Comparison of a standard psychiatric evaluation to rating scales and EEG in the differential diagnosis of attention-deficit/hyperactivity disorder

Humberto Quintana a,*, Steven M. Snyder b, William Purnell a, Carolina Aponte a, Janis Sita a

a Department of Psychiatry, Louisiana State University Health Sciences Center, 1542 Tulane Avenue, Room 314a, New Orleans, LA 70112, USA
b Department of Psychology, University of North Texas Health Science Center at Fort Worth, Fort Worth, TX 76107, USA

Received 22 April 2005; received in revised form 11 November 2005; accepted 20 April 2006

Abstract

The objective was to investigate the effectiveness of rating scales and electroencephalography (EEG) in detecting the presence of attention-deficit/hyperactivity disorder (ADHD) within a diverse clinical sample. A standard psychiatric evaluation was used to assess 26 children/adolescents who presented to a clinic because a parent suspected the presence of ADHD. EEG data was collected in a blinded protocol, and rating scales were collected as well. Although all subjects had presented with ADHD-like symptoms, only 62% were diagnosed with ADHD, while the remaining 38% had other disorders or no diagnosis. Rating scales readily classified inattentive, impulsive, and/or hyperactive symptoms as being due to ADHD, regardless of the actual underlying disorder, leading to a sensitivity of 81% and a specificity of 22%. Previous studies have observed that there is an EEG marker that identifies ADHD vs. controls, and this marker was present in 15 out of 16 of the ADHD subjects (sensitivity = 94%) and in none of the subjects with ADHD-like symptoms due to other disorders (specificity = 100%). In the detection of ADHD in a diverse clinical sample, rating scales and EEG were both sensitive markers, whereas only EEG was specific. These results may have important implications to ADHD differential diagnosis.

© 2006 Elsevier Ireland Ltd. All rights reserved.

Keywords: Attention-deficit/hyperactivity disorder; Children; Adolescents; Psychometrics; Electroencephalography; Diagnosis

1. Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common disorder of childhood and adolescence with a conservatively estimated prevalence of 3–6% (Barkley, 1990; Bradley and Golden, 2001; Goldman et al., 1998). ADHD is generally diagnosed through the identification of symptoms of inattention, hyperactivity, and impulsivity based on criteria of the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) (APA, 1994). The fact that many ADHD symptoms are common to other psychiatric disorders as well as to regular childhood/adolescent behavior creates a challenging scenario for the diagnosis of ADHD (Munoz-Millan and Casteel, 1989). Indeed, patients who present to clinics with purported ADHD-like symptoms stand a significant chance of expressing a disorder other than ADHD, such as oppositional defiant disorder (ODD), an anxiety disorder, conduct disorder (CD), a mood disorder, adjustment disorder,
reading disorder, and dyslexia (Cantwell, 1996; Goldman et al., 1998; Munoz-Millan and Casteel, 1989; Pary et al., 2002; Rucklidge and Tannock, 2002). As such, the ruling out of other psychiatric disorders remains a necessity in the accurate diagnosis of ADHD (APA, 1994; Cantwell, 1996; Paule et al., 2000).

While professional guidelines for ADHD diagnosis recommend a thorough examination to determine the possibility of alternative diagnoses, the time necessary for such an evaluation may be one limiting factor in the clinical setting. Jensen (2000) reported that although ADHD can be rigorously and reliably diagnosed under optimal conditions, such best practices do not appear to be taking place in the real world. In order to provide an accurate diagnosis, the current study required a 2- to 3-h standard psychiatric evaluation. In contrast, primary care physicians who provide care to the majority of the children/adolescents with mental health concerns are typically allotted about 15 min of time with each patient (Mechanic et al., 2001). Goldman et al. (1998) state that it is clear that ADHD cannot be diagnosed in a typical 15-min primary care office visit. Compounding these circumstances is the fact that physicians are expected to provide this service with minimal formal training in mental health evaluation relative to that of psychiatrists and clinical psychologists. For instance, the American Academy of Pediatrics recommends following DSM-IV criteria in the diagnosis of ADHD (AAP, 2000), yet physicians have reported that they use DSM-IV criteria for only 28–38% of patients with problems of attention and hyperactivity (Chan et al., 2002; Wasserman et al., 1999). Faced with constraints of time and experience, physicians have reported that they often rely on behavior rating scales, with 70% using ADHD specific rating scales and 60% using global rating scales (Chan et al., 2002).

Rating scales can be useful in the identification of symptoms of certain disorders, such as ADHD (Bussing et al., 1998; Vaughn et al., 1997). Rating scales are widely used by clinicians to guide parents and teachers in the identification of ADHD symptoms occurring in the home and school settings (Doyle et al., 1997). Professional guidelines, such as those offered by the American Academy of Pediatrics, recommend rating scales as one of many possible means of determining if the child is expressing ADHD symptoms (AAP, 2000). Numerous studies have been conducted to establish the validity and reliability of ratings scales as an assessment tool for ADHD as covered in a recent 10-year review (Collett et al., 2003). Of most clinical relevance were those studies which tested the diagnostic efficiency of the rating scales. The ADHD-IV rating scale was evaluated with an overall accuracy in the range of 64–77% for ADHD vs. controls (DuPaul et al., 1998). The Conners’ rating scales were evaluated with an experimental design using discriminant analysis for a reported accuracy of 83–93% for ADHD vs. controls (Conners et al., 1998a,b, 1997). However the appropriate use of the discriminant analyses in these studies has been called into question (Snyder et al., 2004). A broad review of the literature demonstrates that when taking statistical methods and experimental designs of these studies into consideration, the expected range of accuracy for rating scales is 55–79% in the identification of ADHD vs. controls (Bussing et al., 1998; Doyle et al., 1997; DuPaul et al., 1998; Eiraldi et al., 2000; Luk and Leung, 1989; McCann et al., 2000; Rucklidge and Tannock, 2002; Snyder, 2004; Sprafkin et al., 2002; Vaughn et al., 1997).

