

Calibration of ADHD Assessments Across Studies: A Meta-Analysis Tool

Jennifer Kim Penberthy,^{1,7} Daniel Cox,¹ Marc Breton,² Raina Robeva,³
M. Layne Kalbfleisch,^{5,6} Tim Loboschewski,⁴ and Boris Kovatchev¹

When analyzed separately, data from small studies provide only limited information with limited clinical generalizability, due to small sample size, differing assessments, and limited scope. In this methodological paper we outline a theoretical framework for performing meta-analysis of data obtained from disparate studies using disparate tests, based on calibration of the data from such studies and tests into a unified probability scale. We apply this method to combine the data from five studies examining the diagnostic abilities of different assessments of Attention Deficit/Hyperactivity Disorder (ADHD), including behavioral rating scales and EEG assessments. The studies enrolled a total of 111 subjects, 56 ADHD and 55 controls. Each individual study had a small sample focused on a specific age/gender group, for example 8 boys ages 6–10, and generally had insufficient power to detect statistically significant differences. No gender, or age comparisons were possible within any single study. However, when calibrated and combined, the data resulted in a clear separation between ADHD versus non-ADHD groups in males below the age of 16 ($p < 0.001$), males above the age of 16, ($p = 0.015$), females below the age of 16, ($p = 0.0014$), and females above the age of 16, ($p = 0.0022$).

We conclude that if data from various studies using various tests are made comparable, the resulting combined sample size and the increased diversity of the combined sample lead to increased significance of the statistical tests and allow for cross-sectional comparisons, which are not possible within each individual study.

KEY WORDS: attention-deficit/hyperactivity disorder (ADHD); diagnostic assessment; electroencephalography (EEG); meta-analysis.

¹Center for Behavioral Medicine Research, Department of Psychiatric Medicine, University of Virginia Health System, Charlottesville, Virginia.

²Department of Systems Engineering, University of Virginia, Charlottesville, Virginia.

³Department of Mathematical Sciences, Sweet Briar College, Sweet Briar, Virginia.

⁴Department of Psychology, Sweet Briar College, Sweet Briar, Virginia.

⁵Krasnow Institute & Graduate School of Education, George Mason University, Fairfax, Virginia.

⁶Department of Pediatrics, The George Washington University School of Medicine and Health Sciences, Washington, D.C.

⁷Address all correspondence to Jennifer Kim Penberthy, Ph.D., Department of Psychiatric Medicine, PO Box 800137, UVA Health System, Charlottesville, Virginia 22908; e-mail: jkp2n@virginia.edu.

INTRODUCTION

Attention Deficit/Hyperactivity Disorder (ADHD) is the most common neurobehavioral disorder of childhood, and is among the most prevalent chronic health conditions affecting school age children (American Academy of Pediatrics, 2000). The symptoms of hyperactivity, impulsivity, and inattention that characterize ADHD individuals also can contribute to multiple other negative impacts, such as substance abuse, low academic achievement, interpersonal conflicts, low self-esteem, high injury rates, and so forth (Barkley, 1990; Barkley, Fischer, Edelbrock, & Smallish, 1990; Barkley, Guevremont, Anastopoulos, DuPaul, & Shelton, 1993; Schachar, Taylor, Wieselberg, Ghorley, & Rutter, 1987; Almond, Tanner, & Goffman, 1999). Like most psychiatric disorders, the diagnosis of ADHD relies on subjective criteria. Unlike a neurological condition such as stroke, in which examination and neuroimaging provide clear, objective criteria in diagnosis, ADHD lacks the “hard evidence” that aids in evaluation and treatment. The difficulty in clinical diagnosis is reflected in the frequent shifts in the diagnostic criteria for ADHD. The DSM-III (APA, 1980), DSM-III-R, (APA, 1987), and DSM-IV (APA, 1994) all present different conceptualizations of ADHD. The most current criteria from the *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition (DSM-IV, 1994) distinguishes three subtypes of ADHD: predominantly inattentive type, predominantly hyperactive-impulsive type, and combined type. DSM-IV diagnostic criteria for ADHD include “a persistent pattern of inattention and/or hyperactivity-impulsivity that is more frequent and severe than is typically observed in individuals in a comparable level of development. Evidence of six of nine inattentive behaviors and/or six of nine hyperactive-impulsive behaviors must have been present before age seven, and must clearly interfere with social, academic, and/or occupational functioning” (APA, 1994). Consequently, diagnosis of ADHD is typically based on retrospective reports from parents and teachers of a child’s behavior, and is highly dependent upon the subjective judgments about the degree of relative impairment. Due to the subjective nature of assessment, precision in diagnosis has been elusive. Determining a biological measure that could aid in the diagnosis of ADHD would help to refine diagnostic criteria and may provide more specific diagnostic tests for ADHD and other disorders of attention and self-regulation. Although research supports a neurological basis for ADHD (Hynd et al., 1991; Castellanos, 1997; Berquin et al., 1998; Mostofsky, Reiss, Lockhart, & Denckla, 1998; Raskin, Shaywitz, Shaywitz, Anderson, & Cohen, 1984; Crawford & Barbasz, 1996; Heilman, Voeller, & Nadeau, 1991; Zametkin et al., 1990; Castellanos et al., 2001; Castellanos et al., 2002), there are only a few large studies examining EEG as a quantifiable physiological marker of ADHD. Of even greater importance, there is currently no uniform methodology for calibrating or standardizing the multiple disparate ADHD assessment tools currently available for clinicians and researchers.

Unfortunately, while neuroanatomical findings support the notion that ADHD is a distinct clinical syndrome and add to our understanding of the etiology of ADHD, neuroimaging techniques are too expensive for general use in diagnosing ADHD, are restricted to a few centers, and lack clear specificity and sensitivity in the diagnosis of ADHD. One technique suggested by a National Institute of Mental Health committee as a possible method to identify functional measures of child and adolescent psychopathology (Jensen et al., 1993) is that of quantitative EEG. Compared to methods of functional neuroimaging (such as positron emission tomography or single photon emission computed tomography),

quantitative EEG is easier to perform, less expensive, does not involve radioactive tracers, and is noninvasive (Kuperman, Gaffney, & Hamdan-Allen, 1990).

There are a few basic methods of analyzing EEG data that have been employed in previous research — visual inspection of raw data and quantitative analyses of EEG data, including spectral and coherence methods of analysis. To date, none of these methods have revealed pervasive or consistent patterns of EEG abnormalities with sufficient specificity or sensitivity to separate children with ADHD from normal subjects. Despite overall inconsistencies in the literature, one important consistency is the finding of patterns of higher levels of theta relative to beta (Bresnahan, Anderson, & Barry, 1999; Lazzaro et al., 1998; Mann, Lubar, Zimmerman, Miller, & Muenchen, 1992; Monastra et al., 1999).

Mann et al. (1992) tested 25 nine to twelve year-old boys with Predominantly Inattentive-type ADHD across three settings: baseline resting, reading, and drawing, and compared their EEG data to 27 controls matched for age and grade level. Five minutes of EEG data were recorded for each of the three tasks. Power spectral analysis was employed to topographically represent the brain and aid in analysis of EEG for individual and group comparisons. Findings include an increased theta (4–7.75 Hz) in the ADHD group when compared to normal control subjects, at both absolute and relative percent power calculations. In other words, the presence of excess theta was detected as an absolute calculation of the EEG band power and the percent calculated relative to the individual subject's EEG data. Results also indicated decreased beta (12.75–21 Hz) in the ADHD group when compared to normals, signifying less active mental processing in the ADHD children. The differences in both theta and beta were greater when subjects were tested while reading or drawing, but were decreased when they were at rest during visual fixation. This pattern predicted accurate classification of ADHD subjects 80 percent of the time and of controls 74 percent of the time. Mann et al. suggest that the pattern of theta activity detected in their data is parallel in nature to the decreases in glucose metabolism detected in the prefrontal and premotor areas of the brain by Zametkin et al. (1990). Mann et al. concludes that their study supports the use of quantitative EEG for further elucidation of the neurophysiology of ADHD.

