The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: A systematic review

Authors

Hall CL[1], Valentine AZ[1], Groom M[2], Walker G[3], Sayal K[4], Daley D[5]., & Hollis C[6]

Affiliations

*[1]Charlotte L Hall, Research Fellow, Division of Psychiatry and Applied Psychology, CLAHRC-EM, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

*[1]Althea Z Valentine, Research Assistant, Division of Psychiatry and Applied Psychology, CLAHRC-EM, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

[2]Madeleine J Groom, Assistant Professor, Division of Psychiatry and Applied Psychology, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

[3]Gemma M Walker, Research Assistant, Division of Psychiatry and Applied Psychology, CLAHRC-EM, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

[4]Kapil Sayal, Professor of Child & Adolescent Psychiatry, Developmental Psychiatry, School of Medicine, University of Nottingham & CANDAL (Centre for ADHD and Neuro-developmental Disorders across the Lifespan), Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

[5]David Daley, Professor of Psychological Intervention and Behaviour Change, School of Medicine, University of Nottingham & CANDAL (Centre for ADHD and Neuro-developmental Disorders across the Lifespan), Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

[6]Chris Hollis, Professor of Child & Adolescent Psychiatry and Consultant Psychiatrist, University of Nottingham, CANDAL (Centre for ADHD and Neuro-developmental Disorders across the Lifespan) & MindTech, Developmental Psychiatry, University of Nottingham, Queen's Medical Centre, Nottingham, NG7 2UH, UK

***Joint 1st Author**

Corresponding Author

Dr Charlotte L Hall

CLAHRC-EM, Institute of Mental Health, University of Nottingham Innovation Park, Triumph Road, Nottingham, NG7 2TU, UK

Email: Charlotte.hall@nottingham.ac.uk. Phone: 0115 823 2438. Fax: 0115 823 1289

Abstract

Attention deficit hyperactivity disorder (ADHD) is typically diagnosed using clinical observation and subjective informant reports. Once children commence ADHD medication, robust monitoring is required to detect partial or non-responses. The extent to which neuropsychological continuous performance tests (CPTs) and objective measures of activity can clinically aid the assessment and titration process in ADHD is not fully understood. This review describes the current evidence base for the use of CPTs and objectively measured activity to support the diagnostic procedure and medication management for children with ADHD. Four databases (PsycINFO, Medline, Allied and Complementary Medicine (AMED) and PsycARTICLES) were systematically searched to understand the current evidence base for: (1) the use of CPTs to aid clinical assessment of ADHD; (2) the use of CPTs to aid medication management; (3) the clinical utility of objective measures of activity in ADHD. Sixty relevant articles were identified. The search revealed six commercially available CPTs that had been reported on for their clinical use. There were mixed findings with regard to the use of CPTs to assess and manage medication, with contrasting evidence on their ability to support clinical decision making. There was a strong evidence base for the use of objective measures of activity to aid ADHD/non-ADHD group differentiation, which appears sensitive to medication effects and would also benefit from further research on their clinical utility. The findings suggest that combining CPTs and an objective measure of activity may be particularly useful as a clinical tool and worthy of further pursuit.

Keywords: Attention Deficit Hyperactivity Disorder (ADHD); Continuous Performance Tests (CPT), Activity, Objective Measures, Systematic Review

Introduction

Attention deficit hyperactivity disorder (ADHD) is a common mental health disorder of childhood affecting approximately 4-8% of school age children [1]. This neurodevelopmental disorder is characterised by three core symptom domains; inattention, hyperactivity and impulsivity. NICE guidelines provide a blueprint for the diagnosis and management of ADHD in England and Wales and indicate the need for young people with ADHD to have access to the best evidence-based care in order to fulfil their potential and prevent poor outcomes [1]. However, in practice, delivery and quality of care is patchy, with little consistency in assessment, diagnosis or management[2].

It has been suggested that ADHD is 'symptom complex', stemming from multiple causes, such as genetics, biological and psychosocial influences [3], resulting in a range of presenting behaviours [4,5]. Given the variation in causes and behavioural consequences of ADHD, there is no single test used to diagnose the disorder and the clinician's judgment is currently the most widely accepted method of assessment. For the clinician to determine a diagnosis of ADHD they will generally gather information from the parents, teachers (and the child themselves where age appropriate), make clinical observations, conduct school observations, and may use tests of behaviour and neuropsychological functioning. However, there is a paucity of clinical guidance on which combination of measures should be used in the diagnostic assessment of ADHD. Furthermore, this approach is heavily reliant on subjective measures, which can lead to discrepancies in the diagnosis of ADHD [4] and the process of interview and data collection is lengthy and difficult to conduct in real world settings. Additionally, once on medication, monitoring may not be adequate or frequent enough to detect early non- or sub-optimal response [6]. Given the subjective nature of ADHD assessment, and the heterogeneous nature of the disorder, it is not surprising that it can be seen as a controversial diagnosis amongst clinicians and the public. The demonstration of a reliable and valid objective measure to aid this assessment may help partially counteract these attitudes.

Objective measures have the potential to augment and streamline current practice in order to shorten assessment time, increase diagnostic accuracy, reduce delays in treatment, and optimise treatment response. Continuous performance tests (CPTs) are neuropsychological tests that measure the individual's attention and impulsivity in a sustained task, and can be used alongside clinical inquiry as part of the diagnostic procedure. Typically, a CPT is a computer-based programme which involves rapid presentation of visual or auditory stimuli. Participants are asked to respond when a given target occurs but to withhold the response to non-targets. A standard CPT typically records two of the core features of ADHD; selective attention is measured through the child's omission errors (responding when the target is present), sustained attention is measured through decrements in performance during the course of the test, including measures of reaction time and reaction time variability, and impulsivity is measured through commission errors (responding when the target is not present). The number of correct responses is also recorded.

Several studies and systematic reviews have demonstrated the ability of CPTs to differentiate children with ADHD from other diagnoses or healthy controls [7,8]. Similarly, several papers report CPTs to be sensitive to the effects of ADHD medication [8,9]. There are several commercially available CPTs including the Test of Variables of Attention (TOVA [10]), the Conner's CPT (CCPT; Conners [11]), the Gordon's Diagnostic System

(GDS [12]) and the QbTest (QbTech Ltd). Indeed CPTs are considered to be one of the most popular clinic based measures to assess sustained attention [13]. However, the majority of CPTs do not measure the patient's activity level, a distinguishing feature of ADHD which may help in accurate differentiation from other diagnoses. Current approaches to the objective measurement of activity in ADHD have focused on actigraphy devices or infra-red devices to capture movement. Two CPTs, the QbTest and the Quotient ADHD system, combine a standard CPT with an infra-red camera which tracks the path of a reflector attached to the participant's head to measure motor activity during the course of the test, allowing information to be gathered on all three core symptoms of ADHD.

Ogundele and colleagues [4] provide a generic overview of the role of CPTs in ADHD and conclude that CPTs can provide a more objective insight into the young person's behaviour and allow parents to gain a better insight into the nature of their child's condition which may also help improve adherence to treatment. They suggest that combining CPTs with clinical judgement and rating scales may provide the optimal format for diagnosing and managing ADHD. Indeed, there has been an increase in the popularity of the CPTs as a clinical tool [11]. However, most research and previous systematic reviews have focused on the psychometric properties of the CPT rather than how these tools may aid the clinical assessment or management process of ADHD [14]. Given the increase in the use of these tools and that some UK NHS ADHD clinics have incorporated combined CPT and objective activity measures as part of standard diagnostic assessment, it is important to understand the clinical worth of these tests.

The aim of this systematic review is to provide an overview of the evidence for commercially available CPTs that have been used for aiding the clinical diagnostic and medication process for children and young people with ADHD, and in doing so we will identify gaps in the evidence to provide future lines of investigation. The review focuses on CPTs that have specifically been investigated for their use as a clinical tool, as opposed to neuropsychological test. Given that to date few reviews have focussed on the objective measure of activity in ADHD and this variable is increasingly considered an important adjunct to CPTs in ADHD assessment [15] we also searched for existing measures which objectively quantify activity/hyperactivity. This review was carried out as part of the National Institute for Health Research (NIHR) CLAHRC-EM (Collaborations for Leadership in Applied Health Research and Care – East Midlands).

Method

A systematic search was conducted using the following databases: PsycINFO, Medline, Allied and Complementary Medicine (AMED) and PsycARTICLES using the Ovid search engine in June 2014 and updated prior to submission in June 2015. Different combinations of various search terms were used to ensure that all relevant papers referring to commercially available CPTs for aiding the diagnosis and treatment of ADHD or measures of activity movement (such as hyperactivity) were found (see Table 1). Informal searching was also undertaken, including hand searching of article references and web search engines (Google). Inclusion criteria were: publication in a peer reviewed journal in English; a study of any design focusing on the clinical use of CPTs in children or adolescents (aged up to 18 years), or objective measures of activity in children or

adolescents. At least some participants in each paper were either diagnosed with ADHD or had been referred for an ADHD assessment. Papers were excluded if they did not meet the inclusion criteria, did not include a human sample and did not use a commercially available CPT. Titles and abstracts were reviewed by two authors (CLH, AZV) according to the inclusion and exclusion criteria; where there was disagreement, this was resolved by discussion between the two reviewers. The full text of retained papers was reviewed by these two authors. Data extraction was performed by CLH and AZV using a data collection form, any discrepancies were discussed in a meeting until consensus was reached.

<<Insert Table 1 about here>>

Given that research evidence on the clinical utility of CPTs is likely to incorporate a variety of different methodologies, it was considered unwise to include study design as an inclusion criterion. As such, no systematic judgments were made about the strengths and weaknesses of each design, however, each method was considered in relation to any unique or specific elements of that study in its ability to answer the question posed in this review.

Results

The combined searches resulted in a total of 1421 articles (after removal of duplicates) of which 60 articles were identified as relevant and included in the review (Figure 1).

<<Insert Figure 1 about here>>

Of the included studies, 24 reported on the clinical utility of CPTs to diagnose (Table 2), 12 reported information relevant to the clinical utility of CPTs to aid medication management (Table 3) and 25 reported objective measures of activity in ADHD (Tables 4 & 5). Although most studies were based in the USA (30), papers were from a variety of locations including: Germany (6), the UK (4), Canada (3), Japan (2), Norway (2), Taiwan (3), Australia (1), Denmark (1), Israel (1), Romania (1), South Korea (2), Brazil (1), Spain (1) and Sweden (1). In addition, one meta-analysis, conducted in Italy, was included in the review.

The search identified six commercially available CPTs that had been specifically investigated for their clinical utility the: TOVA, GDS, CCPT, Integrated Visual and Auditory Continuous Performance Test (IVA+CPT), Quotient ADHD system (or McLean Motion and Attention Test; MMAT) and the QbTest. Although not the purpose of this paper, it is worth mentioning that the included papers from this search cite good psychometric properties (e.g., test-retest reliability) of each CPT (e.g., TOVA [16]; GDS [17]; CCPT [18]; IVA+CPT;[19]; QbTest [20]; MMAT [21].

The clinical use of CPTs to aid the assessment of ADHD

The 24 papers that reported on CPTs clinical utility to aid ADHD assessment reported on a combined total of five CPTs, of which six papers report on the TOVA [22-25,16,26], 11[1] report on the GDS [27-32,17,33-37], five report on the CCPT [38,7,14,39,40], one reports on IVA+CPT [19] and one reports on the QbTest [41]. A summary of the papers is presented in Table 2. For the purpose of this review, we considered relevant outcomes to include papers that reported: sensitivity/specificity, positive and negative predictive power, methods to distinguish group membership or severity of symptoms and feasibility of incorporating the test in routine practice. Papers that also qualitatively reported on their experience of using CPTs to aid assessment were also included. It should be noted that the above outcomes are likely to differ according to the sample used in the study, for example, clinic-referred children are more likely to have a presenting disorder and co-morbidities than a community sample. We have highlighted the sample type in Table 2.

<<Insert Table 2 about here>>

Test of Variables of Attention (TOVA)

The TOVA presents a target stimulus in the form of a small square with a hole near the top or bottom of the square. Participants are requested to respond by pressing a hand-held switch when the hole is on the top of the square and not respond when the hole is on the bottom. The stimuli are presented for 100 milliseconds (msec) at a between stimuli interval of 2,000msec; the test lasts for a duration of approximately 22.5 minutes. The TOVA utilises both a 'rare target' and 'response inhibition' paradigm by presenting 22.5% targets and 77.5% non-targets in the first half of the test and then 77.5% of targets and 22.5% of non-targets in the second half of the test. Variables include omission errors (number of missed targets), commission errors (response to non-targets), response time (mean response latency), variability (standard deviation of response times), number of multiple responses (number of stimuli to which the participant responded more than once) and anticipatory responses (very short latency responses). The TOVA also computes an ADHD score using the formula of response time Z score (Half 1) +d'Z score (half 2) + variability Z score (total). The ADHD score is a comparison between the participant's performance to that of a known ADHD sample based on 178 children (148 male, 30 female; aged 6-15 years) who had been diagnosed with ADHD by a senior healthcare professional. Six studies report information relevant to the clinical utility of the TOVA. Schatz et al.[24] report the TOVA's sensitivity (correct identification of ADHD) and specificity (correct rejection of ADHD) as 85.7% and 70.0% respectively in a clinically diagnosed sample of ADHD children and controls. Using the Receiver-Operating Characteristic (ROC) curve, they found the TOVA was a poorer predictor of ADHD diagnosis than the Conners' Hyperactivity Index, which showed no false positives. Logistic regression showed that only the TOVA variable of 'variability' was able to significantly predict group membership, with the addition of other TOVA variables not improving the ability to predict group membership.

[1] Excluding an erratum to the *Rielly et al* article [27], the correct specificity figures were obtained from the *Rielly et al* erratum [28].

