitive level, and there was a large attrition concerning incomplete Qb data, reducing statistical power. In the drop-out analysis with Mann Whitney U-test, age was the only variable found to be statistically different between drop-outs and individuals with complete data. No statistically significant difference was found regarding parent and teacher SNAP-IV scores, QbTest results or gender.

CONCLUSION

As previously suggested, the Conners' CPT II could be clinically useful as a supplemental diagnostic assessment tool in unclear cases of ADHD (Jarrett *et al*., 2016). Where parent and teacher SNAP-IV ratings were consistent, adding Conners' CPT II did not add incremental value to the ADHD diagnosis. On the other hand, where parent and teacher SNAP-IV ratings were inconsistent, adding Conners' CPT II provided useful information either supporting or rejecting the ADHD diagnosis, and possibly decreasing uncertainty in the clinic. When evaluating treatment with central stimulant medication by titration, parents' assessment of inattention can give important information in identifying the optimal clinical dose. However, when the parents' assessment fails to identify such a dose, the QbTest, especially a combination of Qb Inattention and Qb Activity, may provide reliable data which can be of help in determining the optimal dose.

The authors received financial support from Psychiatry Region Skåne, the Research and Development Department of the Southern Sweden Health Care Region, and the Lindhaga Foundation.

REFERENCES

Akobeng, A. K. (2007). Understanding diagnostic tests 2: Likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Paediatrica*, *96*, 487–491.

Ashby, D. (2006). Bayesian statistics in medicine: A 25 year review. *Statistics in Medicine*, *25*, 3589–3631.

American Psychiatric Association. (1994). *Diagnostic and Sstatistical manual of mental disorders: DSM–IV: prepared by the Task Force on DSM–IV* (4th edn). Washington, DC: American Psychiatric Association.

Boyle, M. H., Cunningham, C. E., Georgiades, K., Cullen, J., Racine, Y. & Pettingill, P. (2009). The Brief Child and Family Phone Interview (BCFPI): 2. Usefulness in screening for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, *50*, 424–431.

Brocki, K. C., Tillman, C. M. & Bohlin, G. (2010). CPT performance, motor activity, and continuous relations to ADHD symptom domains: A developmental study. *European Journal of Developmental Psychology*, *7*, 178–197.

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*, 317–328.

Conners, C. K., Epstein, J. N., Angold, A. & Klaric, J. (2003). Continuous performance test performance in a normative epidemiological sample. *Journal of Abnormal Child Psychology*, *31*, 555–562.

Conners, K. & MHS staff. (2000). *Conners' Continuous Performance Test for Windows (CPT II)*, Toronto: Multi Health Systems Inc.

Conners, K. & MHS staff. (2004). *Conners' Continuous Performance Test (CPT II): Version 5 for Windows, Technical guide and software manual*. Toronto: Multi Health Systems Inc.

Cunningham, C. E., Boyle, M. H., Hong, S., Pettingill, P. & Bohaychuk, D. (2009). The Brief Child and Family Phone Interview (BCFPI): 1. Rationale, development, and description of a computerized children's mental health intake and outcome assessment tool. *Journal of Child Psychology and Psychiatry*, *50*, 416–423.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E. & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin*, *141*, 858–900.

De Los Reyes, A., Youngstrom, E. A., Pabon, S. C., Youngstrom, J. K., Feeny, N. C. & Findling, R. L. (2011). Internal consistency and associated characteristics of informant discrepancies in clinic referred youths age 11 to 17 years. *Journal of Clinical Child & Adolescent Psychology*, *40*, 36–53.

Fernández-Jaén, A., Fernández–Mayoralas, D. M., Pardos, A., Calleja-Pérez, B. & Jareño, N. M. (2009). Clinical and cognitive response to extended–release methylphenidate (Medikinet®) in attention deficit/hyperactivity disorder: efficacy evaluation. *Advances in Therapy*, *26*, 1097–1110.

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*, 543–555.

Frazier, T. W., Youngstrom, E. A. & Hamilton, J. D. (2006). Evidence-based assessment of attention–deficit/hyperactivity disorder: using multiple sources of information. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*, 614–620.

Gaub, M. & Carlson, C. L. (1997). Behavioral characteristics of DSM–IV ADHD subtypes in a school-based population. *Journal of Abnormal Child Psychology*, *25*, 103–111.

Hall, C. L., Valentine, A. Z., Groom, M. J., Walker, G. M., Sayal, K., Daley, D. & Hollis, C. (2016). The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: a systematic review. *European Child and Adolescent Psychiatry*, *25*, 677–699.

Homack, S. & Riccio, C. A. (2006). Conners' Continuous Performance Test (2nd edn; CCPT–II). *Journal of Attention Disorders*, *9*, 556–558.

Huang-Pollock, C. L., Karalunas, S. L., Tam, H. & Moore, A. N. (2012). Evaluating vigilance deficits in ADHD: A meta-analysis of CPT performance. *Journal of Abnormal Psychology*, *121*, 360–371.

Hunsley, J. & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology*, *3*, 29–51.