Lazzaro et al. (1998) explored the possibility of maturational lag and cortical under-arousal as causes of ADHD. EEG activity of 26 adolescent (mean age of 13.4) unmedicated ADHD males and 26 age and gender matched controls were examined in a resting eyes open condition. In the ADHD subjects, increased levels of theta activity in the anterior region of the brain and reduced levels of beta activity in the posterior region were discovered. Lazzaro, et al. suggests that these results show evidence of a maturational lag (due to the presence of theta in the anterior regions of the brain and a reduction of beta in the posterior regions) and reduced cortical arousal in ADHD.

Bresnahan, Anderson, and Barry (1999) investigated patterns of activity with quantitative EEG in subjects, age 6 to 42 years of age. Analyzing data in 25 children, 25 adolescents, and 25 adults with ADHD and their age-matched, normal control group, from the midline EEG channels (Fz, Pz, Cz) they observed increased theta activity and decreased beta activity across all age groups. The decline in beta activity reduced with age. Bresnahan, et al. conclude that because the hyperactivity component in ADHD tends to decrease with age while the impulsivity tends to persist, that their data suggests reduced beta activity may be related to hyperactivity and that increased theta activity may be related to impulsivity. These patterns replicate findings of Mann et al. (1992).

Monastra et al. (1999) report similar results from their study on 482 subjects (ages 6–30) classified into the three groups of ADHD Combined-type, ADHD Predominantly Inattentive-type, and controls. Monastra and his group are the only ones to provide information regarding sensitivity and specificity (of 86% and 98%, respectively) when classifying individuals with Attention-Deficit Hyperactivity Disorder (ADHD) against controls by means of the Attentional index (AI). This index, which is the numerical inverse of a previously published ratio used by NASA (Engagement Index) to track attentional changes in pilots (Pope & Bogart, 1991; Pope, Comstock, Bartolome, Bogart, Burdette, 1994) was significantly larger for the ADHD groups than for the control group. Further, the authors set normal limits for their AI based on the mean and standard deviation (*SD*) of their control group and used these normal limits to classify their ADHD subjects. The EEG data were used to test the hypothesis that frontal cortical slowing (presence of excess theta) can differentiate ADHD subtypes from controls. Analysis of variance demonstrated cortical slowing that differentiated ADHD subjects, regardless of sex or age. Specifically, consistent with the presented hypothesis, statistical analysis revealed that the ADHD groups (inattentive and combined type) displayed significantly higher levels of slow-wave (theta) *relative* to fast-wave (beta) EEG activity. A methodological concern in this study maybe the control group used for setting normal limits for the diagnostic criteria. Instead of age and gender-matched random sample, the control group consisted of individuals seeking a differential diagnosis for possible ADHD. In addition, the control group of the study was relatively small (only 85 subjects versus 397 ADHD) and the normal limits for the AI were set on the basis of quite small control subgroups half of which were 11 subjects or less.

In a series of follow-up experiments, Monastra, Lubar, and Linden (2001) proceed to further explore results from their initial study in order to “replicate prior findings and examine the reliability and validity” (p. 138). For the first experiment, a cross-validation of the initial study, Monastra et al. report findings from 129 individuals, ages 6–20 (98 males and 31 females), who were diagnosed as ADHD-inattentive, ADHD- hyperactive or combined type, or normal control. These researchers determined that “groups of participants with either inattentive or combined types of ADHD could be differentiated from nonclinical control groups on the basis of an attentional index derived from a spectral analysis of power in the theta and beta bands” (p. 141). Their second experiment examined the criterion-related validity of the simplified QEEG process by comparing classifications based on EEG data with those derived from more traditional data used to diagnose ADHD, such as behavioral ratings, patient history, and so forth. Monastra et al. employed 55 ADHD individuals, ages 6 to 20 (50 males and 5 females) who had been previously classified as ADHD by the QEEG, ADDES, and CPT tests. Based on QEEG data obtained from individuals with ADHD tested 30 days apart, this reliability study found a test-retest correlation coefficient of .96 ($p < .05$). Their third experiment investigated the test-retest reliability of the attentional index the authors derived from the QEEG process. Two hundred and eighty-five individuals between the ages of 6 and 20, who were diagnosed with ADHD, were assessed with QEEG and other commercially available tests that were developed to assess attentional deficits (e.g. behavioral and CPT measures). Findings indicate that classification agreement was noted in 83% of the participants evaluated with the QEEG and the ADDES. Classification consistency was not obtained when QEEG result were compared with those derived from Conner’s CPT (Conners, 1997). The authors conclude, however, that because all of the participants in this experiment had been diagnosed with ADHD, issues of test

sensitivity and specificity could only be partially addressed. They also note that due to an absence of a “gold standard” for diagnosing ADHD, it is difficult to adequately address issues of test sensitivity and specificity.

Therefore, although there is evidence of the effectiveness of the theta/beta ratio in diagnosing ADHD, it is also evident that none of the currently existing psychological and physiological markers for ADHD has been proven to hold immediate promise of becoming the “gold standard” for diagnosing ADHD. Currently, there does not appear to be an established, well-tested, comprehensive assessment protocol, which includes multiple EEG data in the diagnoses of ADHD, and which can consistently produce findings at levels sufficient for clinical diagnosis.

An approach of combining several disparate measures into a single coherent assessment may be of interest, especially if proven to hold a compounded classification accuracy that is significantly greater than the accuracy of each individual measure. In our recent paper (Robeva, Penberthy, Loboschewski, Cox, & Kovatchev, 2004) we proposed a new psychophysiological assessment procedure for diagnosing ADHD, capable of combining various disparate ADHD markers into a powerful assessment tool whose classification power exceeds the classification power of its separate components. The formal framework of that combined assessment employs a Bayesian algorithm that allows for linking of disparate assessment instruments into one unified probability assessment. Important features of this sequential assessment are that it is test-order-invariant, and can accommodate missing data. Another important feature of this sequential assessment is that it can be used to unify separate studies, each of which may utilize different assessments or tests, thereby increasing the total number of subjects in a meta-analysis and the diversity of the studied population. As a result the power of the statistical tests increases, new comparisons between subgroups of the unified large sample become possible, and the clinical validity and generalizability of the results increases. The key to this meta-analysis approach is the calibration of data from various studies into a single scale - probability for ADHD.

We present a strategy designed to incorporate into a single analysis data from various small studies and various tests and markers for ADHD, none of which could claim perfect sensitivity and specificity to ADHD. For example, in Robeva et al. (2004) we employed data from twelve female college students using both self-report and others' report of symptoms, EEG information when transitioning from one cognitive task to another, behavioral information, cognitive performance, and history of symptoms. Classification of subjects as ADHD or non-ADHD improved with the sequential steps, achieving 100% accuracy by the final EEG measures. One of the problems with the study just reported, as with many studies of ADHD, is the small sample size employed and the subsequent limited generalizability of findings. Therefore, we evaluated our proposed Bayesian model as a method for combining not only related diagnostic tests, but also related studies that include such tests.