Of the five studies that used the TOVA to specifically aid clinical diagnosis, four reported the TOVA to have some clinical utility in aiding the assessment of ADHD [25,26,16,22]. Specifically, the TOVA was noted to be easily grasped by children and was useful in aiding the assessment process in children of different ethnicities and ages as it avoided the need for linguistic skills and could be easily incorporated into a busy outpatient clinic [25]. TOVA was noted by two studies to provide a unique and important source of information to the assessment process. Forbes et al. [26] combined TOVA with standard rating scales and found that the Revised Conners' Teacher Rating Scale and TOVA did not identify identical groups of children, with each correctly classifying children misclassified by the other measure. They suggest that rating scales and TOVA are measuring distinct but important aspects of ADHD and the addition of TOVA makes a valuable contribution to assessment of ADHD. In support of this, Chae et al. [16] used TOVA combined with parent and teacher rating scales to diagnose ADHD in gifted children. They found the TOVA to be a valid tool in determining ADHD diagnosis, and reported that 13.2% of gifted children who were rated as having ADHD on the rating scales, but had normal TOVA scores were not classified as having ADHD. The authors suggest that ADHD may be over-diagnosed if an objective measure such as TOVA is not used in diagnosing ADHD in gifted children, again supporting the TOVA alongside rating scales as a unique source of information.

Two papers report on the clinical utility of TOVA in determining ADHD sub-types, with mixed results. Porumb [22] report a single case study of an 11-year-old girl referred for an ADHD assessment. The author reports that although the TOVA did not predict the number of ADHD symptoms, it was useful in determining sub-type and in differentiating ADHD from other attention-based symptoms. In contrast, the Forbes et al. [26] study reported above, which is based on data from 146 private-practice referred children found no difference on TOVA variables between ADHD and ADD group.

Another study [23] found that the TOVA was not clinically useful in distinguishing children with ADHD from children with sub-clinical level ADHD-type difficulties. The two groups were sampled using an epidemiologically derived sample of children and selecting those at high risk for ADHD. The study compared TOVA performance between children meeting the DSM-IV criteria for ADHD and those displaying sub-clinical levels of behavioural and cognitive problems. They found no significant difference in TOVA scores between these two groups [23], suggesting the TOVA would not be a useful clinical adjunct when determining borderline cases. Furthermore, Preston et al. [23] found the TOVA could not predict the number of ADHD symptoms, independent of group membership. They concluded that the TOVA does not increase diagnostic accuracy in determining ADHD or sub-clinical difficulties in attention/impulsivity/activity, which are the sample likely to be seen by a clinician.

Gordon's Diagnostic System (GDS)

The GDS was designed specifically for clinical use and allows multiple tasks to be administered. The GDS provides normative data based upon protocols of 1,300 non-hyperactive boys and girls aged 4-16 years-old. The GDS consists of three sub-tasks: the Vigilance Task, the Delay Task and the Distractibility Task. The Vigilance Task presents numbers on a display screen at the rate of 1/sec. The stimulus is presented for 800msec with a 200msec delay. The participant is required to press a button whenever the number *1* is followed by the number *9*. There are a total of 45 target pairs presented during the task. For young children there is a variant of

this task which asks for participants to respond every time a *1* is presented. The Vigilance Task records the number of correct responses, number of omissions and commissions. The Delay Task requests the participant to not respond. The participant is asked to press a button and wait before re-pressing. If they refrain from re-pressing for a minimum of 6sec a light flashes and reward points are allocated. If the participant responds prior to this the timer resets and no points are awarded. The task records three primary scores: the number of responses, the number of correct responses and an Efficiency Ratio (percentage of correct responses). A variant of the Vigilance Task was also created to assess the participant's distractibility. The Distractibility Task is identical to the standard Vigilance Task, except for the presence of flashing digits that are presented at random intervals on the outer edges of the screen. Administration of all three sub-tasks takes approximately 26 minutes.

Twelve papers report on the utility of the GDS to aid clinical assessment of ADHD. One paper presents a ROC curve to determine if a child had severe ADHD or was in the community control sample [32]. For the three GDS indices of efficiency ratio (generated from Delay Task), commission errors and correct response, they found areas under the ROC curve ranging from 0.72-0.73, meaning a randomly selected individual with ADHD would have a poorer GDS score than a control child 72-73% of the time. The scores for sensitivity varied between the three variables from 49-59% and specificity between 81-87%, suggesting, as a single test the GDS was unable to accurate assign group classification.

Mayes et al. [33] investigated the GDS accuracy in differentiating subtypes of combined and inattentive ADHD. They found the ADHD combined subgroup had greater impulsivity than the inattentive subgroup (measured through Delay Task), but were equally impaired on vigilance and distractibility, which measures attention, indicating the validity of the Delay Task to aid sub-group differentiation. Classification accuracy of the GDS in determining the sub-groups was 69.7% for the Delay Task alone and 70% for Delay plus Distractibility. When Combined with WISC Freedom from Distractibility/Working Memory Index and Processing Speed Index, the GDS could differentiate ADHD subtypes with 72% accuracy.

Trommer et al. [34] found that 28.6% children diagnosed with attention deficit disorder (ADD) performed in the normal range on the task, with 35.7% performing within the borderline range and 35.7% performing in the abnormal range. Furthermore, 66.6% of non-ADD children performed in the abnormal range on the CPT, questioning the sensitivity of this task in aiding group differentiation. However, in order to determine the potential effect of IQ on GDS score, Mayes et al [17] scored the GDS relative to IQ to determine classification accuracy. Using a 13-point or larger discrepancy between IQ and the GDS to classify ADHD combined-type resulted in an accuracy rate of 86% (sensitivity = 90%, specificity = 86%). Of the children who had an abnormal GDS score (below 90), 90% had a diagnosis of ADHD (positive predictive power) and 52% of healthy control children did not have ADHD (negative predictive power). These figures for positive and negative predictive power rose to 91% and 67% respectively when IQ minus the GDS composite was 13 or more points. Matier-Sharma et al. [35] report correct classification of 62.5% of ADHD and 94.9% of healthy control children on the GDS. Through a series of papers using similar data sets, Grodzinsky and colleagues report GDS scores ranging from 83-87% positive predictive power and 59-61% negative predictive power [36,37].

Three studies support the use of the GDS as part of the assessment process [27,28,31,30]. Two papers [27,30] found the GDS to be useful for excluding an ADHD diagnosis. Using the IQ GDS cut-point described above, Mayes and Calhoun [30] found that 87.8% of children were correctly identified as having ADHD. When combining this score with IQ minus Freedom from Distractibility >11, diagnostic accuracy was 90.9%. Their findings also showed that the GDS was better at ruling out ADHD than Freedom from Distractibility scores (WISC-III) in a clinic referred sample and suggest that GDS should be used in clinical practice with children who do not meet ADHD cut-off on Freedom of Distractibility to be confident the child does not have ADHD. Rielly et al. [27,28] conducted a classification analysis comparing GDS scores with parent and teacher ratings of ADHD in children with language disorder. They found that the GDS showed low positive predictive values (20.0% to 36.8%) and high negative predictive values (71.9% to 87.9%), indicating that the test may not be useful in diagnosing ADHD. Based on this, Rielly et al. [27] suggest that the GDS is useful as part of a battery of assessment to rule out ADHD in communicative disorders clinics, but not in confirming the diagnosis, it should be noted that these findings are specific to that found in this specific clinic. The other paper [31] reports a series of case studies that used the GDS to aid the assessment process. Gordon concludes that whereas the GDS only provides a snapshot on behaviour and should not be used conclusively to form a diagnosis, in each case the GDS helped an accurate and efficient assessment, specifically allowing the clinician to use data based on the child's actual behaviour [31].

One study reported limited validity of the GDS to aid the assessment process in clinical practice [29]. By comparing GDS scores on a clinic sample referred for assessment of ADHD they found that scores on the GDS frequently disagreed with the diagnosis of ADHD based on parent interview and behaviour rating scales. Based on these results the authors suggested that there is poor ecological validity of the GDS, which may be because the test does not correspond to the child's behaviour exhibited at home, or that rating scales assess behaviour over a period of time rather than the snapshot provided by the CPT. Although not suggested by the authors, an alternative may be that the GDS is tapping into a unique aspect of ADHD that is not captured by rating scales.

Conners' CPT (CCPT)

The CCPT consists of 360 trials in which a single letter appears on the screen for 250ms. The participant is asked to press the space bar when every letter appears except X. The CCPT takes approximately 14 minutes to complete and involves frequent responding. It is based on a 'response inhibition' paradigm which contains 90% target stimuli and 10% non-target stimuli. In the CCPT the inter-stimulus interval is either 1.5s or 3s to prevent practice effects. Variables of omission and commission errors, RT, and RT variability are calculated. Additionally, the CCPT computes an overall index of attention problems, which is calculated from RT, RT variability and omission errors. Norms for the CCPT are based on a non-referred community population and clinic-referred cases with an ADHD diagnosis, thus providing *T*-scores. An overall index that exceeds 11 is considered a cut-off for attention problems [14]. Perugini et al. [39] report sensitivity and specificity of 67% and 73% respectively when classifying ADHD boys from community control boys. Whereas Alloway et al. [38] used discriminant function analysis to determine CCPT classification scores for children with ADHD and

healthy controls (drawn from a community sample). They report CCPT scores correctly classified 41% of ADHD children and 65% of healthy controls, Epstein et al. [7] investigated the prediction of ADHD symptoms by CPT performance. They found CPT measures (mean hit RT, hit RT (*SE)*, commission errors, omission errors, *d', β*) were significantly related to DSM-IV ADHD symptoms but did not demonstrate symptom domain specificity [7].

Kleinman et al. [40] compared patterns of attention deficits on CCPT in children with a diagnosis of bi-polar disorder (BD), ADHD and combined BD and ADHD with no co-morbidities. They found no significant difference in CCPT scores between any groups, suggesting that CCPT scores could not differentiate between diagnostic groups. Using cluster analysis based on the CCPT scores, the sample was best clustered into two new groups (A+B), which were independent of their original diagnosis. ADHD and BD+ADHD were evenly split between the two groups; however, Group A had greater functional impairment on CCPT and rating scales. The authors suggest CCPT results may be used to create clinically homogenous groups that could help clinician's to understand the individual's difficulties and implement specific treatments for these difficulties.

One study investigated the utility of the CCPT to aid ADHD assessment [14]. Using a clinic-referred sample they found no association between CCPT scores and parent and teacher ratings of symptoms, nor could the test distinguish between ADHD and clinical controls. They report a further weakness of the CCPT is its association with linguistic capability. McGee et al [14] found that phonological skills were associated with phonological awareness, and suggest that the task should not be used in children with reading disorders. They conclude that the CCPT has questionable value as a diagnostic tool and may be particularly prone to over-diagnosing ADHD in children with reading difficulties. Although the authors also present the auditory CPT (ACPT), they do not discuss this in terms of its use as an assessment tool in clinical practice and it is therefore not reported here.

Integrated Visual and Auditory Continuous Performance Test (IVA+CPT)

The IVA+CPT was designed to aid the diagnosis of ADHD and utilises both auditory and visual stimuli. Participants are asked to click a button when they hear or see the number '1', but not click when they hear or see a number '2'. The test takes approximately 20 minutes to complete. IVA+CPT is normed on data from 1700 men and women aged 6-96 years. The test computes 12 quotients with separate visual and auditory scores. The scores measure omission and commission errors, which are used to calculate a hyperactivity-impulsiveness and attention deficit scale. Each item consists of 3 auditory quotients and 3 visual quotients. One study investigated the clinical utility of quantitative electroencephalography (QEEG) and IVA+CPT to assess ADHD [19] . They found all of the IVA+CPT measures were significantly different between the ADHD and control group, and report sensitivity of 72.9% and specificity 70.9% in detecting ADHD. The findings suggest that the IVA+CPT could be a useful tool in aiding ADHD assessment, however, there is limited research investigating this tool for its clinical utility.

QbTest

The QbTest (Qbtech Ltd) combines a CPT to measure attention and impulsivity with an infra-red motion capture of head movement to measure activity. The QbTest presents a target stimulus (a grey circle) and

a non-target (grey circle with a cross); participants are requested to press a hand-held button every time the target is presented. The stimuli are presented on the screen for 100ms, at an interval of 1,900ms. A total of 450 stimuli are presented with an equal number of targets and non-targets. The stimuli are always presented in the same location with a fixed between-stimuli interval. Physical activity is measured during the CPT via an infrared camera that tracks the path of a reflector attached to the centre of participant's forehead. These elements of the test provide information on each of the three symptom domains of ADHD and provide summary scores for each individual based on deviation from a normative data set, based on age group and gender. The summary score for each symptom domain is comprised of 17 parameters, five activity (time active, distance travelled, score area, number of micro-events, motion simplicity) and 12 CPT (RT, score outliers, RT variation, normalised variation, omission errors, commission errors, normalised commission errors, anticipatory responses, multi-responses, D-Prime modified, longest passivity, total error rate). There are two versions of the task for children and young people; the task for 6–11-year olds is 15 minutes in duration and the task for 12–17-year olds is 20 minutes in duration. The QbTest result is supported by a behavioural observation of events that may affect test performance. The QbTest has been approved by the US Food and Drug Administration (FDA) to supplement standard clinical assessment and treatment monitoring. To the best of our knowledge the QbTest is the only test cleared for both assessment and treatment.

One study investigated the clinical utility of the QbTest to aid assessment. Vogt & Shameli [41] compared practice outcomes of two groups of children; one group had their assessment without QbTest and one group had QbTest added to standard clinical assessment. The outcomes were compared 1-year later and the results showed that seven participants (37%) in the non-QbTest had their diagnosis revised (from non-ADHD to ADHD) 1-year later, in comparison to none in the QbTest group. Clinician's decision to initially reject an ADHD diagnosis was due to a lack of pervasiveness in the child's developmental history, or attribution to emotional disorder. Of the 19 participants who had a QbTest and did not receive an ADHD diagnosis, scores for all three symptom domains (attention, activity, impulsivity) were not outside the normal range. Given the clinical difficulty in differentiating ASD and ADHD, it is particularly interesting to note that five participants with normal attention and impulse scores that were not diagnosed with ADHD received a diagnosis of ASD. The study demonstrates that the QbTest improves diagnostic accuracy and stability, and the addition of the activity measure may be particularly useful in distinguishing between disorders with sensory over-responsivity (such as ASD).