Jarrett, M. A., Van Meter, A., Youngstrom, E. A., Hilton, D. C. & Ollendick, T. H. (2016). Evidence-based Assessment of ADHD in Youth Using a Receiver Operating Characteristic Approach. *Journal of Clinical Child & Adolescent Psychology*, *47*, 808–820.

Jensen, P. S. (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, *56*, 1073–1086.

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S. A., Friedman, L. M. & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical Psychology Review*, *33*, 795–811.

Konrad, K., Günther, T., Hanisch, C. & Herpertz-Dahlmann, B. (2004). Differential effects of methylphenidate on attentional functions in children with attention–deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*, 191–198.

Lipszyc, J. & Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task.

Journal of the International Neuropsychological Society, *16*, 1064–1076.

Losier, B. J., McGrath, P. J. & Klein, R. M. (1996). Error patterns on the continuous performance test in non–medicated and medicated samples of children with and without ADHD: A meta-analytic review. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*, 971–987.

Martel, M. M., Markon, K. & Smith, G. T. (2017). Research review: Multi-informant integration in child and adolescent psychopathology diagnosis. *Journal of Child Psychology and Psychiatry*, *58*, 116–128.

Mash, E. J. & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal of Clinical Child & Adolescent Psychology*, *34*, 362–379.

McCarthy, S., Wilton, L., Murray, M. L., Hodgkins, P., Asherson, P. & Wong, I. C. (2012). Persistence of pharmacological treatment into adulthood, in UK primary care, for ADHD patients who started treatment in childhood or adolescence. *BMC Psychiatry*, *12*, 219. https://doi.org/10.1186/1471-244X-12-219

Nichols, S. L. & Waschbusch, D. A. (2004). A review of the validity of laboratory cognitive tasks used to assess symptoms of ADHD. *Child Psychiatry and Human Development*, *34*, 297–315.

Park, S., Kim, B. N., Cho, S. C., Kim, J. W., Shin, M. S., Yoo, H. J., et al. (2013). Baseline severity of parent-perceived inattentiveness is predictive of the difference between subjective and objective methylphenidate responses in children with attention-deficit/hyperactivity disorder. *Journal of Child and Adolescent Psychopharmacology*, *23*, 410–414.

Pelham, J., William, E., 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*, 449–476.

Posserud, M. B., Ullebo, A. K., Plessen, K. J., Stormark, K. M., Gillberg, C. & Lundervold, A. J. (2014). Influence of assessment instrument on ADHD diagnosis. *European Child and Adolescent Psychiatry*, *23*, 197–205.

Preston, A. S., Fennell, E. B. & Bussing, R. (2005). Utility of a CPT in diagnosing ADHD among a representative sample of high-risk children: a cautionary study. *Child Neuropsychology*, *11*, 459–469.

Ramtvedt, B. E. & Sundet, K. (2014). Relationships between computer-based testing and behavioral ratings in the assessment of attention and activity in a pediatric ADHD stimulant crossover trial. *Clinical Neuropsychology*, *28*, 1146–1161.

Riccio, C. A., Reynolds, C. R., Lowe, P. & Moore, J. J. (2002). The continuous performance test: a window on the neural substrates for attention? *Archives of Clinical Neuropsychology*, *17*, 235–272.

Schatz, A. M., Ballantyne, A. O. & Trauner, D. A. (2001). Sensitivity and specificity of a computerized test of attention in the diagnosis of attention-deficit/hyperactivity disorder. *Assessment*, *8*, 357–365.

Seixas, M., Weiss, M. & Muller, U. (2012). Systematic review of national and international guidelines on attention-deficit hyperactivity disorder. *Journal of Psychopharmacology*, *26*, 753–765.

Silberstein, R. B., Pipingas, A., Farrow, M., Levy, F. & Stough, C. K. (2016). Dopaminergic modulation of default mode network brain functional connectivity in attention deficit hyperactivity disorder. *Brain and Behavior*, *6*, e00582.

Spencer, S. V., Hawk, L. W., Richards, J. B., Shiels, K., Pelham, W. E. & Waxmonsky, J. G. (2009). Stimulant treatment reduces lapses in attention among children with ADHD: The effects of methylphenidate on intra-individual response time distributions. *Journal of abnormal child psychology*, *37*, 805–816.

Swanson, J. M., Kraemer, H. C., Hinshaw, S. P., Arnold, L. E., Conners, C. K., Abikoff, H. B., et al. (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. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*, 168–179.

Ulberstad, F. (2012). *QbTest Technical Manual* (Rev. edn.). Stockholm: Qbtech AB.

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*, 1336–1346.

Youngstrom, E. A. (2013). Future directions in psychological assessment: combining evidence-based medicine innovations with psychology's historical strengths to enhance utility. *Journal of Clinical Child & Adolescent Psychology*, *42*, 139–159.

Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *Journal of Pediatric Psychology*, *39*, 204–221.

Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D. & Jensen-Doss, A. (2015). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*, *22*, 20–35.

Received 23 March 2018, accepted 19 September 2018