This study expands upon our previous proposal of a combined stochastic assessment for diagnosing ADHD, by presenting in detail a method for performing a meta-analysis on separate studies of ADHD, and by illustrating this method with data from five studies performed in our group. These individual studies ranged in sample size from eight to forty-nine, with three of the five studies having a sample size of twelve or less. Our methodology allowed the combination of such small studies, of lower statistical power and limited generalizability, into a larger, more statistically powerful, more diverse, and clinically relevant meta-analysis.

METHODS

Subjects

We have conducted five consecutive studies that introduced and investigated the use of the EEG Consistency Index, and employed multiple assessments amenable to inclusion within our bio-behavioral assessment of ADHD using the stochastic Bayesian approach to combine various measures of ADHD. These studies are described briefly below.

Study I

Four boys (ages 6–10) with ADHD and four age-matched control boys had their EEG data acquired during two 30-min tasks separated by a 5-min break. For the ADHD boys, this procedure was repeated 3 months later, to assess test-retest reliability. The EEG Consistency Index (CI) was based on this data and calculated from information obtained from four electrode sites, CZ, PZ, P3 and P4. Parents completed the ADHD-Symptom Inventory (ADHD-SI, Cox et al., 1999). Results are reported in Cox et al., 1999.

Study II

Seven ADHD males and six non-ADHD males, ages 18–25, participated in a double-blind, placebo versus methylphenidate controlled crossover design study. ADHD subjects had to have previously taken methylphenidate but could not be taking any medication for their condition within the 6 months prior to the study. EEG data was acquired while the subjects were given four tasks of the Gordon Diagnostic System, two easy (auditory and visual) and two hard (auditory and visual). EEG CI was calculated based upon this data obtained from CZ, PZ, P3 and P4, and the relative power of the frequency bands computed as in Study I. Subjects and their parents completed the ADHD-SI. Results are reported in Merkel et al., 2000 and Cox, Merkel, Kovatchev, & Seward, 2000.

Study III

Eighteen boys and 17 girls, ages 8–16, classified as either ADHD or non-ADHD had EEG data collected for 36 min while performing various tasks. The EEG CI was based on this data and calculated from information obtained from eight electrode sites, CZ, PZ, P3, P4, F3, F4, T3, and T4. The relative power of the frequency bands was computed based on the same thresholds employed in Study I and II. Parents and teachers completed the ADHD-SI. Results are reported in Kovatchev et al., 2001.

Study IV

Six ADHD and six non-ADHD Caucasian college age females engaged in a series of short concentration tasks (2–3 min.) with shorter resting intervals (1–2 min) will EEG data were collected. The CI was calculated from this data, with the CI employing data from eight electrode sites CZ, PZ, P3, P4, F3, F4, T3, and T4. Subjects were administered the

ADHD-SI, and the Wender-Utah Rating Scale (WURS; Ward, Wender, & Reimherr, 1993), which is a 61-item retrospective self-report scale with adequate reliability and validity. Results of analyses of these initial subjects are reported in Robeva et al., 2004.

Study V

Seventy-seven children ages 8–12 were administered EEGs while watching a movie for 20 min, resting with eyes open for 5 min, reading silently for 10 min, resting with eyes open for 5 min, then performing creative drawing tasks for 10 min. This pattern was repeated once, for a total test time of 100 min. The EEG CI was calculated when contrasting EEG during the video and reading and during the reading and divergent thinking tasks, utilizing information collected from CZ, PZ, P3 and P4. Parents and teachers were administered the AD/HD Rating Scale-IV (DuPaul, Power, Anastopoulos, & Reid, 1998), and parents completed the ADHD-SI.

Although these five studies were conducted sequentially over a period of years, each was conducted in the same laboratories and under the supervision of the same group of researchers. A special emphasis was placed on collecting multiple assessments for diagnostic purposes, and each subject in all five studies has at least one biological assessment in the form of EEG data, and at least one ADHD symptom assessment in the form of a completed symptom questionnaire, such as the AD/HD Rating Scale-IV (DuPaul et al., 1998), Wender-Utah rating scales (Ward, Wender, & Reimherr, 1993), or the ADHD-Symptom Inventory (ADHD-SI, Cox et al., 1999). There were, however, differing assessment tools and EEG acquisition procedures utilized in each study, and the total assessment package is not the same for all subjects in all five studies. As stated, the majority of these individual studies had a sample size of 12 or less and although some significant results were obtained, generalizability and validity of each individual study is compromised by small sample sizes.

The inclusion and exclusion criteria were consistent across these studies: Subjects in the above five studies were placed in the ADHD group if they (a) met DSM-IV diagnostic criteria for ADHD (with or without hyperactivity), as assessed by a structured clinical diagnostic interview performed by a licensed clinical psychologist and/or psychiatrist, and in Study III, the structured interview was additionally confirmed by appropriate scores on the Conners' Rating Scale (Conners, 1997), and Achenbach Child Behavior Checklist (Achenbach & Edelbrock, 1983); (b) had a history of a positive response to methylphenidate, whether they were currently on medication for ADHD or not, and no history of adverse reaction to methylphenidate, (c) did not meet diagnostic criteria for current co-morbid psychological disorders of depression, anxiety, psychosis, drug/alcohol abuse or addiction; and (d) did not have any significant health problems or disorders that might effect the brain or EEG recordings (e.g. Tourette's Syndrome, epilepsy).

Participants in the control groups of the above studies were included if they (a) had no known history of ADHD or disruptive behavioral disorders; (b) had never been prescribed stimulant medication nor taken it for recreational purposes; (c) did not meet diagnostic criteria for current co-morbid psychological disorders, including depression, anxiety, psychosis, drug/alcohol abuse or addiction; and (d) did not have any significant health problems or disorders that might effect the brain or EEG recordings (e.g. Tourette's Syndrome, epilepsy), and (e) demonstrated a lack of ADHD symptoms on our screenings. All participants were either 15 or younger, or 17 and older. No subjects were exactly 16 years of age at the time of the assessments.

In all five studies, individuals in the ADHD group were tested off their medication, which was discontinued under the supervision of a psychiatrist, at least 36 hr prior to the EEG testing. EEG acquisition was conducted for all individuals and involved the subjects being seated in front of the computer and an appropriately sized EEG cap placed on their heads. Electrode placement was in accordance with the international 10–20 system, and the electrode sites were prepared with a ground in front of Cz, and a right earlobe reference electrode. For all five of the studies, EEG signals were amplified and processed by the Lexicor Neurosearch-24 system. The EEG data collection used standard settings of a clinical EEG acquisition with 5 K Ohms impedance criterion measured by a Prep-Check electrode impedance meter.

Procedure

To illustrate the advantage of our approach, we combined data from the five studies described above.

Measures for the Meta-Analysis

From each study, we included at least one measure of symptoms of ADHD and the EEG-Consistency Index. The symptom ratings scales included: ADHD-Symptom Inventory (Cox et al., 1999), AD/HD Rating Scale-IV (DuPaul et al., 1998), and Wender-Utah Rating Scale (Ward, Wender, & Reimherr, 1993). The EEG Consistency Index is a previously reported physiological marker of ADHD (EEG-CI, Cox et al., 1999; Merkel et al., 2000; Kovatchev et al., 2001). Information and details about these assessments are listed below.