In summary, the papers show mixed findings on the clinical utility of CPTs to aid diagnosis. The sensitivity and specificity of the tests to aid group differentiation varies across papers, however, it is important to note that the CPT should not be used as a stand-alone tool for diagnosis. Across the CPTs, studies reported the utility of including a CPT as an objective measure of symptoms that allow the clinician to directly observe behaviour. However, some studies report the CPT cannot accurately differentiate ADHD in clinic referred samples. Additionally, several studies report little association between CPTs and rating scales, some studies have interpreted this as a lack of validity of the CPT, whereas others suggest the CPT may be tapping into a unique factor of ADHD not assessed in rating scales. Further research is required to establish the clinical utility of CPTs in aiding the assessment process in clinical practice.

The clinical use of CPTs to aid medication management in ADHD

Twelve papers looked at the clinical utility of CPTs to monitor medication in ADHD, from which five CPTs were identified, four papers used TOVA [42-45], three papers used CCPT [46,18,47] two papers used MMAT [48,21], two GDS [49,31], and one QbTest [13]. A summary of the papers is presented in Table 3.

<<Insert Table 3 about here>>

Test of Variables of Attention (TOVA)

The four studies that used the TOVA collectively found mixed reports on its clinical utility to monitor medication. Three papers reported that only impulsivity (commission scores) significantly improved on TOVA after Methylphendiate (MPH), two papers showed this when looking at the prolonged use of MPH [42], and one paper when assessing acute responses [44]. Aggarwal and Lillystone [42] suggest that TOVA as a means to assess medication therapy requires more investigation and the test would benefit from a cut-off score that can be used to support stopping medication. Huang et al. [44] report that when assessing the acute effects of MPH, the second half of TOVA was more sensitive in assessing medication response, which may suggest this half of the test should be primarily used to assess medication response. Alternatively, given that the second half of the test occurred later than the first half, it may be that in clinical practice conducting the TOVA one and a half hours after medication is more effective than one hour or that there is a decrement in sustained attention during the course of the test. Wang et al. [45]) showed that during a 24-month period of MPH treatment behavioural ratings of ADHD improved, but there was limited correlation between TOVA scores and behavioural ratings. Teacher-rated inattention was correlated with TOVA omission errors and the overall ADHD score. Parent-rated inattention was correlated with TOVA commission. Measures of psychiatrist rated inattention and hyperactivity and impulsivity rated by parent/teacher/psychiatrist did not correlate with TOVA. On the basis of these findings the authors recommend getting information from multiple informants to establish medication effects.

The fourth paper reports that an objective assessment of medication response is required in children. Manor et al. [43] compared children's subjective reports (Clinical Global Impression of Change) and TOVA scores on functioning after taking MPH. They found children (particularly those under 10 years of age) were unable to accurately assess their symptoms and improvement, thus conclude including an objective test prior to medication initiation and during titration is a valid addition to ADHD monitoring.

Gordon's Diagnostic System (GDS)

Two papers which reported on the GDS describe mixed findings on the use of the tool to aid medication monitoring. Gordon[31] reported the benefit of the GDS in monitoring medication in one case study. Tests off medication showed clear difficulties on both the Delay and Vigilance Task, but whilst on medication his scores were in the normal or near-normal range. Based on this, the young person was recommended to continue stimulant medication therapy and the author concluded that the task was useful in gaining direct observations of this child's behaviour.

In contrast, the findings of Fischer and Newby [49] suggest that the tool is not useful for monitoring early medication responses. The authors describe a multi-method protocol implemented in an RCT investigating individual children's responses to medication in clinic. Children sat weekly clinic tests (including the GDS Vigilance Task) with the addition of parent and teacher rating scales during titration. They found that GDS scores only improved on the highest dose of MPH, however, differences on lower doses were found on other tests (see Table 3 for tests), suggesting that the GDS would not be a valuable tool to aid titration.

Conner's CPT (CCPT)

Three papers support the use of CCPT to aid medication management. Fernández-Jaén et al. [46] aimed to assess the efficacy of MPH-extended release (ER). They compared the performance of two groups on a suite of attention rating scales and the CCPT. One group sat the tests at baseline (prior to treatment onset) and again 3-months later, and three hours after taking MPH-ER (group A). For the second group, the tests were conducted after achieving optimal MPH-ER dose (three hours after taking it) and repeated approximately one month later, with at least a 48 hours suspension in medication before the repeat test (group B). The results showed a significant improvement in CCPT scores after MPH-ER. Furthermore, the CCPT was the only test which changed significantly in patients in both group A and B. The authors concluded the CCPT was useful in determining treatment efficacy, particularly the short-term effects of treatment.

Wang et al. [18] used CCPT to monitor effects of MPH, with children sitting the CCPT at baseline, one month after baseline (after MPH initiation) and three and six months later. They found during the six month period, CCPT scores on impulsivity, hyperactivity and inattention improved, but CCPT assessed distraction did not, suggesting this score is not useful for monitoring ADHD in clinical settings.

Bédard et al. [47] assessed whether ATX and MPH improved performance on CCPT, after patients had been titrated to an optimum dose over a period of 4-6 weeks. They found that medication improved measures of attention on the CCPT, but not inhibitory control, however, the changes in attention scores for patients on ATX were not significant. Scores on the CCPT did not correlate with symptomatic improvement assessed via rating scale (ADHD-RS). On the basis of these findings the authors conclude that CCPT does not reflect ADHD symptom change, and that using CCPT to titrate mediation may lead to suboptimal outcomes. Together, the research reports mixed findings on the clinical utility of the CCPT to aid medication management.

QbTest

Vogt and Williams [13] investigated MPH treatment response after a single dose using the QbTest in routine clinical practice. They found the QbTest to be sensitive to the effects of medication, with significant effects found for all activity and attention variables. Robust treatment effects were reported in 84% of patients, 7% showed a partial response and 9% were identified as non-responders based on measures of activity, attention and impulsivity. They conclude that adding the QbTest, a CPT that combines activity measurement to routine clinical practice, provides important information on behavioural and neuropsychological responses to treatment.

Quotient ADHD System/ MMAT (formerly known as the McLean Motion and Attention Test (MMAT and OPTAx)

The Quotient ADHD system or MMAT takes approximately 15 minutes to complete and combines a CPT based on Greenberg's Minnesota Computer Assessment task [50] with an infrared motion analysis system (Qualysus, Gothenberg, Sweden). During the CPT, participants are asked to respond to target shapes by pressing a space bar, and inhibit response to non-targets. Unlike the QbTest, the stimuli are presented at random locations on the screen. Half the presented stimuli are targets. Stimuli are presented on the screen for 200ms with a random interval between stimuli presentation. The CPT measures attention, accuracy (percentage of correct responses), omission and commission errors and latency, variability and coefficient variation in responding. The infrared motion analysis tracks the location of reflective marker worn on a cap on the participants head. The infrared system collects and records movement at 50/sec with a resolution of 0.04mm. A report is produced that graphs the participant's performance on the test. The test is commercially available as the 'Quotient ADHD system', however, published reports refer to its previous name 'MMAT'. The test has been cleared by the FDA to support ADHD assessment.

Two papers report on the use of MMAT in aiding medication management, both report it to be a clinically useful tool. Teicher et al. [48] investigated whether MMAT could aid the titration process. Participants took part in a four-week triple-blind treatment trial, whereby they received one-week of treatment with placebo, and one-week of low, medium and high MPH doses respectively. At the end of each treatment week the MMAT was conducted and parents completed the Clinical Global Improvement (CGI) scale at each time point. Results showed that in 9/11 (81.8%) cases MMAT identified the dosage parents perceived to be most beneficial, suggesting the utility of this tool in monitoring medication. The authors state that MMAT can accurately assess a response to 0.4mg/kg dose of MPH and at the time of publishing, they were using this to start the titration process in clinic. Tabori-Kraft et al. [21] aimed to evaluate the utility of MMAT in routine medication management. Doses were optimised during the first 6-weeks of treatment based on parent and teacher reports and clinical observation. After optimisation the child sat the MMAT, before and after stimulant uptake. They found that children were less active on stimulant medication and based on CGI scores, 95% of participants were reported to be "much" or "very much" improved. The authors conclude that the MMAT should be used in medication management to support subjective ratings of improvement. They also note that children were pleased with their MMAT report and improvement on and off medication, which may encourage treatment adherence.

In summary, the clinical utility of CPTs in aiding medication management in clinical practice is under investigated. Most research has been conducted on TOVA which shows limited sensitivity to medication effects, with only impulsivity scores improving as a result of medication. However, the QbTest and MMAT which combine a measure of activity alongside the domains in other CPTs (attention and impulsivity) have been shown to be clinically useful in supporting titration. Combined, the papers demonstrate the need for an objective measure of medication response, which may be particularly useful when the CPT combines an activity parameter (such as the QbTest and MMAT) to measure all three symptom domains of ADHD.

Objective measures of activity in ADHD

As demonstrated, there has been some evidence investigating the clinical utility of CPTs in aiding ADHD assessment and monitoring. However, traditionally CPTs only measure two of the core symptom domains of ADHD attention and impulsivity). Our search revealed that more recently two CPTs have been

developed that additionally provide an objective measurement of activity (QbTest and MMAT). Yet in comparison to objective measures of attention and impulsivity the objective measurement of activity has been under-investigated. Therefore we conducted an additional search with the aim of identifying objective methods of assessing activity in ADHD. Given the comparative lack of research in objective measures of activity, we did not limit our search to papers that used the measures to clinically aid assessment or titration *per se*, but only include papers that discuss the relevance of activity measures to differentiate between ADHD and non-ADHD groups or medication effects. Where papers also met our first search criteria, their findings are not repeated below. This is the case for four papers [21,51,41,13].

Our search revealed 25 papers and two measures for objectively assessing activity in ADHD: accelerometer-based devices (actigraphy and inertial measurement units (IMUs)) and infra-red motion analysis (MMAT and QbTest). A summary of the papers is presented in Table 4 and Table 5. An additional paper [52] reported a review on actigraphy and motion analysis, this was excluded from our review as not all reported papers were on ADHD, however, the relevant articles presented in this review were checked and identified in our search.

Accelerometer-based devices

The actigraph was the most commonly cited method of measuring activity. An actigraph measures motor activity through an accelerometer. Of the actigraphy studies, ten papers looked at the role of objectively measuring activity in diagnosing ADHD or differentiating it from controls [53-62] and six papers objectively measured activity and medication [63-68]. Three papers reported using an actigraph which was worn either on the participant's wrist or ankle [53,54,64], ten report using an actigraph worn around the patient's middle [65,67,58,56,59,57,63,60,61,66] and two report use a combination of an actigraph worn on the patient's leg and around their middle [55,62]; one reports a meta-analysis review[68].

Porrino et al.[58] was the earliest paper to show that children with ADHD experienced greater levels of activity than children without ADHD during a structured task. One paper used actigraphy to measure children's activity whilst they attended a full-day ADHD clinical diagnostic assessment [54]. They found no difference in the activity levels between those who were and were not diagnosed with ADHD in the morning sessions and no difference in activity between the ADHD sub-types. However, in the afternoon session, children who were diagnosed with ADHD were significantly more active than those without ADHD, suggesting some sensitivity of this measure to ADHD, but also highlighting the potential influence on CPT scores depending on the time of day. Two papers used actigraphy whilst patients performed laboratory tests (stop-signal task, choice-task and cognitive-experimental tasks) and found recorded levels of activity could differentiate children with ADHD from controls [53,55], with Wood et al. [55] reporting ROC area under cure (AUC) values of up to 0.8 for actigraphy, differentiating between ADHD, their siblings and controls. Five papers used an actigraph whilst the patients sat a non-commercially available CPT [56,57,59-61]. Halperin et al. [56] found that only activity could differentiate the ADHD group from non-ADHD psychiatric controls, but could not differentiate ADHD from children with anxiety disorders [59]. Rajendran et al. [62] assessed whether scores on NPESY, actigraphy and a non-commercial CPT could identify preschool children at risk of ADHD when they were school ages. Typically developing and Hyperactive/Inattentive (HI) subgroups (assessed via a rating scale) of 3-4 year olds were

assessed for 6 years. Using latent profile analysis, they found that the nature/severity of dysfunction in HI pre-schoolers did not predict later ADHD diagnosis. All measures could significantly differentiate typically developing from HI preschoolers, yet in the HI only preschool actigraphy measures could predict ADHD outcome 4 and 5 years later.

Hall et al. [60] found children with ADHD were more impulsive (as measured by CPT) and more active (measured through an actigraph) during the CPT. Marks et al. [61] found that CPT measured inattention correlated with activity. Innoue et al. [57] found that activity measurement could differentiate ADHD from healthy controls with sensitivity and specificity of 65% and 76% respectively, additionally combining actigraphy with CPT increased the accuracy of ADHD diagnosis than a standard CPT alone. The findings suggest the utility of adding an objective measure of activity to a standard CPT to aid assessment.

Six papers looked at the role of actigraphy in determining medication effects [65,63,64,67,66]. Porrino et al. [67] found a single dose of Dextroamphetamine (Dex) reduced activity by 28%, similarly Donnelly et al. [66] also found a reduction in activity after Dex. Borcherding et al. [65] reported that MPH resulted in a greater reduction in motor activity than Dex, as measured by actigraphy. De Crescenzo et al.[68] conducted a meta-analytic review focussing on the role of actigraphy in detecting changes in activity and sleep patterns in RCTs investigating the effect of MPH. They conclude that actigraphy was able to assess the effect of MPH on sleep and activity, which could be used to support clinical diagnosis and treatment follow-up. Rapoport et al. [63] compared activity of hyperactive boys on Dex to a placebo during a gym class. They found that boys receiving Dex were less active than boys on the placebo and decrease in activity correlated with teacher reported improved behaviour, suggesting that stimulant medication affects activity regardless of the test situation. Konrad et al. [64] found a linear effect of MPH dose on actigraphy scores and a relationship between actigraphy scores obtained at school and in a neuropsychological test, suggesting that measuring activity in a test situation can be a clinically useful way of observing hyperactivity. They also found changes in teacher-rated hyperactive-impulsive symptoms and inattentive symptoms could be explained by changes in actigraphy scores, highlighting the need to include objective measures of activity. Combined, the papers suggest that actigraphy is a valid tool to assess medication response.