ADHD Symptom Inventory (ADHD-SI)

The ADHD-Symptom Inventory is an 18-item scale developed from DSM-IV criteria for ADHD and was introduced by Cox et al. (1999). The ADHD-SI has good test-retest reliability ($r = .87$, $p < .005$). In addition, the ADHD-SI correlates highly with the Hyperactivity Index of the Conners' ($r = .72$, $p = .044$), the Attention Problems subscale of the Achenbach's Child Behavior Checklist ($r = .82$, $p = .013$), and the Impulsivity-Hyperactivity subscale of the Conners' ($r = .88$, $p = .004$). Additionally, in previous research the ADHD-SI discriminated between diagnostic status the most strongly ($t = 3.7$, $p < .01$) among the four psychometrics with no overlap in scores (Cox et al., 1999, Merkel et al., 2000).

AD/HD Rating Scale-IV

The AD/HD Rating Scale-IV is similar to the ADHD-SI, both scales being developed independently and concurrently at different laboratories. This rating scale has demonstrated adequate reliability and validity (DuPaul et al., 1998). The scale items reflect the DSM-IV criteria and respondents are asked to indicate the frequency of each symptom on a 4-point Likert scale. The Home and School Versions of the scale both consist of two subscales:

Inattention (nine items) and Hyperactivity-Impulsivity (nine items). The manual provides information regarding the factor analysis procedures to develop the scales, as well as information regarding the standardization, normative data, reliability, validity, and clinical interpretation of the scales (DuPaul et al., 1998).

Wender-Utah Rating Scale (WURS)

The WURS test is a 61-item retrospective self-report scale where individuals rate the severity of ADHD symptoms experienced when they were children using a 5-point Likert scale. For adults, WURS has been shown to be a valid retrospective screening and dimensional measure of childhood ADHD symptoms (Stein, Fischer, & Szumowski, 1999; Stein, Fischer, & Szumowski, 2000), to replicate and correlate with Connors Abbreviated Parent and Teacher Questionnaire and demonstrate internal consistency reliability (Fossati et al., 2001), and to exhibit good construct validity (Weyandt, Linterman, & Rice, 1995).

EEG Consistency Index

The EEG-CI is an EEG-based measure of ADHD (Cox et al., 1999; Kovatchev et al., 2001; Merkel et al., 2000). The CI ranges from 0 to 100%; a CI < 40% indicates ADHD (Kovatchev et al., 2001). The CI of a person is computed using data from two adjacent disparate cognitive tasks. We used our previously published algorithm with threshold parameter of 1.0 and no cutoff (Kovatchev et al., 2001). These settings correspond to the procedures employed by our previous studies (Cox et al., 1999; Kovatchev et al., 2001; Merkel et al., 2000). A brief description of the CI follows. Details can be found in the above referenced articles.

The CI is based on the notion that the EEG data stream can be represented by a three-dimensional numeric array—at any given moment one dimension is frequency of brain waves, another is spatial—the location of the electrode on a subject’s head, and the third is time. ADHD can cause disruption in the frequency or spatial dimension or in both. This disruption is most evident when the tested subject transitions from one cognitive task to another, the two tasks being separated by a rest period of approximately 3–5 min. The transition is deemed “consistent” if the differences between the means of the power spectra from the adjacent tasks shift coherently from low to high or vice versa, for example a consistent transition would mean that most frequency bands and most channels would display similar unidirectional shifts, while an inconsistent transition will result in scattered power changes across the EEG bands and channels.

Given this heuristic explanation, the algorithm for computing the CI works as follows:

- 1) Discrete spectra, including residual power, are calculated for all EEG channels through a standard FFT algorithm. The relative power of the theta, low alpha, high alpha, and beta frequency bands was computed.
- 2) *Power change distances* (PCD) between two contiguous tasks separated by a break (slow forward and fast forward tracing in our study) are computed for each EEG band and channel. Each PCD is normalized using the formula below, where $M1$ and $M2$ are the mean powers at two contiguous tasks, $SD1$ and $SD2$ are their standard

deviation, and $N1$ and $N2$ are the epoch counts at these tasks.

$$PCD = \frac{M1 - M2}{\sqrt{\frac{SD 1^2}{N1} + \frac{SD 2^2}{N2}}}$$

- 3) PCD undergo filtering to eliminate changes below a “noise threshold” of 1.0: The PCD values that are larger by an absolute value than the threshold will be marked with 1 or -1 depending on their direction, while all PCD below threshold will be marked by zero. This procedure transforms the PCD sequence into a sequence of 1, 0, -1 that indicates, for each EEG band and channel, whether a significant power change was observed while the person shifted from the first task to the next. *The CI is defined as the count of non-zero components of this sequence.* The maximum CI equals the number of EEG channels multiplied by the number of EEG bands used during spectrum discretization (the splitting of the power spectrum into frequency bands). For our study, with 8-channel EEG equipment and 4 bands, the CI ranges from 0 to 32. In order to make the results comparable across different experiments, the CI is expressed in terms of percentage from its maximum value.

Standardizing the Scores of Different Tests

In order to integrate the various study data from the psychometric tests and the CI results into a single assessment of ADHD, we first need to standardize the output of these tests. In order to do so, we translate the output of each test into a *probability for ADHD*. Heuristically, justification of this paradigm would be the following: any test, for ADHD or otherwise, produces a score that is contingent upon certain (psychological, biologic, etc.) characteristics of the tested subject. Therefore, a subject with a certain condition (ADHD) is expected to yield a higher (or a lower score depending on the direction of the test), compared to a subject without that condition. However, the relationship between the condition and the test score is not always exact—it may happen that a subject without ADHD scores high on the WURS scale, or vice versa. Thus, this relationship is probabilistic in nature and is best quantified as a conditional probability of earning certain score, given a preexisting condition, which is a number between 0 and 1 (or 0 to 100%), that is one of continuum-many possible values.

During each step of the overall ADHD assessment, each subject receives a certain test score and the magnitude of this score depends on whether the subject has ADHD, as well as on the severity of the disorder. In other words, the probability of earning a certain score depends on the subjects’ condition, ADHD or non-ADHD. In addition, each test has a suggested cutoff value, and scores greater (or lower) than this cutoff value are accepted as indicators of ADHD. Thus, for each test we can define a function that represents the conditional probability of earning the specified score given ADHD. As in (Robeva et al., 2004), we use a linear mapping of a test score into a [conditional] probability ranging from 0 to 1 with the test cutoff value mapped to 0.5, and the test maximal (or minimal) value indicating ADHD mapped to 1. Published diagnostic thresholds or cut-offs are used to produce the piecewise linear and continuous shape for the function. Figure 1 presents two examples of these mappings, where the X axis is the scoring cut-offs for the specified