O'Mahony et al.[69] investigated the use of IMUs to discriminate ADHD and non-ADHD participants. The IMUs were worn on the waist and dominant ankle across five context (in the waiting room with parent, in the waiting room with supervisor only, in the consultants room with psychiatrist, in consultants room with psychiatrist and parent, and taking the TOVA). Using a support vector machine to classify participants as ADHD or non-ADHD, sensitivity of IMU was 94.44% and 95.65%. Specifically, analyses of motion during the CPT (TOVA) provided a better performance classification than that achieved during free time. The findings support the use of an activity measure during CPT to aid diagnostic assessment.

Infra-red motion analysis

Of the infra-red motion analysis studies, two papers report the use of a specifically designed CPT and combined infra-red motion tracker to investigate group differences [51,70]. Teicher et al. [51] used an early version of the MMAT system and found children with ADHD moved significantly more than children without-

ADHD during the CPT, with head movement patterns significantly correlating with teacher ratings of activity, this combined CPT and infra-red motion analysis differentiated ADHD children from normal controls with 88.9% sensitivity and 100% specificity. Reh et al. [70] looked at the factor structure of the QbTest and found QbTest scores on activity significantly correlated with teacher ratings of hyperactive behaviours, their analysis supported a three-factor model (attention, impulsivity and activity) that was able to explain 76% of variance within the sample of ADHD children, with the activity parameter explaining the largest amount of variance. The authors conclude the addition of a measurement of activity combined with the CPT constitutes a significant advantage for QbTest to aid differentiation of ADHD and non-ADHD cases.

Six further studies looked at the effects of medication on combined CPTs and infra-red motion analysis [71-76]. Heiser [71] used MMAT to assess the effect of MPH on ADHD symptoms and found a significant improvement on activity parameters after MPH, suggesting MMAT is a useful measure of medication response. Wehmeier et al. [74] investigated the efficacy of Atomextine (ATX) on executive functioning. Participants were randomised into receiving eight weeks of titrated ATX or placebo. The QbTest was administered at baseline and at five points during subsequent weeks (during weeks 1, 2, 4, 6 and 8) and sat in the morning, at noon and late noon/early evening. They found that both groups experienced a circadian pattern of outcomes during the day, peaking at 10am, declining and then peaking again at 5pm before then declining. ATX was also shown to have a positive effect on activity ratings. The sensitivity of the activity measure to ATX was also reported in their subsequent paper [72]. Gunther et al. [75] used QbTest to compare long-acting MPH with immediate release, and found similar levels of activity between the two groups; the authors conclude that it is important that all three symptom domains of ADHD are assessed during titration and medication initiation in order to get the most optimal formulation for each individual. Together, these papers support the use of infra-red motion analysis to aid differentiation between ADHD and non-ADHD children and to investigate the effect of medication.

Ramtvedt and Sundet [73] compared scores on the QbTest with behavioural ratings in children on MPH, Dex and a placebo. They found the QbTest scores improved after MPH and Dex but not after placebo, with the activity measure having the largest effect size. Additionally, although convergent validity for rating scales and the QbTest activity were found on a group level, the measures were not equivalent on an individual level for each child. In a paper from the same trial, Ramtvedt et al. [76] used QbTest to investigate the clinical gains of including both Dex and MPH in stimulant trials but do not report specifically on activity scores. Their findings support the inclusion of the QbTest to monitor medication, which can provide additional information to that assessed in rating scales.

In summary, objectively assessed activity appears to be sensitive to group differences (ADHD vs. non-ADHD) and medication effects, suggesting it may be a valuable tool in aiding both the assessment and medication management.

Discussion

Given the continuing and increasing use of CPTs to aid clinical practice in ADHD, this review sought to bring together the current published data on the clinical utility of the CPT in supporting assessment and

medication management in children referred or diagnosed with ADHD. This is the first review to focus on the clinical application of these tests, and in doing so we can reflect on the current evidence base and identify strengths and weaknesses of the CPT as a clinical tool to inform areas for future investigation. A secondary aim of this review was to investigate the evidence supporting the objective measure of activity in ADHD.

In total, our search revealed six commercially available CPTs that had been investigated for their clinical utility (TOVA, GDS, CCPT, MMAT, IVA+CPT and QbTest), of which the TOVA and GDS had the largest evidence base for its clinical utility. Together, the evidence indicates mixed findings on the utility of the CPT as an assessment tool. Some evidence reported the CPT could not differentiate ADHD from non-ADHD in a clinic referred sample [26,23,14], suggesting the tool has limited utility in clinical practice. There were mixed findings on the sensitivity and specificity of the CPTs in predicting group membership, however, it is important to remember that the CPT should not be used as a stand-alone diagnostic tool, but interpreted within the clinical context. Additionally, some studies reported little association between CPT scores and parent/teacher rating scales [14,29,16]. However, the interpretation of this result differed between studies. Whereas some questioned the validity of the CPT, others suggested it may be tapping into unique aspects of ADHD not assessed in rating scales. It is difficult to ascertain which may be the correct interpretation, and further investigation is required to compare independently rated clinical diagnosis with clinical diagnosis obtained using CPT scores. One study [41] combined a CPT with an objective measure of activity (QbTest) to aid assessment, and found the tool improved the clinician's ability to accurately diagnose ADHD. On the basis of one study it is unwise to draw firm conclusions; however, it appears that combining CPTs with objective measures of activity may be significantly beneficial.

With regard to medication management, the majority of research has focussed on the TOVA, of which only scores on impulsivity were shown to improve as a result of medication. A lack of research in this area was noted from the review and further research is required to fully understand the utility of CPTs to aid medication management. However, CPTs were noted to be useful in providing an objective report in medication response, which can be difficult to assess in children who find it hard to accurately report their symptoms and improvements [43]. Additionally, CPTs that compute a report documenting symptom changes (such as QbTest and MMAT) may increase the child's motivation and encourage treatment adherence [21]. Again, the papers that combined a CPT with motion detection (QbTest and MMAT) were the most clinically useful in investigating medication effects, even being sensitive to partial responses and responses on low medication doses [13,48] suggesting their utility to aid titration in clinical practice.

The secondary search on objective measures of activity in ADHD revealed two methods to currently assess activity: accelerometer based devices and infra-red motion analysis combined with a CPT (QbTest and MMAT). Both methods were reported to be useful in aiding the differentiation of ADHD/non-ADHD and assessing the effects of medication. To date, QbTest is the only CPT that has FDA clearance to aid both the assessment and treatment monitoring phase of ADHD in children and young people. Clearly, activity is only one of three core symptom domains of ADHD; however, it is currently over-looked in the majority of laboratory tests of ADHD. Despite this, the evidence suggests that combining an objective assessment of activity with a CPT assessing impulsivity and attention may provide the most clinically useful tool in aiding assessment and medication management. It should be noted that no CPT or assessment tool of ADHD is designed to be used as

the sole basis for an ADHD diagnosis. However, the addition of such a test to augment clinical practice has some supporting evidence base, that is worthy of further investigation, particularly when combined with a measure of activity.

Our review noted a lack of research investigating the clinical utility of these measures, with most research focussing on their psychometric properties. Whilst this is undoubtedly important, for CPTs to be validated as a clinical tool, it is necessary to further understand how they aid clinical practice and in which part of the assessment or management pathway they may be best placed to optimise patient outcome in a manner that is time and cost-effective. We specifically note a lack of randomised controlled trials (RCT) to support the clinical use of these measures and the majority of research focuses on group differences, however, in clinical practice, the primary concern is diagnostic accuracy for an individual patient. Future RCTs are required to ascertain the clinical utility of CPTs in aiding diagnostic accuracy for individual patients referred for an ADHD assessment.

The quality of this review and its findings are dependent on the quality of the papers assessed. Given the lack of trials investigating the clinical utility of the CPT we did not restrict our data to RCT evidence only. Additionally, we did not undertake any quality assessment of the papers and their findings. The papers included in this review incorporate samples of clinic-referred and community-based samples, which may have influenced the clinical utility of the CPT. For example; it could be argued that the ability to differentiate ADHD/non-ADHD is easier in a community sample than a clinic sample. However, we could not ascertain if this was the case given the limited number of community-based samples in the review. In addition, some of the studies used specific samples, such as gifted children [16] or children with language disorders [27] which limits the generalisability of that papers findings, but is nonetheless relevant to the understanding of the use of the CPT in clinical practice. Furthermore, we chose to limit our search to only papers that were available in English and thus may have excluded other relevant papers reported in a different language. However, this review is unique in its focus on the use of objectively measured attention, impulsivity and activity to aid clinical practice and provides an important source for future research, healthcare professionals and service providers.

In conclusion, published studies support that CPTs provide an objective method to assess attention and impulsivity, but there are mixed reports on whether they are a useful adjunct to clinical practice. The clinical utility of these tests to aid assessment and medication requires further investigation, with a particular need for RCTs investigating whether the addition of CPTs to standard practice can streamline the ADHD pathway and whether they can facilitate speedier, more accurate diagnosis and subsequent cost-savings and improved patient outcome. Adding an objective measure of activity allows the CPT to measure all three core symptom domains of ADHD and initial evidence suggests this may improve the clinical utility of the CPT.

Acknowledgements

The study was funded by the National Institute of Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC EM). The research reported in this paper was also conducted by the NIHR MindTech Healthcare Technology Co-operative. This work has been funded by the National Institute for Health Research. The views represented are the views of the authors alone and do not necessarily represent the views of the Department of Health in England, NHS, or the National Institute for Health Research. The funding body had no role in the design, collection, analysis and interpretation of data, or in the writing of the manuscript.

Declaration of Conflict and Competing Interests

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical Standards

The paper is not a human or animal study so ethical approval was not required. However, all work was undertaken under the auspices of the AQUA-Trial, which has ethical approval.

References

1.NICE (2008) Attention deficit hyperactivity disorder:diagnosis and managment of ADHD in children, young people and adults. *Clinical Guideline 72*. In:National Institute for Health and Clinical Excellence London

2. Magyary D, Brandt P (2002) A decision tree and clinical paths for the assessment and management of children with ADHD. Issues in Mental Health Nursing 23:553-566

3. Faraone SV, Biederman J (1998) Neurobiology of attention-deficit hyperactivity disorder. Biol Psychiat 44:951-958

4. Ogundele MO, Ayyash HF, Banerjee S (2011) Role of computerised continuous performance task tests in ADHD. Prog Neural Psychiatr 15:8-13

5. Weinberg WA, Brumback RA (1992) The myth of attention deficit-hyperactivity disorder: Symptoms resulting from multiple causes. J of Child Neurol 7:431-445

6. Hall CL, Walker GM, Valentine AZ, Guo B, Kaylor-Hughes C, James M, Daley D, Sayal K, Hollis C (2014) Protocol investigating the clinical utility of an objective measure of activity and attention (QbTest) on diagnostic and treatment decision-making in children and young people with ADHD—'Assessing QbTest Utility in ADHD'(AQUA): a randomised controlled trial. BMJ Open 4:e006838

7. Epstein JN, Erkanli A, Conners CK, Klaric J, Costello JE, Angold A (2003) Relations between Continuous Performance Test performance measures and ADHD behaviors. J Abnorm Child Psych 31:543-554

8. Losier BJ, McGrath PJ, Klein RM (1996) Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: a meta-analytic review. J Child Psychol Psyc 37:971-987

9. Riccio CA, Waldrop JJ, Reynolds CR, Lowe P (2001) Effects of stimulants on the continuous performance test (CPT): implications for CPT use and interpretation. J Neuropsych Clin N 13:326-335

10. Dupuy T, Greenberg L (1993) TOVA manual for test of variables of attention computer program. University of Minnesota, Minneapolis

11. Conners C (1995) The Conners Rating Scales: Instruments for the assessment of childhood psychopathology. Duke University: Durham, USA

12. Gordon M (1982) The Gordon diagnostic system. Gordon Systems, New York

13. Vogt C, Williams T (2011) Early identification of stimulant treatment responders, partial responders and non-responders using objective measures in children and adolescents with hyperkinetic disorder. Child and Adolescent Mental Health 16:144-149

14. McGee RA, Clark SE, Symons DK (2000) Does the Conners' Continuous Performance Test aid in ADHD diagnosis? J Abnorm Child Psych 28:415-424

15. Lis S, Baer N, Stein-en-Nosse C, Gallhofer B, Sammer G, Kirsch P (2010) Objective measurement of motor activity during cognitive performance in adults with attention-deficit/hyperactivity disorder. Acta Psychiat Scan 122:285-294

16. Chae PK, Kim J-H, Noh K-S (2003) Diagnosis of ADHD among gifted children in relation to KEDI-WISC and T.O.V.A. performance. Gifted Child Quart 47:192-201

17. Mayes S, Calhoun L, Crowell EW (2001) Clinical validity and interpretation of the Gordon Diagnostic System in ADHD assessments. Child Neuropsychol 7:32-41

18. Wang L-J, Huang Y-S, Chiang Y-L, Hsiao C-C, Shang Z-Y, Chen C-K (2011) Clinical symptoms and performance on the Continuous Performance Test in children with attention deficit hyperactivity disorder between subtypes: a natural follow-up study for 6 months. BMC Psychiatry 11:65

19. Kim J, Lee Y, Han D, Min K, Kim D, Lee C (2015) The utility of quantitative electroencephalography and Integrated Visual and Auditory Continuous Performance Test as auxiliary tools for the Attention Deficit Hyperactivity Disorder diagnosis. Clinical Neurophysiology 126:532-540

20. Ulberstad F (2012) QbTest Technical Manual. Stockholm, Sweden: Qbtech AB

21. Tabori-Kraft J, Sørensen MJ, Kærgaard M, Dalsgaard S, Thomsen PH (2007) Is OPTAx™ useful for monitoring the effect of stimulants on hyperactivity and inattention? Eur Child Adolesc Psy 16:347-351

22. Porumb M (2007) Using T.O.V.A. for the assessment of ADHD: A case study. Cognitie Creier Comportament 11:571-584

23. Preston AS, Fennell EB, Bussing R (2005) Utility of a CPT in diagnosing ADHD among a representative sample of high-risk children: a cautionary study. Child Neuropsychol 11:459-469

24. Schatz AM, Ballantyne AO, Trauner DA (2001) Sensitivity and specificity of a computerized test of attention in the diagnosis of attention-deficit/hyperactivity disorder. Assessment 8:357-365

25. Wada N, Yamashita Y, Matsuishi T, Ohtani Y, Kato H (2000) The test of variables of attention (TOVA) is useful in the diagnosis of Japanese male children with attention deficit hyperactivity disorder. Brain & Development 22:378-382

26. Forbes GB (1998) Clinical utility of the Test of Variables of Attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. J Clin Psychol 54:461-476

27. Rielly NE, Cunningham CE, Richards JE, Elbard HJ, MAhoney WJ (1999) Detecting Attention Deficit Hyperactivity Disorder in a Communications Clinic: Diagnostic Utility of the Gordon Diagnostic System. J Clin Exp Neuropsyc 21:685-700

28. Rielly NE, Cunningham CE, Richards JE, Elbard H, Mahoney WJ (2000) "Detecting attention deficit hyperactivity disorder in a communications clinic: Diagnostic utility of the Gordon Diagnostic System": Erratum. J Clin Exp Neuropsyc 22:435

29. DuPaul GJ, Anastopoulos AD, Shelton TL, Guevremont DC, Metevia L (1992) Multimethod assessment of attention-deficit hyperactivity disorder: The diagnostic utility of clinic-based tests. J Clin Child Psychol 21:394-402

30. Maye Mayes S, Calhoun SL (2002) The Gordon Diagnostic System and WISC-III Freedom from Distractibility Index: validity in identifying clinic-referred children with and without ADHD. Psychol Rep 91:575-587

31. Gordon M (1986) How is a computerized attention test used in the diagnosis of attention deficit disorder? Journal of Children in Contemporary Society 19:53-64

32. El-Sayed E, van't Hooft I, Fried I, Larsson JO, Malmberg K, Rydelius PA (1999) Measurements of attention deficits and impulsivity: a Swedish study of the Gordon Diagnostic System. Acta Paediatr 88:1262-1268

33. Mayes SD, Calhoun SL, Chase GA, Mink DM, Stagg RE (2009) ADHD subtypes and co-occurring anxiety, depression, and oppositional-defiant disorder: Differences in Gordon Diagnostic System and Wechsler Working Memory and Processing Speed Index scores. Journal of Attention Disorders 12:540-550

34. Trommer BL, Hoeppner J-AB, Lorber R, Armstrong K (1988) Pitfalls in the use of a continuous performance test as a diagnostic tool in attention deficit disorder. J Dev Behav Pediatr 9:339-345

35. Matier-Sharma K, Perachio N, Newcorn JH, Sharma V, Halperin JM (1995) Differential diagnosis of ADHD: Are objective measures of attention, impulsivity, and activity level helpful? Child Neuropsychol 1:118-127

36. Barkley RA, Grodzinsky G (1994) Are tests of frontal lobe functions useful in the diagnosis of Attention Deficit Disorders? The Clinical Neuropsychologist 8:121-139

37. Grodzinsky GM, Barkley RA (1999) Predictive power of frontal lobe tests in the diagnosis of attention deficit hyperactivity disorder. The Clinical Neuropsychologist 13:12-21

38. Alloway TP, Gathercole SE, Holmes J, Place M, Elliott JG, Hilton K (2009) The diagnostic utility of behavioral checklists in identifying children with ADHD and children with working memory deficits. Child Psychiat Hum D 40:353-366

39. Perugini EM, Harvey EA, Lovejoy DW, Sandstrom K, Webb AH (2000) The predictive power of combined neuropsychological measures for attention-deficit/hyperactivity disorder in children. Child Neuropsychol 6:101-114

40. Kleinman A, Caetano SC, Brentani H, de Almeida Rocca CC, Dos Santos B, Andrade ER, Zeni CP, Tramontina S, Rohde LAP, Lafer B (2015) Attention-based classification pattern, a research domain criteria framework, in youths with bipolar disorder and attention-deficit/hyperactivity disorder. Aust Nz J Psychiat doi: 10.1177/0004867414557957

41. Vogt C, Shameli A (2011) Assessments for attention-deficit hyperactivity disorder: Use of objective measurements. The Psychiatrist 35:380-383

42. Aggarwal A, Lillystone D (2000) A follow-up pilot study of objective measures in children with attention deficit hyperactivity disorder. J Paediatr Child H 36:134-138

43. Manor I, Meidad S, Zalsman G, Zemishlany Z, Tyano S, Weizman A (2008) Objective versus subjective assessment of methylphenidate response. Child Psychiat Hum D 39:273-282

44. Huang Y-S, Chao C-C, Wu Y-Y, Chen Y-Y, Chen C-K (2007) Acute effects of methylphenidate on performance during the Test of Variables of Attention in children with attention deficit/hyperactivity disorder. Psychiat Clin Neuros 61:219-225

45. Wang L-J, Chen C-K, Huang Y-S (2015) Neurocognitive performance and behavioral symptoms in patients with attention-deficit/hyperactivity disorder during twenty-four months of treatment with methylphenidate. J Child Adol Psychop 25:246-253

46. Fernandez-Jaen A, Fernandez-Mayoralas DM, Pardos A, Calleja-Perez B, Munoz Jareno N (2009) Clinical and cognitive response to extended-release methylphenidate (Medikinet) in attention deficit/hyperactivity disorder: efficacy evaluation. Advances in Therapy 26:1097-1110

47. Bédard ACV, Stein MA, Halperin JM, Krone B, Rajwan E, Newcorn JH (2015) Differential impact of methylphenidate and atomoxetine on sustained attention in youth with attention-deficit/hyperactivity disorder. J Child Psychol Psyc 56:40-48

48. Teicher MH, Polcari A, McGreenery CE (2008) Utility of objective measures of activity and attention in the assessment of therapeutic response to stimulants in children with attention-deficit/hyperactivity disorder. J Child Adol Psychop 18:265-270

49. Fischer M (1991) Assessment of stimulant response in ADHD children using a refined multimethod clinical protocol: Erratum. J Clin Child Psychol 25:121

50. Greenberg L (1986) An objective measure of methylphenidate response: clinical use of the MCA. Psychopharmacol Bull 23:279-282

51. Teicher MH, Ito Y, Glod CA, Barber NI (1996) Objective measurement of hyperactivity and attentional problems in ADHD. J Am Acad Child Psy 35:334-342

52. Teicher MH (1995) Actigraphy and motion analysis: new tools for psychiatry. Harvard Review of Psychiatry 3:18-35

53. Alderson RM, Rapport MD, Kasper LJ, Sarver DE, Kofler MJ (2012) Hyperactivity in boys with attention deficit/hyperactivity disorder (ADHD): the association between deficient behavioral inhibition, attentional processes, and objectively measured activity. Child Neuropsychol 18:487-505

54. Dane AV, Schachar RJ, Tannock R (2000) Does actigraphy differentiate ADHD subtypes in a clinical research setting? J Am Acad 39:752-760

55. Wood AC, Asherson P, Rijsdijk F, Kuntsi J (2009) Is overactivity a core feature in ADHD? Familial and receiver operating characteristic curve analysis of mechanically assessed activity level. J Am Acad Child Psy 48:1023-1030

56. Halperin JM, Matier K, Bedi G, Sharma V, Newcorn JH (1992) Specificity of inattention, impulsivity, and hyperactivity to the diagnosis of attention-deficit hyperactivity disorder. J Am Acad Child Psy 31:190-196

57. Inoue K, Nadaoka T, Oiji A, Morioka Y, Totsuka S, Kanbayashi Y, Hukui T (1998) Clinical evaluation of attention-deficit hyperactivity disorder by objective quantitative measures. Child Psychiat Hum D 28:179-188

58. Porrino LJ, Rapoport JL, Behar D, Sceery W, Ismond DR, Bunney WE, Jr. (1983) A naturalistic assessment of the motor activity of hyperactive boys. I. Comparison with normal controls. Arch Gen Psychiat 40:681-687

59. Halperin JM, Newcorn JH, Matier K, Sharma V, McKAY KE, Schwartz S (1993) Discriminant validity of attention-deficit hyperactivity disorder. J Am Acad Child Psy 32:1038-1043

60. Hall S, Halperin J, Schwartz S, Newcorn J (1997) Behavioral and executive functions in children with attention-deficit hyperactivity disorder and reading disability. Journal of Attention Disorders 1:235-243

61. Marks DJ, Himelstein J, Newcorn JH, Halperin JM (1999) Identification of AD/HD subtypes using laboratory-based measures: A cluster analysis. J Abnorm Child Psych 27:167-175

62. Rajendran K, O'Neill S, Marks DJ, Halperin JM (2015) Latent profile analysis of neuropsychological measures to determine preschoolers' risk for ADHD. J Child Psychol Psyc doi: 10.1111/jcpp.12434

63. Rapoport JL, Tepsic PN, Grice J, Johnson C, Langer D (1980) Decreased motor activity of hyperactive children on dextroamphetamine during active gym program. Psychiat Res 2:225-229

64. Konrad K, Gunther T, Heinzel-Gutenbrunner M, Herpertz-Dahlmann B (2005) Clinical evaluation of subjective and objective changes in motor activity and attention in children with attention-deficit/hyperactivity disorder in a double-blind methylphenidate trial. J Child Adol Psycho 15:180-190

65. Borcherding BG, Keysor CS, Cooper TB, Rapoport JL (1989) Differential effects of methylphenidate and dextroamphetamine on the motor activity level of hyperactive children. Neuropsychopharmacol 2:255-263

66. Donnelly M, Rapoport JL, Potter WZ, Oliver J, Keysor CS, Murphy DL (1989) Fenfluramine and dextroamphetamine treatment of childhood hyperactivity: Clinical and biochemical findings. Arch Gen Psychiat 46:205-212

67. Porrino LJ, Rapoport JL, Behar D, Ismond DR, Bunney WE (1983) A naturalistic assessment of the motor activity of hyperactive boys: II. Stimulant drug effects. Arch Gen Psychiat 40:688-693

68. De Crescenzo F, Armando M, Mazzone L, Ciliberto M, Sciannamea M, Figueroa C, Janiri L, Quested D, Vicari S (2014) The use of actigraphy in the monitoring of methylphenidate versus placebo in ADHD: a meta-analysis. ADHD Attention Deficit and Hyperactivity Disorders 6:49-58

69. O'Mahony N, Florentino-Liano B, Carballo JJ, Baca-García E, Rodríguez AA (2014) Objective diagnosis of ADHD using IMUs. Med Eng Phys 36:922-926

70. Reh V, Schmidt M, Lam L, Schimmelmann BG, Hebebrand J, Rief W, Christiansen H (2013) Behavioral Assessment of Core ADHD Symptoms Using the QbTest. Journal of Attention Disorders: doi: 1087054712472981

71. Heiser P, Frey J, Smidt J, Sommerlad C, Wehmeier P, Hebebrand J, Remschmidt H (2004) Objective measurement of hyperactivity, impulsivity, and inattention in children with hyperkinetic disorders before and after treatment with methylphenidate. Eur Child Adolesc Psy 13:100-104

72. Wehmeier PM, Schacht A, Wolff C, Otto WR, Dittmann RW, Banaschewski T (2011) Neuropsychological outcomes across the day in children with attention-deficit/hyperactivity disorder treated with atomoxetine: results from a placebo-controlled study using a computer-based continuous performance test combined with an infra-red motion-tracking device. J Child Adol Psychop 21:433-444

73. Ramtvedt BE, 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. The Clinical Neuropsychologist 28:1146-1161

74. Wehmeier PM, Schacht A, Ulberstad F, Lehmann M, Schneider-Fresenius C, Lehmkuhl G, Dittmann RW, Banaschewski T (2012) Does Atomoxetine Improve Executive Function, Inhibitory Control, and Hyperactivity?: Results From a Placebo-Controlled Trial Using Quantitative Measurement Technology. J Clin Psychopharm 32:653-660

75. Gunther T, Kahraman-Lanzerath B, Knospe EL, Herpertz-Dahlmann B, Konrad K (2012) Modulation of attention-deficit/hyperactivity disorder symptoms by short- and long-acting methylphenidate over the course of a day. J Child Adol Psychop 22:131-138

76. Ramtvedt BE, Røinås E, Aabech HS, Sundet KS (2013) Clinical gains from including both dextroamphetamine and methylphenidate in stimulant trials. J Child Adol Psychop 23:597-604

TABLE 1: Search strategies to identify studies for inclusion

Searches:
1. ADHD OR attention deficit hyperactivity disorder OR attention OR hyper$ OR ADD$ OR attention deficit disorder
2. Continuous Performance Test OR Continuous Performance Task OR Auditory Continuous Performance Test OR ACPT OR CANTAB OR Conner$ CPT OR CCPT OR Gordon$ Diagnostic System OR GDS OR Integrated Visual and Auditory test OR IVA OR QbTest OR Seidel Continuous Attention Test OR SCAT OR Test of Variables of Attention OR TOVA
3. diagnosing OR assessing OR medication management OR titration OR treatment monitoring OR response OR drug OR medication
4. motor activity OR measure$ OR quantifying