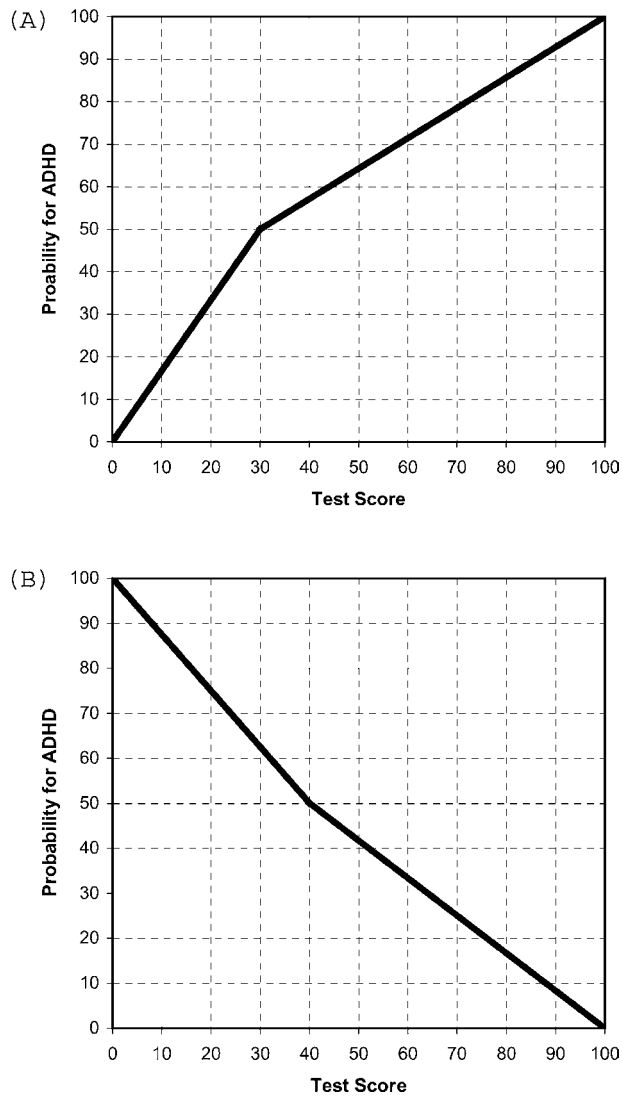


Fig. 1. Examples of test score standardization. (A) Calibration of the WURS Scale. (B) Calibration of the CI.

assessment, and the Y axis is the probability of ADHD (0–100%). The necessity of such published thresholds has made it challenging for us to include the AI of Monastra et al. (1999) in our model, due to unavailability of published clearly defined minimum and maximum values, and a threshold for ADHD. Thus, because the AI is still a relative (to a control group marker) measure, it is not included in the algorithm as such, but could easily be incorporated with more information regarding the defined values and thresholds for scoring.

The computing of the conditional probabilities of Fig. 1 is carried out as follows:

- 1) ADHD-Symptom Inventory: The score on the ADHD-SI ranges from 0 to 36 with scores >12 indicating ADHD. The mapping formula is then:

$$\begin{cases} x \leq 12 & P(\text{ADHD} | x) = \frac{x}{24} \\ x \geq 12 & P(\text{ADHD} | x) = \frac{x}{48} + \frac{1}{4} \end{cases}$$

- 2) AD/HD Rating Scale-IV Inattentive Type: The score ranges from 0 to 100 with scores (>93 indicating ADHD. The mapping formula is then:

$$\begin{cases} x \leq 93 & P(\text{ADHD} | x) = \frac{x}{186} \\ x \geq 93 & P(\text{ADHD} | x) = \frac{x}{14} - \frac{43}{7} \end{cases}$$

- 3) WURS scale: The score ranges from 0 to 100 with scores >30 indicating ADHD. The mapping formula is then:

$$\begin{cases} x \leq 30 & P(\text{ADHD} | x) = \frac{x}{60} \\ x \geq 30 & P(\text{ADHD} | x) = \frac{x}{140} + \frac{2}{7} \end{cases}$$

- 4) The EEG Consistency Index (CI): The Consistency Index ranges from 0 to 100% with a CI $<40\%$ indicating ADHD. The mapping formula is then:

$$\begin{cases} x \leq 40 & P(\text{ADHD} | x) = 1 - \frac{x}{80} \\ x \geq 40 & P(\text{ADHD} | x) = \frac{5}{6} - \frac{x}{120} \end{cases}$$

Once the standardization of measures is completed, a Bayesian algorithm for calculating the combined probability for ADHD for each individual links the results from the individual assessments in each study. We present an outline for the algorithm below. Additional details can be found in Robeva et al., 2004.

The algorithm works as follows: At step 0, a prior probability for ADHD $P_{\text{ADHD}}^0 = 0.5$ is assigned to each subject regardless of whether she is ADHD or control. Then, after the first test $P_1^{\text{test}} = P P(\text{ADHD}/\text{test score})$ and $P_2^{\text{test}} = 1 - P_1$ are used to calculate a posterior probability P_{ADHD}^1 for ADHD, using the formula:

$$P_{\text{ADHD}}^1 = \frac{P_1^{\text{test}} P_{\text{ADHD}}^0}{P_1^{\text{test}} P_{\text{ADHD}}^0 + P_2^{\text{test}} (1 - P_{\text{ADHD}}^0)}$$

From here on the procedure is recursive - after each step the posterior probability becomes a prior probability for the next step; for example in the formula above P_{ADHD}^0 is replaced by P_{ADHD}^1 , P_1^{test} and P_2^{test} derived from the second test, and so forth. In general, the posterior probability from step $(n - 1)$ becomes a prior probability in step (n) and then posterior probability is computed for step (n) using the results from the assessment at step (n) . At each step we may have a “gray zone” of a nondefinitive assessment, however, at each sequential step the gray zone will become smaller and the final result is an assessment that is substantially more precise than any of its individual steps. The final outcome of the Bayesian algorithms is a combined probability for ADHD assigned to each subject (0 to 100%), for example a placement of each subject on a continuum of disruption, with greater number and severity of disruptions resulting in placement on the high extreme end of the continuum. Since the final outcome is based upon the combination of all of the above psycho-physiological data, it has increased specificity/sensitivity beyond any single measure.

Statistical Analysis

T-tests were used to compare the probabilities for ADHD/no ADHD, estimated by each tests, across the ADHD versus No ADHD groups. ANOVA with factors gender and age group was used to compare the probabilities of ADHD across studies.

RESULTS

Probabilities for ADHD Estimated by Each Test Within Each Study

In order to illustrate the calibration process described in the section on “Methods” Table I presents the probabilities for ADHD estimated by each test for ADHD versus Control subjects. T-tests compare these probabilities within each study. As evident from Table I, although a number of tests reach significance in most studies, a correction for the

Table I. Probabilities for ADHD Within Each Study for Each Test

Study		ADHD-SI	WURS	DuPaul (<i>I</i>)	CI	Combined probability
I	ADHD	0.63	—	—	0.79	0.78
	Control	0.04			0.24	0.014
	<i>t, p</i>	0.0146			0	0.0009
II	ADHD	0.7	—	—	0.88	0.93
	Control	0.16			0.57	0.21
	<i>t, p</i>	0.0008			0.095	0.0024
III	ADHD	0.86	—	—	0.81	0.94
	Control	0.19			0.51	0.24
	<i>t, p</i>	0			0.0019	0
IV	ADHD	—	0.56	—	0.83	0.84
	Control		0.18		0.18	0.033
	<i>t, p</i>		0.0029		0.0019	0
V	ADHD	—	—	0.68	0.65	0.74
	Control			0.19	0.54	0.29
	<i>t, p</i>			0	0.19	0

significance level accounting for multiple parallel tests would eliminate many significant results.

The last column of Table I (Combined probability) presents the joint results from ADHD questionnaire plus EEG assessment within each study. The combined probability for ADHD is computed using a previously reported Bayesian algorithm (Robeva et al., 2004). It is evident, that a combination of tests within a study generally provides a better ADHD-control separation than each individual test (in terms of greater *t*-value and smaller *p*-value).

Cross-Sectional Comparisons

We now have the ingredients needed to perform a meta-analysis combining data across studies and across tests: (1) The data from disparate tests within each study are combined into a single probability for ADHD, and (2) The data across studies are mapped (calibrated) onto the same probability scale, which makes them directly comparable. In order to illustrate such a meta-analysis, Tables II and III present the results from ($2 \times 2 \times 2$) ADHD versus control, gender, and age-group cross-sectional comparisons (ANOVA with 7,103 degrees of freedom).