$ indicates a wildcard character

FIGURE 1: Flow diagram illustrating study selection process

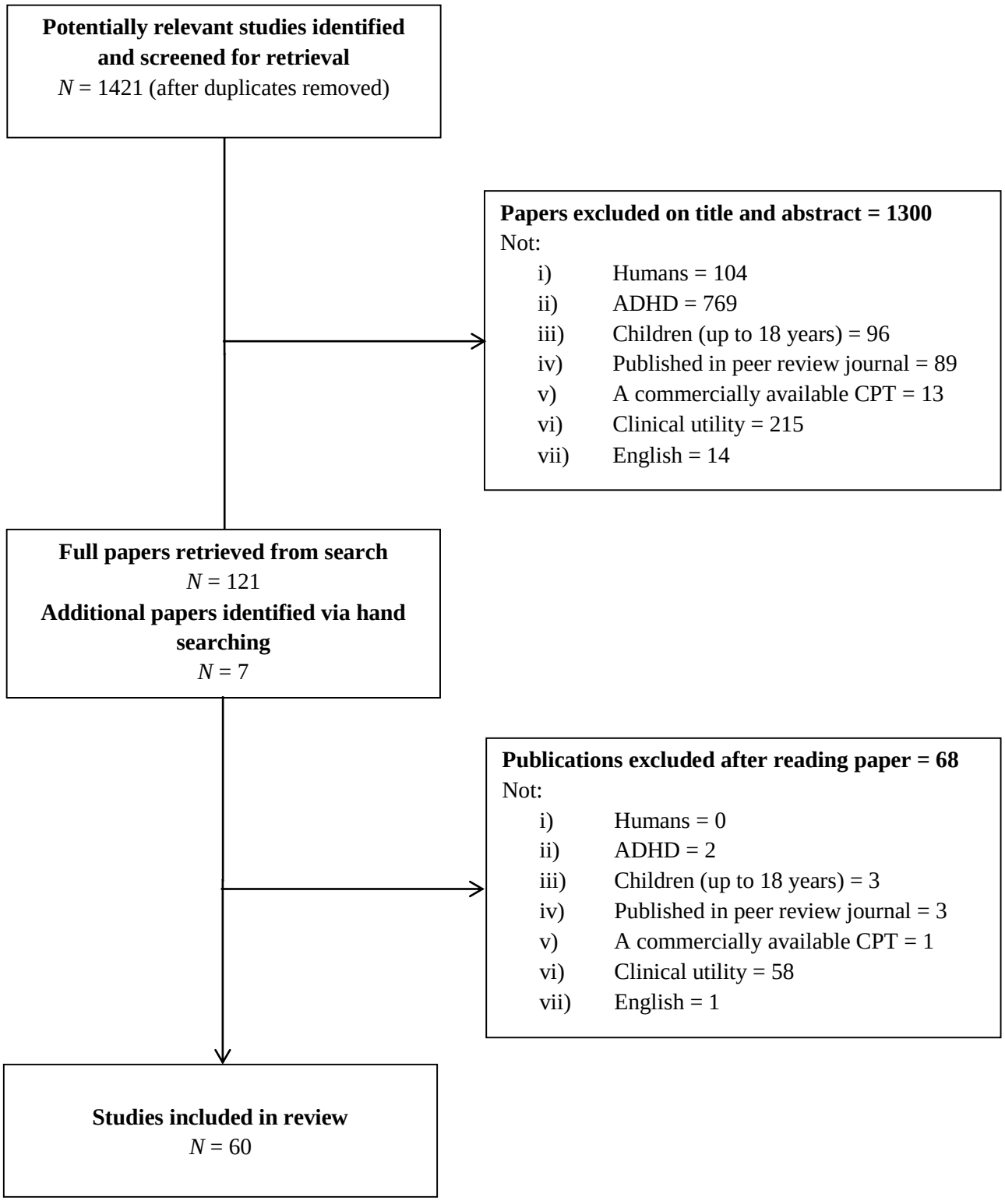


TABLE 2: Continuous performance tasks used for diagnosing ADHD

Reference	Design	CPT test	Other assessments	Sample	Main relevant outcomes
Alloway et al. (2009) [38]	Cross-sectional	CCPT	BRIEF; CTRS-R; WMRS	91 children (66 male, 25 female): 46 ADHD; 25 low working memory; 20 control (community sample). Age: 8-11 years.	CCPT scores did not accurately discriminate between ADHD and working memory impairment. CCPT correctly classified 41% of children with ADHD, 65% controls and 68% working memory impaired children.
Barkley & Grodzinsky (1994) [36]	Cross-sectional	GDS	CAP; COWAT; Grooved Pegboard Test; HMS; PM; Rey; Stroop; TMT; WCS; WISC-R	47 children (all male): 24 ADD/ADHD and 11 LD (clinic referred); 12 control (community sample). Age: 6-11 years.	GDS had reasonably high levels of positive predictive power for ADD/ADHD, but was not useful in discriminating between groups. GDS had low negative predictive power and high false negative rates. Abnormal scores may indicate an ADHD diagnosis, but normal scores could not be said to rule out a diagnosis.
Chae et al. (2003) [16]	Cross-sectional	TOVA	K-CBCL; KEDI-WISC SRBCSS; TRFC	177 children (113 male, 64 female): 106 gifted; 71 non-gifted (community sample). Age: 6-9 years.	13% of gifted children would have been classed as having ADHD using the other assessments alone, but were classified as not having ADHD as they had normal TOVA scores. Gifted children tend to perform better on TOVA. By utilisation of a higher TOVA standard 9% of gifted children were diagnosed as having ADHD.
DuPaul et al. (1992) [29]	Cross-sectional	GDS	ADHD Rating Scale-Teacher; CBCL-P/TRF; Clinical Interview; HSQ – P; MFFT; SSQ – Teacher; WISC-R	68 children (58 male, 10 female) with ADHD (clinic referred). Age: 6-11 years.	Scores on GDS frequently disagreed with the diagnosis of ADHD based on parent interview and behaviour rating scales. The percentage agreement between classifications rendered by clinic test scores and a diagnosis of ADHD based on parent and/or teacher report ranged from 12% to 62%.
El Sayed et al. (1999) [32]	Cross-sectional	GDS	Clinical interview (DSM-IV); WISC; other assessments (e.g. fine/gross motor assessment)	159 children (126 male, 33 female): 71 ADHD (clinic referred); 88 control (community sample). Age: 6-16 years.	GDS was well accepted in both control and clinical groups. GDS was able to differentiate between children with ADHD and control children. Areas under the ROC curve for GDS performance was between 0.72-0.73 for the GDS scores. GDS sensitivity of 49-59% and specificity of 81-87%.

Reference	Design	CPT test	Other assessments	Sample	Main relevant outcomes
Epstein et al (2003) [7]	Cross-sectional	CCPT	Clinical Interviews (CAPA);	816 children (421 male, 395 female): 21 ADHD; 795 control (community sample). Age: 9-17 years.	CCPT measures were significantly related to DSM-IV ADHD symptoms but did not demonstrate symptom domain specificity. CCPT differentiated children with ADHD from children with no diagnosis.
Forbes et al. (1998) [26]	Cross-sectional	TOVA	ACBC; ACTeRS; B-G; CBCL; Clinical Interview; CTRS-R DCBRS; PPVT-R	146 children (110 male, 36 female): 117 ADHD/ADD; 29 no ADHD/ADD diagnosis (clinic referred). Age: 6-12 years.	Teacher rating scales and TOVA did not identify identical groups of children, with each correctly classifying children misclassified by the other measure. TOVA could not distinguish between ADHD, ADD and control children.
Gordon (1986) [31]	Case study series	GDS	Clinical Interview/case histories	4 children (3 male, 1 female) (clinic referred). Age: 7-13 years.	GDS results led to a diagnosis of ADHD in children who showed ADHD symptoms during initial observations, as well as in those who did not show symptoms. Consistencies and inconsistencies in behavioural responses led to further assessments being performed. GDS helped clinical assessment.
Grodzinsky et al., (1999) [37]	Cross-sectional	GDS	COWAT; HMS; PM; Rey; Stroop; TMT; WCS	130 children (all male): 66 ADHD (clinic referred and/or diagnosed); 64 control (community sample). Age: 6-11 years.	Significantly more participants with ADHD had abnormal GDS scores than the control group. Positive predictive power based on correct responses was 87% and based on commission errors was 83%. Negative predictive power (NPP) based on correct responses was 61% and based on commission errors was 59%.
Kim et al (2015) [19]	Cross-sectional	IVA+CPT	QEEG; Korean DISC-IV	157 children (118 males, 39 females) 85 ADHD; 72 control (sample attending attention support group). Age: 9 years (mean).	IVA+CPT measures were significantly different between the ADHD and control group. IVA+CPT sensitivity of 72.9% and specificity 70.9% in detecting ADHD.

Reference	Design	CPT test	Other assessments	Sample	Main relevant outcomes
Kleinman et al. (2015) [41]	Cross-sectional	CCPT	K-SADS-PL; K-SADS-E; CGAS; Wechsler Abbreviated Scale of Intelligence; CDRS-R; YMRS; SNAP-IV	84 children (54 male, 30 female): 23 ADHD; 10 Bipolar (BD); 33 BD+ADHD (all clinic referred); 18 control (community sample). Age: 12-17 years.	No significant difference in CCPT scores between any groups, The sample was best clustered into two new groups (A+B), which were independent of their original diagnosis. ADHD and BD+ADHD were evenly split between the two groups.
Matier-Sharma et al. (1995) [35]	Cross-sectional	GDS	Actigraph	129 children (100 male, 29 female): 40 ADHD; 57 non-ADHD; 14 excluded subthreshold ADHD (clinic referred); 18 control (community sample). Age: 6-13 years.	Children with ADHD made significantly more errors and were more active than the other groups. GDS correctly classified 63% of ADHD and 94% of control children. Poor performance on GDS generally indicated ADHD but normal performance could not rule out ADHD. Specificity was between 0.83-0.94 for ADHD vs control children.
Mayes & Calhoun (2002) [30]	Cross-sectional	GDS	WISC-III	230 children (180 male, 50 female): 184 ADHD; 46 without ADHD (clinic referred). Age: 6-16 years.	IQ minus GDS composite score > 13 correctly identified 88% of children as having ADHD, this increased to 91% when combined with Freedom from Distractibility >11. GDS was useful for excluding an ADHD diagnosis. GDS was better at ruling out ADHD (70%) than WISC-III (31%).
Mayes et al. (2001) [17]	Cross-sectional	GDS	Clinical interview (DSM-IV) & observation; PBS; WISC-III; WIAT	211 children (approximately 163 male): 165 ADHD; 46 without ADHD (clinic referred). Age: 6-16 years.	GDS sensitivity was 90% and specificity 70%. 90% of children with an abnormal GDS score had a diagnosis of ADHD and 52% of children with a normal GDS score did not have ADHD. When IQ minus GDS score IQ >13, this increased positive predictive power to 91% and negative predictive power 67%.
Mayes et al. (2009) [33]	Retrospective analysis of clinical data	GDS	Clinical interview (DSM-IV); WISC-III/WISC-IV; WIAT/WIAT-II; PBS	587 children (gender not specified) with ADHD with or without a co-morbid disorder (clinic referred). Age: 6-16 years.	Classification accuracy for combined and inattentive ADHD was 70% for both Delay and for Delay + Distractibility errors, falling to 46% accuracy in predicting diagnosis for ADHD subgroups. Overall, GDS differentiated between ADHD subgroups with 72% accuracy.

Reference	Design	CPT test	Other assessments	Sample	Main relevant outcomes
McGee et al. (2000) [14]	Cross-sectional	CCPT	CBCL; Clinical Interviews (DSM-IV); CPRS; CTRS; GORT; WIAT; WISC	100 children (79 male, 21 female): 40 ADHD; 14 RD; 14 ADHD + RD; 32 control (clinic referred). Age: 6-11 years.	The CCPT did not distinguish between ADHD and clinical controls. Children with ADHD did not have higher scores than clinical controls; children with RD did. The CCPT task should not be used in children with reading disorders.
Perugini et al. (2000) [39]	Cross-sectional	CCPT	COWAT; K-ABC; Stroop test; Trail Making Test; WISC-III	43 children (all male): 21 diagnosed with ADHD; 22 control (community sample). Age: 6-12 years.	CCPT sensitivity of 67% and specificity of 73%. CCPT was more accurate in predicting ADHD than other measures alone, but the predictive validity increased further when measures were considered in combination.
Porumb (2007) [22]	Case study	TOVA	NEPSY	1 child (female) with ADHD (clinic referred). Age: 11 years.	TOVA did not predict the number of ADHD symptoms, but did assist in diagnosis, particularly in discriminating from other disorders. TOVA could identify the ADHD subtype.
Preston et al. (2005) [23]	Cross-sectional	TOVA	Clinical interview (DISC-IV); SNAP IV-P SNAP IV-T	167 children (67 male, 100 female): 116 ADHD; 51 sub-clinical ADHD control (community sample). Age: 6-14 years.	No significant difference in TOVA performance of children with ADHD than those with sub-clinical levels. TOVA did not predict number of ADHD symptoms independent of group membership. TOVA did not increase diagnostic accuracy in determining ADHD or sub-clinical difficulties in attention/impulsivity/activity.
Rielly et al. (1999) [27]	Cross-sectional	GDS	DBDRS-P; DBDRS-T WISC-R	99 children (all males) with language disorder (clinic referred) Age: 7-9 years.	Comparisons with rating scales showed low positive predictive values (20%-37%) but high negative predictive values (72%-88%). GDS accurately disconfirmed ADHD; likelihood ratios for normal scores were low to moderate (0.41 to 1.16).
Schatz et al. (2001) [24]	Cross-sectional	TOVA	CPRS	48 children (gender not specified): 28 diagnosed with ADHD; 20 control (community sample). Age: 5-17 years.	TOVA sensitivity of 86% and specificity of 70%. TOVA variability and RT differentiated between ADHD and non-ADHD groups. Variability predicted group membership.

Reference	Design	CPT test	Other assessments	Sample	Main relevant outcomes
Trommer et al. (1988) [34]	Cross-sectional	GDS	REITAN tests; WISC-R	20 children (14 male, 6 female): 14 ADD; 6 non-ADD (clinic referred). Age: 5-14 years.	The sensitivity of GDS was questioned. Using GDS, 29% of children diagnosed with ADD performed in the normal range, 36% in the borderline range and 36% in the abnormal range. 67% of non-ADD children performed in the abnormal range.
Vogt & Shameli (2011) [41]	Pre-test post test	QbTest	Clinical Interview; CPRS-R (long & short) CTRS-A; SDQ	108 children (gender: not reported): 70 ADHD; 38 not ADHD (clinic referred). Age: 9/10.5 years (mean).	QbTest reduced the risk of unidentified ADHD. 7 participants who had not received a QbTest had their diagnosis revised (from non-ADHD to ADHD) 1-year later. 5 participants with normal attention and impulse scores were diagnosed with ASD. QbTest improves diagnostic accuracy and stability.
Wada et al. (2000) [25]	Cross-sectional	TOVA	Clinical Interview; CPRS; CTRS; WISC-R	36 children (all male) 17 ADHD; 19 control Age: 6-12 years	TOVA was easily understood by both children and parents and was easy to administer in children with varying ethnicities and ages. Children with ADHD had higher means on all TOVA variables than those without ADHD.

ACBC Achenbach Child Behaviour Checklist, *ACTeRS* ADD-H Comprehensive Teacher's Rating Scale, *B-G* Bender-Gestalt, *BRIEF* Behavior rating inventory of executive function, *CAPA* Child and Adolescent Psychiatric Assessment, *CBCL(-P)* Child Behaviour Checklist (-Parent), *CBCL-TRF* Child Behaviour Checklist-Teacher Report Form, *CDRS-R* Children;s Depression Rating Scale-Revised, *CGAS* Children's Global Assessment Scale, *COWAT* The Controlled Oral Word Association Test or F-A-S, *CPRS/CTRS(-R)(-A)* Conners' Parent/Teacher Rating scale (-Revised) (-Abbreviated), *DBDRS-P/DBDRS*-T Disruptive Behaviour Disorders Rating Scale – Parent/Teacher, *DCBRS* Devereux Child Behaviour Rating Scale, *DISC-IV* The Diagnostic Interview Schedule for Children, *GORT* Gray Oral Reading Test, *HMS* Hand Movements Scale, *HSQ(-P)* Home Situations Questionnaire (-Parent), *K-ABC* Kaufman Hand Movements Scale from the K-ABC, *K-CBCL* Korea Children Behaviour Checklist, *KEDI-WISC* Korean Educational Development Institute-Welscher Intelligence, *K-SADS-PL/-E* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version/-Epidemiological version, *MFFT* Matching Familiar Figures Test, *NEPSY* A Developmental NEuroPSYchological Assessment, *PBS* Pediatric Behavior Scale, *PM* Porteus Mazes, *PPVT-R* Peabody Picture Vocabulary Test-Revised, *RD* Reading disorder, *REITAN TESTS* The Category and Tactile Performance Tests from the Halstead-Reitan Neuropsychological Test Battery for Older Children or Reitan-Indiana Neuropsychological Test Battery for Young Children, *Rey* Rey-Osterrieth Complex Figure, *SDQ* Strengths and Difficulties Questionnaire, *SNAP-IV-P/T* Swanson, Nolan & Pelham-IV (Parent/Teacher), *SRBCSS* Scale for Rating the Behavioural Characteristics of Superior Students, *SSQ-T* School Situations Questionnaire (-Teacher), *Stroop* Stroop Color-Word Test, *TMT* Trail Making Test, *TRFC* Teacher Report Form, *WIAT* Wechsler Individual Achievement Test, *WISC-R/-III/-IV* Wechsler Intelligence Scale for Children-Revised/-III/-IV, *WCS* Wisconsin Card Sorting Test, *WMRS* Working Memory Rating Scale, *YMRS* Young Mania Rating Scale

TABLE 3: Continuous performance tasks used for medication management in ADHD

Reference	Design	CPT test	Other assessments	Sample	Medication	Main relevant outcomes
Aggarwal and Lillystone (2000) [42]	Pre-post test	TOVA	CBCL; Clinical interviews (DSM IV); CPRS 28; CTRS 28	18 children (16 male, 2 female) with ADHD. Age: 8-16 years.	Dex or MPH	Significant improvement in mean commission errors after prolonged use of medications. Omission errors, response time and variability did not significantly improve after medications. Significant positive correlation between commission and omission scores.
Bédard et al. (2015) [47]	Double-blind crossover RCT	CCPT	K-SADS-PL; clinician judgement	101 children (77 male, 24 females) with ADHD. Age: 6-17 years.	MPH + ATX	Medication improved measures of attention on the CCPT, but not inhibitory control, however, the changes in attention scores for patients on ATX were not significant. Scores on the CCPT did not correlate with symptomatic improvement assessed via rating scale (ADHD-RS).
Fernández-Jaén et al. (2009) [46]	Pre-post-test	CCPT	ADHD-RS; D2 Test; EMA	94 children (69 male, 25 female) with ADHD. Age: 6-18 years.	MPH-ER	84% showed lower ADHD-RS score on MPH-ER. 77% improved omission and 55% commission after MPH-ER. CCPT was the most effective test for monitoring short-term effects of MPH.
Fischer & Newby (1991) [49]	Double blind randomised placebo controlled trial	GDS-Vigilance Task	CBCL; CPRS-R; CTRS-R; HSQ; Side effects rating scale; SSQ	161 children (141 male, 20 female) with ADHD. Age: 2-17 years.	MPH	GDS scores improved on both low and high doses of MPH compared to placebo. However, only the higher doses of MPH resulted in significant improvements on GDS scores.
Gordon (1986) [31]	Case study series	GDS	Clinical interview/case histories	4 children (3 male, 1 female) with ADHD Age: 7-13 years	Stimulant Medication	Delay and Vigilance Task scores were in the normal or near-normal range on medication but were impaired without medication. GDS led to recommendations to continue stimulant medications.
Huang et al. (2007) [44]	Pre-post test	TOVA	Clinical Interviews (DSM-IV & K-SADS-E); WISC-III	57 children (50 male, 7 female) with ADHD Age: 6-13 years	MPH	One dose of MPH enhanced commission errors, response time and ADHD scores, but not omission errors, response time variability or d′. Second half of TOVA was more sensitive in determining MPH effects.

Reference	Design	CPT test	Other assessments	Sample	Medication	Main relevant outcomes
Manor et al. (2008) [43]	Pre-post test	TOVA	CGI-C; Clinical Interview; CPRS; CTRS; SDQ-P; SDQ-T	165 children (110 male, 55 female) with ADHD Age: 5-18 years	MPH	All TOVA indices improved after MPH. Largest improvement in the commission mean score and the least improvement in the omission mean score. The older the patient the more prone he/she was to perceive improvement.
Tabori-Kraft et al. (2007) [21]	Pre-post-test	MMAT	CGI; Clinical Interview; DuPaul Questionnaire; WISC-III	23 children (22 male, 1 female) with HKD or ADD Age: 7-12 years	Dex MPH	Children showed a significant improvement on MMAT when on medication compared with no treatment. The improvement measured by MMAT was supported by clinical assessment. Children were pleased with their OPTAx report
Teicher et al. (2008) [48]	Triple-blind within-subject efficacy study	MMAT	CGI; Clinical Interview (DSM-IV, K-SADS-E 5th)	11 children (all males) with ADHD. Age: 6-12 years	MPH	MMAT identified the dosage parents perceived to be most beneficial in 82% of cases. MMAT can accurately assess a response to 0.4mg/kg dose of MPH.
Vogt and Williams (2011) [13]	Pre-post-test	QbTest	CPRS- R (short & long); CTRS-A; SDQ-P	44 children (36 male, 8 female) with HKD Age: 7-18 years	MPH	84% had improvements on MPH with measures (activity, attention, impulse control) reverting to the "normal" population mean. 7% demonstrated a partial response to MPH. 9% were non-responders.
Wang et al. (2011)[18]	Pre-post-test	CCPT	ADHD-RS; CBCL; SNAP-IV	50 children (40 male, 10 female) with ADHD Age: 6-12 years	MPH	Over 6 months of MPH treatment, CCPT scores on impulsivity, hyperactivity and inattention improved, but CCPT assessed distraction did not.
Wang et al. (2015) [45]	Longitudinal	TOVA	SNAP-IV; ADHD-RS	181 children (151 male, 30 female) with ADHD Age: 8-18 years	MPH	Limited correlation between TOVA scores and behavioural ratings.

ADHD-RS Attention-Deficit/Hyperactivity Disorder Rating Scale, *ATX* atomexetine, *CBCL* Child Behaviour Checklist, *CGI(-C)* Clinical Global Impression (-change scale), *CPRS/CTRS (-A)(R)(28)* Connors' Parent/Teacher Rating Scale(-Abbreviated) (-Revised) (-28 item), *Dex* Dexamphetamine, *HSQ* Home Situations Questionnaire, *K-SADS-PL/-E* Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime version/-Epidemiological version, *MPH* Methylphenidate, *MPH-ER* Methylphenidate Extended Release, *SDQ-P/T* Strengths and Difficulties Questionnaire- Parent/Teacher, *SNAP-IV* Swanson, Nolan & Pelham-IV, *SSQ* School Situations Questionnaire, *WISC-III* Wechsler Intelligence Scale for Children-III, *D2 Test* Brickenkamp's D2 Test, *EMA* Escalas Magallanes de Attencion Visual Face Perception Test

TABLE 4: Objective measures of activity in ADHD using actigraphy

Reference	Design	Activity test	Other assessments	Sample	Main relevant outcomes
Alderson et al. (2012) [53]	Cross-sectional	MicroMini Motionlogger (R) Actigraph (Ambulatory Monitoring, 2004)	CBCL; Clinical Interview (DSM-IV, K-SADS); CSI- P; CSI- T TRF; WISC-III/IV	22 children (all males): 11 with ADHD (clinic/community referred); 11 control (community sample). Age: 8-12 years.	Children with ADHD were more active than typically developing children across all experimental conditions. All children exhibited significantly higher activity rates under all experimental tasks relative to control.
Borcherding et al. (1989) [65]	Double blind crossover trial	Actigraphy	n/a	18 children with ADHD (all male) (community and clinic sample). Age: 9.6 years (mean).	MPH resulted in a greater reduction in motor activity as measured by actigraphy than Dex.
Dane et al. (2000) [54]	Repeated measures	Actigraph (Ambulatory Monitoring Inc, 1996)	Clinical Interview (DSM-IV); IOWA-C-10; OCHS; PICS; TTI	64 children (49 male, 15 female): 42 with ADHD; 22 control (clinic referred). Age: 7-12 years.	No significant differences in activity between children with ADHD and controls in the morning session. Children with ADHD were significantly more active than controls during the afternoon session, but there were no differences between ADHD subtypes.
DeCrescenzo et al. (2014)[68]	Meta-analysis of placebo controlled crossover trials (MPH/Placebo)	Actigraphy	n/a	8 papers 393 children (187 male, 206 female) with ADHD. Age: 6-12 years.	Actigraphy was able to assess the effect of MPH on sleep and activity. Actigraphy showed MPH can negatively affect total sleep and reduce mean activity in ADHD children compared to a placebo.
Donnelly et al. (1989)[66]	Double blind crossover placebo controlled trial	Actigrapy	CBRS; CGAS; CPQ; CPRS	20 ADHD children (male) (community & clinic) Age: 6-12 years.	Immediate marked improvement in disruptive, overactive behaviour after Dex. No effect with FEN.
Hall et al. (1997) [60]	Cross-sectional	Actigraphy	CBCL; CBQ; CTQ; PPVT-R; RCPM; WRAT-R	70 children (47 male, 23 female) (clinic referred). Age: 6-13 years.	Children with ADHD were more impulsive regardless of whether there was reading disability present. ADHD children more active than non-ADHD.

Reference	Design	Activity test	Other assessments	Sample	Main relevant outcomes
Halperin et al. (1992) [56]	Cross-sectional	Solid state Actigraphs Halperin CPT	CBCL; Clinical Interview (DSM-III-R); CTRS; PIAT-R; PPVT-R; RCPM; WRAT-R	102 children (83 male, 19 female): 31 with ADHD and 53 without ADHD (clinic referred); 18 control (community sample). Age: 6-13 years.	Patient groups were inattentive in comparison to control children but differences were not significant. The ADHD group was more active than both non-ADHD patients and control children; who did not differ from one another.
Halperin et al. (1993) [59]	Cross-sectional	Solid state actigraph (Ambulatory Monitoring Inc) Halperin CPT;	RCPM or WISC-R	66 children (52 males, 14 females): 13 ADHD; 20 anxiety disorder; 15 disruptive disorders (ODD/CD) (clinic referred); 18 control (community sample). Age: 6-13 years.	Activity measures were able to distinguish the ADHD children from the control children, but did not distinguish between patient groups.
Inoue et al. (1998) [57]	Cross-sectional	Solid-state actigraph CPT	MFFT; WISC-R-JV	72 children (all males): 20 with ADHD (clinic referred); 52 control (community sample). Age: 6-12 years.	Actigraph activity measured during the 10 minutes of CPT, and the commission and omission errors of the CPT differentiated the ADHD group from the control group with sensitivity of 65% and specificity of 76%.
Konrad et al. (2005) [64]	Double-blind Placebo-controlled (MPH/Placebo)	Actigraphy (Cambridge Neurotechnology, version 2.56, Cambridge, UK)	CBCL; Clinical Interview (DSM IV); FBB-HKS; K-DIPS; WISC III IQ	44 children (37 male, 7 female) with ADHD (clinic referred).Age: 8-12 years.	MPH produced significant improvements in both objective and subjective measures of ADHD. There was a linear effect of MPH dose on activity and a relationship between actigraphy scores at school and in test situation.
Marks et al. (1999)[61]	Cross-sectional	Actigraphy	CAS-T & P; CBCL; CPT; IOWA-CTQ; WISC	66 children (all males) with ADHD (clinic referred). Age: 7-11 years.	CPT measured inattention correlated with activity. Lab based measures enabled clinicians to distinguish between four groups of ADHD (hyperactive-inattentive, impulsive-inattentive, inattentive only, hyperactive only) and may provide a useful tool in assessment/diagnosis.
O'Mahony et al (2014) [69]	Cross-sectional	Intertial measurement units	ADHD-rating scale; TOVA; clinical judgment	43 clinic-referred children: 24 ADHD (17 male), 19 non ADHD (male). Age: 6-11 years	Sensitivity of IMU was 94.44% and 95.65% in classifying ADHD or non-ADHD.