Specifically, Table II presents the means and the standard deviations for each cross-section of the combined across studies population. It is evident that the control population is quite homogeneous—there are no groups different in terms of their probabilities for ADHD. The ADHD population, however is not internally homogeneous: while boys and girls under the age of 16 are similar, and girls under 16 are similar to boys above 16, girls above 16 stand alone as a separate subgroup, which is somewhat closer to controls in terms of its probability for ADHD.

Table III presents the significance of the main effects in the analysis, ADHD versus Control, Gender, and Age group, as well as their first-order interactions. As seen from Table III, the ADHD-control differentiation is much stronger in the combined data than it is in each separate study (Table I). In addition, we observe several significant interaction effects, which was not possible within each of the separate studies.

Finally, Figs. 2A–2C present the process of increasing separation of ADHD and control subjects along the steps of the Bayesian model. Each figure represents the distribution of the subjects' probabilities for ADHD in the ADHD versus control group. In Fig. 2A, there is a significant number of unclassified cases (e.g., subjects with probability for ADHD of 0.5) and a number of cases that are not clearly classified as ADHD or control. With application of more tests, the number of unclassified cases decreases and the number of clearly classified cases increases (the extreme left and right columns). Thus, with the application of consecutive tests the differentiation between ADHD and non-ADHD control groups increases, achieving a best separation at the end of the procedure (Fig. 2C).

DISCUSSION

Diagnosing ADHD presents a challenge to traditional assessment paradigms because there is no single assessment tool or medical test that definitively establishes its presence (Hinshaw, 1994). Instead, there are multiple tests of varying design, each of which has its own administration, scoring system, and diagnostic criteria. Unfortunately, none of these

Table II. Cross-Sectional Probabilities for ADHD and Cross-Sectional Similarities

	ADHD				Control			
	Age < 16 years		Age > 16 years		Age < 16 years		Age > 16 years	
	Boy	Girl	Boy	Girl	Boy	Girl	Boy	Girl
Probability for ADHD Mean(SD)	0.80 (0.26)	0.87 (0.08)	0.93 (0.15)	0.67 (0.35)	0.23 (0.30)	0.32 (0.19)	0.21 (0.36)	0.30 (0.36)

Table III. Significance of Main Effects and First-Order Interactions

	ADHD-Control	Gender	Age group
Main Effect	$F = 119.4$ ($p = 0.0001$)	$F = 0.096$ ($p = 0.76$)	$F = 0.57$ ($p = 0.45$)
First-order interactions			
ADHD-control		$F = 0.014$ ($p = 0.91$)	$F = 2.66$ ($p = 0.11$)
Gender	$F = 0.04$ ($p = 0.91$)		$F = 2.82$ ($p = 0.096$)
Age group	$F = 2.66$ ($p = 0.11$)	$F = 2.82$ ($p = 0.096$)	

individual assessments has proven to be 100% accurate in diagnosing ADHD. This is to be expected, however, since ADHD is considered to be a physiologically based disorder with a multi-factorial etiology that includes neurobiology as an important factor, and would not be easily classified by only one assessment tool. In fact, the reliability of the ADHD diagnosis based on one method or test alone is quite low, and lower still when chance agreement is considered. For example, previous research has found 78% agreement between a structured interview and a discharge diagnosis of ADHD (Welner, Reich, Herjanic, & Jung, 1987) and 70 to 80% accuracy (with considerable variation depending on age range) of laboratory measures of attention in correctly predicting an ADHD diagnosis (Fischer, Newby, & Gordon, 1995).

What is needed is a methodology for combining disparate assessments and tests in order to not only provide a more accurate diagnosis, but to also enable the combination of multiple studies of ADHD assessments, increasing thus the sample size and providing more power, generalizability, and possibilities for cross-sectional comparisons. Similarly, researchers continue to look for novel methods to classify and predict diagnoses of ADHD in the fields of imaging and genetics that will more closely link assessment data with underlying neurobiological markers (Castellanos & Tannock, 2002). This manuscript outlines a theoretical framework for performing meta-analysis of data obtained from disparate studies based on calibration of the data from such studies into a single probability scale. We illustrate this framework with existing data from a series of smaller studies examining the diagnosis of ADHD, but this method could be applied to any set of studies where one wishes to combine data.

Each of the five studies we combined used a behavioral assessment of ADHD, such as ADHD-SI, AD/HD Rating Scale-IV, Wender-Utah Rating Scale, and an EEG assessment of ADHD—the CI. These two assessments were joined, within each study, using a previously reported Bayesian algorithm (Robeva et al., 2004), resulting in a combined probability for ADHD for each subject. In general, this combined probability presents a better assessment of ADHD than each of the separate tests it includes. Such a procedure is especially useful in situations such as diagnosing ADHD, when there is no single conclusive assessment, but rather a number of imperfect tests that marginally address the outcome of interest, and where researchers may have multiple related tests performed on a single subject, which they wish to combine into a more comprehensive assessment of this subject.

Equally important, once the data output from each study is standardized, the data can be combined across different studies. In addition to significantly increasing the sample size, such approach allows the data to be examined in subgroups divided by age and/or gender, diagnostic group, and so forth. In our example, in all five studies, the subjects were classified into groups of ADHD (Total $N = 56$) versus Control (Total $N = 55$). However, different studies focused on different age and gender groups. The standardization of the

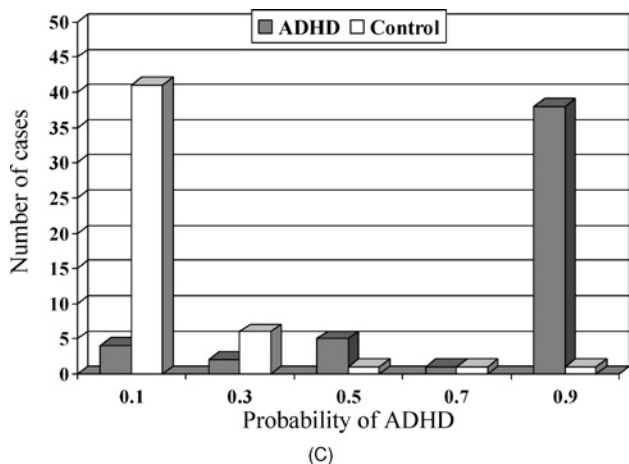
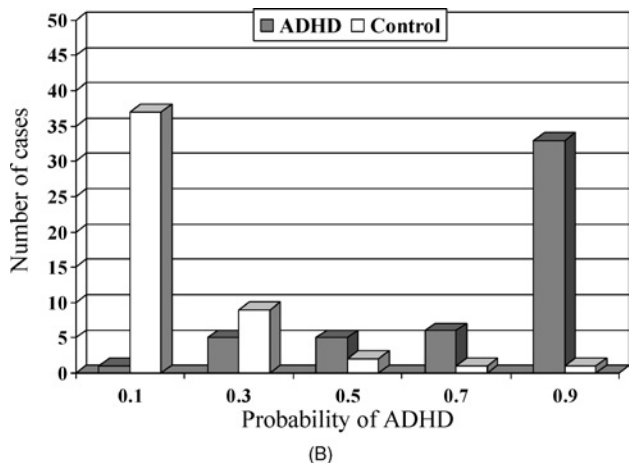
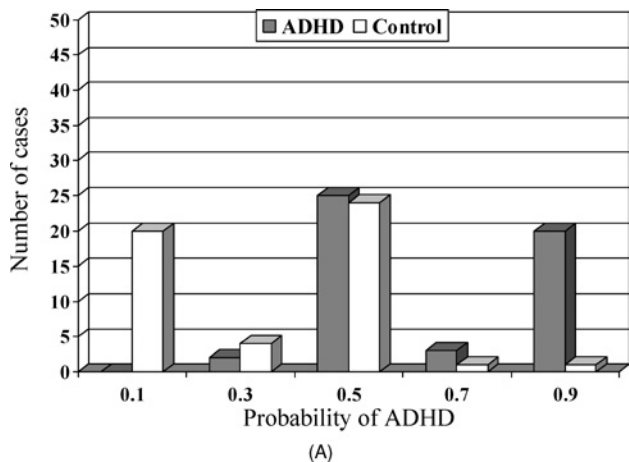


Fig. 2. (A) Subject distribution after Test 1. (B) Subject distribution after Test 1 + 2. (C) Subject distribution after Test 1 + 2 + 3.

data allowed cross-sectional comparisons, which were not possible with the original data. For example we found cross-sectional differences and similarities among the various age and gender groups between control and ADHD. Although all ages and genders of the control groups appear quite homogeneous—there are no groups different in terms of their probabilities for ADHD, the ADHD groups, while all demonstrating probability levels consistent with ADHD, vary by gender and age in the exact probability level. Specifically, it is of interest that the boys and girls under the age of 16 are similar in their probabilities levels, and girls under 16 are similar to boys above 16, girls above 16 stand alone as a separate subgroup, which is somewhat closer to controls in terms of its probability for ADHD. This may be reflective of a subgroup of ADHD or may reflect differences in maturation (Clark et al., 2001). While developmental sex differences have been reported in the imaging literature (Giedd, Castellanos, Rajapakse, Vaituzis, & Rapoport, 1997), studies investigating gender differences in ADHD report no significant differences between boys and girls on symptomatology, neurobiology, or response to stimulant medication (Castellanos et al., 2002; Castellanos et al., 2000; Castellanos et al., 2001; Sharp et al., 1999).

The proposed assessment does not aim at replacing any established practices for screening and diagnosing of ADHD but instead at demonstrating that the outcomes of related studies can be combined in a manner that allows meta-analysis of different types of data which may not be collected in the same manner in each study, and which can include physiological data as well as symptom reports.

A major limitation of the illustration of meta-analysis is the small sample size of males older than 16 years of age, and females of any age. Despite the small sample size in these categories, however, all presented results are statistically significant. Overall, results of this meta-analysis appear to confirm that when analyzed separately, each individual study included in the meta-analysis has lower power and less statistical ability to determine accurate diagnosis of subjects, as well as poor generalizability. However, when combined, these studies provide an improved ability to accurately diagnose ADHD, as well as significantly increased power and generalizability.

It should be emphasized that the application of the proposed meta-analysis tool is not limited to the specific tests used in the studies presented—the meta-analysis is capable of accommodating a variety of ADHD or other tests. Specifically, almost any individual test or assessment could be employed within this model, assuming that the output of such test can be standardized into probabilities for the specific disorder or disease. As such, this meta-analysis procedure may provide a much-needed tool for combining related studies with similar or disparate tests and assessments in a number of research areas, which may otherwise have small, less generalizable studies of limited power and utility.

ACKNOWLEDGMENTS

This study was supported by a grant “Quantifying Cognitive & Attentional Impairments: A Psycho-Physiological Procedure” from the Carilion Biomedical Institute, Roanoke, Virginia and by grant J-451 “Encephalographic and Psychometric Differences Between Female College Students With and Without Attention Deficit/Hyperactivity Disorder” from the Thomas F. Jeffress and Kate Miller Jeffress Memorial Trust, Richmond, VA.

REFERENCES

- Achenbauch, T. M., & Edelbrock, C. (1983). *Manual for the child behavior checklist*. Burlington, VT: University of Vermont Department of Psychiatry.
- Almond, B. W.Jr., Tanner, J. L., & Goffman, H. F. (1999). *The family is the patient: Using family interviews in children's medical care* (2nd ed., 307–313). Baltimore, MD: Williams & Wilkins.
- American Academy of Pediatrics. (2000). Clinical practice guideline: Diagnosis and evaluation of the child with Attention-Deficit/Hyperactivity Disorder. *Pediatrics*, *105*(5), 1158–1170.
- American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1987). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., Revised). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Barkley, R. A. (1990). A critique of current diagnostic criteria for attention deficit hyperactivity disorder: clinical and research implications. *Journal of Developmental and Behavioral Pediatrics*, *11*, 343–352.
- Barkley, R. A., Fischer, M., Edelbrock, C. S., & Smallish, L. (1990). The adolescent outcome of hyperactive children diagnosed by research criteria: I. An 8-year prospective follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 546–557.
- Barkley, R. A., Guevremont, D. C., Anastopoulos, A. D., DuPaul, G. J., & Shelton, T. L. (1993). Driving-related risks and outcomes of attention deficit hyperactivity disorder in adolescents and young adults: A 3- to 5-year follow-up survey. *Pediatrics*, *92*, 212–218.
- Berquin, P. C., Giedd, J. N., Jacobsen, L. K., Hamburger, S. D., Krain, A. L., Rapoport, J. L., & Castellanos, F. X. (1998). Cerebellum in attention-deficit hyperactivity disorder: A morphometric MRI study. *Neurology*, *50*, 1087–1093.
- Bresnahan, S. M., Anderson, J. W., & Barry, R. J. (1999). Age-related changes in quantitative EEG in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *46*, 1690–1697.
- Castellanos, F. X. (1997). Toward a pathophysiology of attention-deficit/hyperactivity disorder. *Clinical Pediatrics (Phila)*, *36*, 381–393.
- Castellanos, F. X., Marvasti, F. F., Ducharme, J. L., Walter, J. M., Israel, M. E., Krain, A., Pavlovsky, C., & Hommer, D. W. (2000). Executive function oculomotor tasks in girls with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(5), 644–650.
- Castellanos, F. X., Giedd, J. N., Berquin, P. C., Walter, J. M., Sharp, W., Tran, T., Vaituzis, A. C., Blumenthal, J. D., Nelson, J., Bastain, T. M., Zijdenbos, A., Evans, A. C., & Rapoport, J. L. (2001). Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, *58*(3), 289–295.
- Castellanos, F. X., Lee, P. P., Sharp, W., Jeffries, N. O., Greenstein, D. K., Clasen, L. S., Blumenthal, J. D., James, R. S., Ebens, C. L., Walter, J. M., Zijdenbos, A., Evans, A. C., Giedd, J. N., & Rapoport, J. L. (2002). Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Medical Association*, *288*(14), 1740–1748.
- Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience*, *3*, 617–628.
- Clark, A. R., Barry, R. J., McCarthy, R., & Selikowitz, M. (2001). Age and sex effects in the EEG: differences in two subtypes of attention-deficit/hyperactivity disorder. *Psychophysiology*, *38*, 212–221.
- Conners, C. K. (1997). *Conners' Rating Scales - Revised; technical manual*. North Tonawanda, NY: Multi-Health Systems.
- Cox, D. J., Kovatchev, B. P., Morris, J. B., Phillips, C., Hill, R., & Merkel, L. (1999). Electroencephalographic and psychometric differences between boys with and without Attention-Deficit/Hyperactivity Disorder (ADHD): A pilot study. *Applied Psychophysiology and Biofeedback*, *23*, 179–188.
- Cox, D. J., Merkel, R. L., Kovatchev, B., & Seward, R. (2000). Effect of stimulant medication on driving performance of young adults with attention-deficit hyperactivity disorder: A preliminary double-blind placebo controlled trial. *Journal of Nervous and Mental Disorders*, *188*, 230–234.
- Crawford, H., & Barabasz, M. (1996). Quantitative EEG magnitudes in children with and without attention deficit disorder during neurological screening and cognitive tasks. *Child Study Journal*, *26*, 71–86.
- DuPaul, G. J., Power, T. J., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV Checklists, norms, and clinical interpretation*. New York: Guilford Press.
- Fischer, M., Newby, R. F., & Gordon, M. (1995). Who are the false negatives on Continuous Performance Tests? *Journal of Clinical Child Psychology*, *24*, 427–433.
- Fossati, A., Di Ceglie, A., Acquarini, E., Donati, D., Donini, M., Novella, L., & Maffei, C. (2001). The retrospective assessment of childhood attention-deficit hyperactivity disorder in adults: Reliability and validity of the Italian version of the Wender Utah Rating Scale. *Comprehensive Psychiatry*, *42*, 326–336.