Reference	Design	Activity test	Other assessments	Sample	Main relevant outcomes
Porrino et al. (1983a) [58]	Cross-sectional	Actigraphy (solid state)	Activity diary (hourly); CPT; CPRS; CTRS-A.	24 children (all males): 12 hyperactive; 12 control (community and clinic sample). Age: 6-12 years.	Hyperactive children had higher activity levels than matched controls across both day and night. Differences in activity were greater during structured school tasks.
Porrino et al. (1983b) [67]	Double blind ABAB design (Dex/Placebo)	Actigraphy	CTRS; CPRS; TRS	12 hyperactive children (all male) (community and clinic sample). Age: 6-12 years.	A single dose of Dex reduced activity for 8 hours after administration with a rebound effect occurring in the early evening and during sleep. CPT changes were minimal due to timing of test.
Rajendran et al. (2015) [62]	Longitudinal	Actigraphy	NEPSY; Kiddie CPT; K-SADS: present & lifetime; Nakao-Treas Socioeconomic prestige index	214 children (156 male, 58 female): 105 hyperactive/ inattentive, 51 control (community sample). Age: 3-4 years at baseline	Actigraphy could significantly differentiate typically developing from hyperactive/inattentive pre-schoolers. Only actigraphy could predict ADHD outcome 4 and 5 years later.
Rapoport et al. (1980) [63]	Double-blind, Placebo-controlled	Actigraphy	Behaviour ratings; CTRS; Skill ratings; WISC	10 children (all males) with hyperactivity/ impulsivity (clinic referred). Age: 9 years (m)	Dex use produced lower activity levels relative to placebo. The decrease in activity correlated with teacher reported improved behaviour.
Wood et al. (2009) [55]	Cross-sectional	2 Actigraphs: 1) MTI Health Services, V323 2) Health One Technology, Pensacola, FL	Clinical Interview (DSM-IV); CPRS-LV CTRS-LV; PACS	507 children (approx 332 males): 116 with ADHD & 119 siblings (International Multicentre ADHD Genetics project); 250 control (community sample). Age: 6-18 years.	High ADHD actigraph familial correlations in activity. Actigraph measures yielded an area under the curve of up to 0.8, indicating an ability to distinguish between cases and controls.

CAS-T/P Children's Aggression Scale-Teacher/Parent, *CBCL* Child Behavior Checklist, *CBQ* Child Behaviour Questionnaire, *CBRT* Conners brief rating scale (Teacher) *CGAS* Clinical Global Impression Scale, *CSI-P/T* Child Symptom Inventory – Parent/Teacher, *CTQ/CPQ* Conners Teacher's/Parent's Questionnaire, *CTRS/CPRS (-LV/A)* Conners' Teacher Rating Scale / Conners' Parent Rating Scale (-Long Version/-Abbreviated), *FEN* Fenfluramine hydrocholoride, *IOWA-C-10* IOWA-Conners rating scale (10 item), *IOWA-CTQ* IOWA-Conners Teacher's Questionnaire, *KITAP* Test of Attentional Performance for Children, *K-SADS(-E)* Kiddie-Schedule for Affective Disorders and Schizophrenia for School-Aged Children (Epidemiologic Version), *MFFT* Matching Familiar Figures Test, *NEPSY* Developmental Neuropsychological Assessment, *OCHS* Ontario Child Health Study Scales, *PACS* Parental Account of Childhood Symptoms, *PIAT-R* Peabody Individual Achievement Test Revised, *PICS* Parent Interview for Child Symptoms-IV, *PPVT-R* Peabody Picture Vocabulary Test-Revised, *RCPM* Raven's Coloured Progressive Matrices, *TRF* Teacher Report Form, *TTI* Teacher Telephone Interview-IV, *WISC-(III/IV/JV/R)* Wechsler Intelligence Scale for Children-(III/IV/Japanese version/Revised), *WRAT-R* Wide Range Achievement Test Revised

TABLE 5: Objective measures of activity in ADHD using infra-red motion analysis

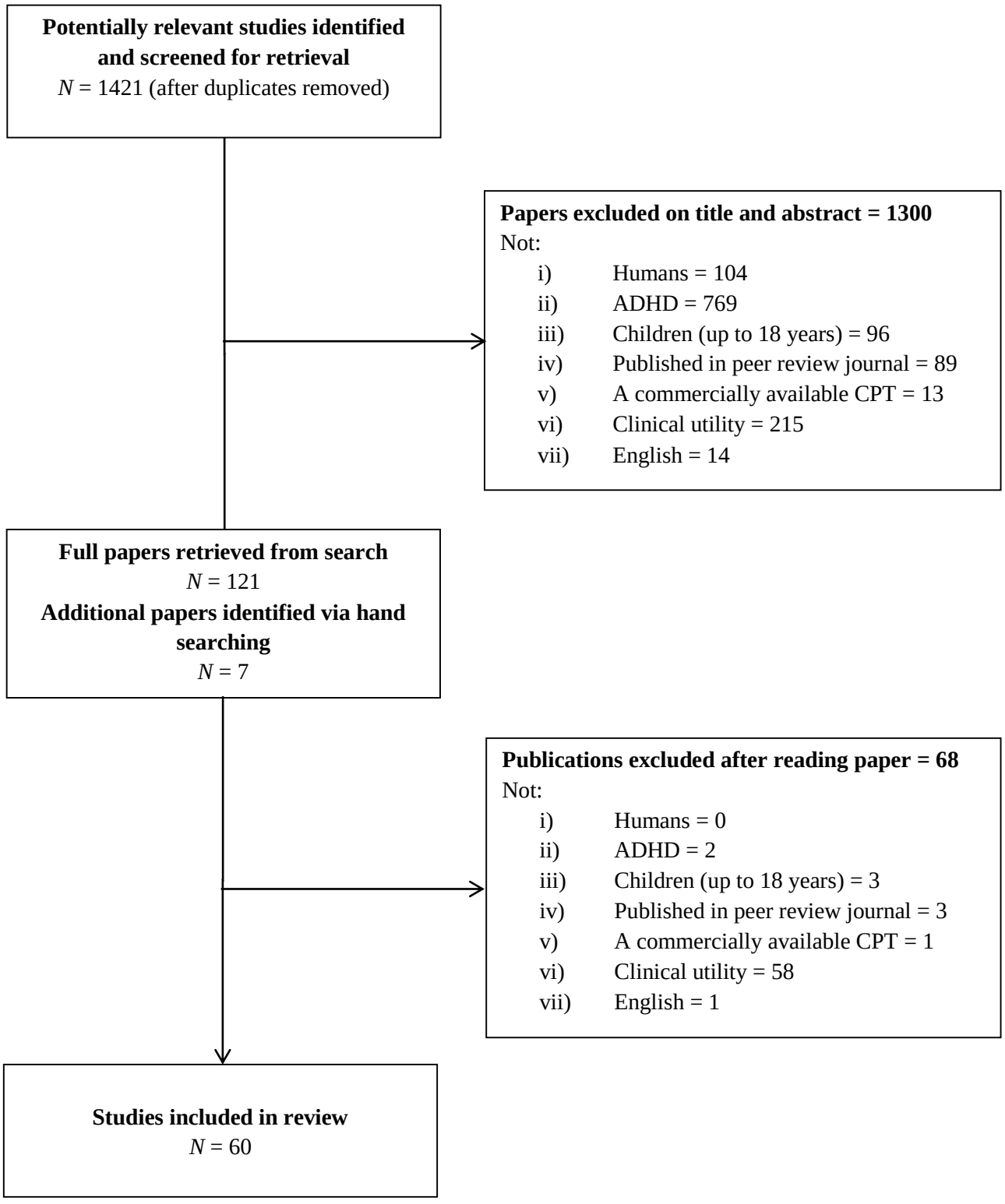
Reference	Design	Activity test	Other assessments	Sample	Main relevant outcomes
Gunther et al. (2012) [75]	Pre-post-test RCT (MPH-ret/MPH-IR)	QbTest	CBCL; Clinical Interview (DSM-IV-TR); FBB-HKS (T) K-DIPS; WISC-III	56 children (43 male, 13 female): 36 with ADHD (clinic referred); 20 control (community sample). Age: 8-12 years.	Significant differences were identified between groups in inattention, hyperactivity and impulsivity. Both MPH-ret and MPH-IR groups improved after medication, and behaviour was comparable to control children. MPH-IR was better than MPH-ret in reducing impulsivity.
Heiser et al. (2004) [71]	Pre-post-test (MPH)	MMAT (OPTAx)	CBCL; Clinical Interview (ICD-10)	25 children (20 male, 5 female) with hyperkinetic disorders (clinic referred). Age: 6-12 years.	Significant improvement in activity, impulsivity, and attentiveness after MPH.
Ramtvedt & Sundet (2014)[73]	Secondary analysis of Ramtvedt et al. (2013) (MPH/Dex/ Placebo)	QbTest	ADHD-questionnaire	36 children (29 male, 7 female) (clinic referred). Age: 9-14 years.	QbTest improved after MPH and Dex but not after placebo. Activity measure had largest effect size. Convergent validity for rating scales and the QbTest were found on a group level, but not on an individual level.
Ramtvedt et al. (2013)[76]	Cross-over RCT (MPH/Dex/ Placebo)	QbTest	ADHD-questionnaire	36 children (29 male, 7 female) with ADHD (clinic referred). Age: 9-14 years.	Significant treatment effects for stimulant drugs on both QbTest and rating scale. Including both stimulants favourable responses increased from 72% to 92%.
Reh et al. (2013) [70]	Cross-sectional	QbTest	Clinical Assessment (DSM-IV-TR); CPRS; CTRS; KITAP; PACS; WISC-IV	Sample 1: 828 children (588 males, 240 females) (clinic referred). Age: 6-12 years. Sample 2: 102 children (79 males, 23 females) diagnosed with ADHD Age: 6-12 years	In children referred for an ADHD diagnosis a three-factor model (attention, impulsivity and activity) explained 76% of the total variance in QbTest data; the activity parameter explained the largest amount of variance.

Reference	Design	Activity test	Other assessments	Sample	Main relevant outcomes
Teicher et al. (1996) [51]	Cross-sectional	M-MAT	CBCL; Clinical Assessment (DSM IV); CPRS; CTRS; IOWA K-SADS-E;	29 children (all males): 18 with ADHD (clinic referred); 11 control (community sample). Age: 6-14 years.	Boys with ADHD moved their head 2.3 times more often than normal children, 3.4 times as far, and had a more linear and less complex movement pattern. Boys with ADHD had slower responses and greater variability on CPT relative to controls.
Wehmeier et al. (2011) [74]	Double-blind Placebo-controlled, RCT (ATX/Placebo)	QbTest	ADHD-RS-IV-P; CGI-S; Clinical Interview (DSM-IV-TR); DCHD; WREMB-R-Inv	128 children (99 male, 29 female) with ADHD (clinic referred).Age: 6-12 years.	Both the ATX and placebo groups showed a circadian pattern of neuropsychological outcomes. ATX was shown to have a positive effect on activity ratings.
Wehmeier et al. (2012)[72]	Double-blind Placebo-controlled, RCT (ATX/Placebo)	QbTest	ADHD-RS; CGI-S; WREMB	135 children (105 male, 30 female) with ADHD (clinic referred). Age: 6-12 years.	Similar to the findings of the previous study actigraphy was sensitive to ATX.

ADHD-RS(IV-P) ADHD Rating Scale (IV-Parent Version), *ATX* Atomexetin, *CBCL* Child Behavior Checklist, *CGI-S* Clinical Global Impressions Scale – Severity, *DCHD* Diagnostic Checklist for Hyperkinetic Disorders, *Dex* Dextroamphetamine, *FBB-HKS* The German Behavioural rating scale for ADHD; Fremdbeurteilungsbogen für Hyperkinetische Störungen (T= Teacher rated) section of DSIYPS (Diagnostik-System für Psychische Störungen im Kindes- und Jugendalter), *K-DIPS* German semi-structured interview, *MPH-IR/MPH-Ret* Methylphenidate immediate release/Retard, TRS Teacher Rating Scale, WISC-*III* Wechsler Intelligence Scale for Children, 3rd edition, *WREMB(-R-Inv)* Weekly Ratings of Morning and Evening Behavior (Revised-Investigator Rated Scale)

Figure

FIGURE 1: Flow diagram illustrating study selection process



Table

TABLE 1: Search strategies to identify studies for inclusion

Searches:
1. ADHD OR attention deficit hyperactivity disorder OR attention OR hyper$ OR ADD$ OR attention deficit disorder
2. Continuous Performance Test OR Continuous Performance Task OR Auditory Continuous Performance Test OR ACPT OR CANTAB OR Conner$ CPT OR CCPT OR Gordon$ Diagnostic System OR GDS OR Integrated Visual and Auditory test OR IVA OR QbTest OR Seidel Continuous Attention Test OR SCAT OR Test of Variables of Attention OR TOVA
3. diagnosing OR assessing OR medication management OR titration OR treatment monitoring OR response OR drug OR medication
4. motor activity OR measure$ OR quantifying

$ indicates a wildcard character