- Giedd, J. N., Castellanos, F. X., Rajapakse, J. C., Vaituzis, A. C., & Rapoport, J. L. (1997). Sexual dimorphism of the developing human brain. *Prog Neuropsychopharmacol Biol Psychiatry*, 21(8), 1185–1201.
- Heilman, K. M., Voeller, K. K., & Nedeau, S. E. (1991). A possible pathophysiologic substrate of attention deficit hyperactivity disorder. *Journal of Child Neurology*, 6, Suppl., S76–S81.
- Hinshaw, S. P. (1994). *Attention deficits and hyperactivity in children*. Thousand Oaks, CA, Sage.
- Hynd, G. W., Semrud-Clikeman, M., Lorys, A. R., Novey, E. S., Eliopoulos, D., & Lytinen, H. (1991). Corpus callosum morphology in attention-deficit-hyperactivity disorder: Morphometric analysis of MRI. *Journal of Learning Disabilities*, 24, 141–146.
- Jensen, P. S., Koretz, D., Locke, B. Z., Schneider, S., Radke-Yarrow, M., Richters, J. E., & Rumsey, J. M. (1993). Child and adolescent psychopathology research: Problems and prospects for the 1990s. *Journal of Abnormal Child Psychology*, 21, 551–580.
- Kovatchev, B. P., Cox, D. J., Hill, R., Reeve, R., Robeva, R. S., & Loboschewski, T. (2001). A psychophysiological marker of Attention Deficit/Hyperactivity Disorder- Defining the EEG consistency index. *Applied Psychophysiology and Biofeedback*, 26, 127–139.
- Kuperman, S., Gaffney, G. R., Hamdan-Allen, G., Preston, D. F., & Venkatesh, L. (1990). Neuroimaging in child and adolescent psychiatry. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 159–172.
- Lazzaro, I., Gordon, E., Whitmont, S., Plahn, M., Li, W., Clarke, S., Dosen, A., & Meares, R. (1998). Quantified EEG activity in adolescent attention deficit hyperactivity disorder. *Clinical Electroencephalography*, 29, 37–42.
- Mann, C. A., Lubar, J. F., Zimmerman, A. W., Miller, C. A., & Muenchen, R. A. (1992). Quantitative analysis of EEG in boys with attention-deficit-hyperactivity disorder: Controlled study with clinical implications. *Pediatric Neurology*, 8, 30–36.
- Merkel, R. L., Cox, D. J., Kovatchev, B. P., Morris, J., Seward, R., Hill, R., & Reeve, R. (2000). The EEG consistency index as a measure of Attention Deficit/Hyperactivity Disorder and responsiveness to medication: A double blind placebo controlled pilot study. *Applied Psychophysiology and Biofeedback*, 25, 133–142.
- Monastra, V. J., Lubar, J. F., Linden, M., VanDeusen, P., Green, G., Wing, W., Phillips, A., & Fenger, T. N. (1999). Assessing attention deficit hyperactivity disorder via quantitative electroencephalography: An initial validation study. *Neuropsychology*, 13, 424–433.
- Monastra, V. J., Lubar, J. F., & Linden, M. (2001). The Development of a quantitative electroencephalographic scanning process for attention deficit-hyperactivity disorder: Reliability and validity studies. *Neuropsychology*, 15(1), 136–144.
- Mostofsky, S. H., Reiss, A. L., Lockhart, P., & Denckla, M. B. (1998). Evaluation of cerebellar size in attention-deficit hyperactivity disorder. *Journal of Child Neurology*, 13, 434–439.
- Pope, A. T., & Bogart, E. H. (1991). Extended attention span training. *Proceedings of Technology 2001 Conference*, 1, 368–304.
- Pope, A. T., Comstock, J. R., Bartolome, D. S., Bogart, E. H., & Burdette, D. W. (1994). Biocybernetic system validates index of operator engagement in automated tasks. In M. Mouloua, & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 300–306), Hillsdale, NJ, Lawrence Erlbaum Associates.
- Raskin, L. A., Shaywitz, S. E., Shaywitz, B. A., Anderson, G. M., & Cohen, D. J. (1984). Neurochemical correlates of attention deficit disorder. *Pediatric Clinics of North America*, 31, 387–396.
- Robeva, R., Penberthy, J. K., Loboschewski, T., Cox, D., & Kovatchev, B. (2004). Sequential Psycho-Physiological Assessment of ADHD: A Pilot Study of Bayesian Probability Approach Illustrated by Appraisal of ADHD in Female College Students. *Applied Psychophysiology and Biofeedback*, 29(1), 1–18.
- Schachar, R., Taylor, E., Wieselberg, M. B., Ghorley, G., & Rutter, M. (1987). Changes in family functioning and relationships in children who respond to methylphenidate. *Journal of the American Academy of Child and Adolescent Psychiatry*, 26, 728–732.
- Sharp, W. S., Walter, J. M., Marsh, W. L., Ritchie, G. F., Hamburger, S. D., & Castellanos, F. X. (1999). ADHD in girls: Clinical comparability of a research sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(1), 40–47.
- Stein, M. A., Fischer, M., & Szumowski, E. (1999). Evaluation of adults for ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(1), 940–941.
- Stein, M. A., Fischer, M., & Szumowski, E. (2000). Evaluation of adults for ADHD: Erratum. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 674.
- Ward, M. F., Wender, P. H., & Reimherr, F. W. (1993). The Wender Utah Rating Scale: An aid in the retrospective diagnosis of childhood attention-deficit-hyperactivity disorder. *American Journal of Psychiatry*, 150, 885–890.
- Welner, Z., Reich, W., Herjanic, B., & Jung, K. G. (1987). Reliability, validity, and parent-child agreement studies of the Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry*, 26(5), 649–653.

- Weyandt, L. L., Linterman, I., & Rice, J. A. (1995). Reported prevalence of attentional difficulties in a general sample of college students. *Journal of Psychopathological and Behavioral Assessment, 17*, 293–304.
- Zametkin, A. J., Nordahl, T. E., Gross, M., King, A. C., Semple, W. E., Rumsey, J., Hamburger, S., & Cohen, R. M. (1990). Cerebral glucose metabolism in adults with hyperactivity of child onset. *New England Journal of Medicine, 323*, 1361–1366.

Copyright of Applied Psychophysiology & Biofeedback is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.