One potential limitation in the accuracy of rating scales is that the outcome can be significantly influenced by the bias of the informant (Collett et al., 2003; Doyle et al., 1997; Eiraldi et al., 2000). A telling sign is that there is commonly disagreement between the results of teacher and parent scales, with frequent discrepancies having been reported (AAP, 2000). Another potential shortcoming of rating scales is that the behavioral symptoms identified by the scales are not necessarily specific to ADHD, but instead are common to numerous disorders (Collett et al., 2003; Rucklidge and Tannock, 2002; Vaughn et al., 1997). Much of the previous validation of rating scales has been limited to the identification of ADHD subjects against asymptomatic subjects. Such an experimental design may not reveal the performance of the scales in the clinical setting where non-ADHD patients might often express ADHD-like symptoms that test positive on the scales (Green et al., 1999).

While useful to an informed diagnosis, behavior rating scales provide a diagnostic accuracy that is limited by subjectivity and informant bias, as well as lack of support for specificity in clinical applications. Given the variety of disorders that express ADHD-like symptoms, differential diagnosis in the clinical setting would benefit from the improved accuracy that may be offered by an objective diagnostic aid that is specific to ADHD and free of informant bias. Such a tool has been sought in the realms of genetic and physiological factors (Bradley and Golden, 2001; Cantwell, 1996; Green et al., 1999). One possibility lies in the electroencephalographic (EEG) marker of brain electrical activity in ADHD subjects (Pary et al., 2002). The EEG classification technique is based on the observation of significant differences that ADHD subjects exhibit in their brain wave activity relative to normal subjects. The brain wave activity is monitored non-invasively as changes in electrical potential at the surface of the scalp. EEG digitization
and analysis is then used to extract the complex information within the brain data.

Recent reviews have reported numerous studies which support that EEG results have been significantly associated with the presence of ADHD (Barry et al., 2003; Chabot et al., 2005). Of particular note are those studies which examined the direct application of EEG in the identification of ADHD. The results of these studies support that EEG can be used to identify ADHD with reasonable accuracy (77% to 96%) when utilized vs. asymptomatic controls and learning disorders (Chabot et al., 1996; Chabot and Serfontein, 1996; Clarke et al., 2002b; Kovatchev et al., 2001; Mann et al., 1992; Monastra et al., 2001, 1999).

Previous criticisms of the EEG assessment of ADHD include claims of insufficient sensitivity and specificity (Levy and Ward, 1995) and inconsistent results between previous studies (AAP, 2000). However, results from numerous recent studies support the use of variables of EEG theta and beta power for identifying DSM-IV ADHD (Bresnahan et al., 1999; Clarke et al., 1998, 2001a,b,c,d, 2002a,b; El-Sayed et al., 2002; Lazzaro et al., 1999, 1998; Monastra et al., 2001, 1999). In one study of children determined to have ADHD by a standard diagnostic protocol, 86% were correctly classified using the theta/beta ratio with a set cutoff between populations (Monastra et al., 1999). Ninety-eight percent (98%) of the asymptomatic children were also correctly identified by this scheme. In a study to cross-validate the variable and cutoff with a fresh sample of subjects, 90% sensitivity and 94% specificity were observed (Monastra et al., 2001). In several studies, more elaborate statistical analyses demonstrated that the fraction of DSM-IV ADHD subjects who did not show an increase in the theta/beta ratio instead showed an increase in frontal beta power (Chabot and Serfontein, 1996; Clarke et al., 1998, 2001b,c,d, 2002b).

Given the favorable results of the theta/beta ratio when applied to ADHD vs. controls, the authors chose to replicate the use of the theta/beta ratio and investigate the consistency when applied to a diverse clinical population. Frontal beta power was investigated as well to determine if a fraction of the ADHD subjects was identifiable by this variable.

Criticisms of research designs covering the use of EEG in the assessment of ADHD are summarized in the following claims from the American Academy of Neurology and the American Clinical Neurophysiology Society (AAN/ACNS) (Nuwer, 1997): (1) no blinded comparisons have been made with a clinical standard, and (2) many studies do not use an appropriate spectrum of patients for whom the diagnostic tests would be applied in clinical practice. The authors agree that the AAN/ACNS offer valid criticisms and the appropriate research must be performed to support clinical applications. It is worth noting that the AAN/ACNS criticisms also apply to ADHD rating scales research, as demonstrated by the recent 10-year review of rating scales (Collett et al., 2003) in which studies were limited to comparisons of ADHD vs. controls with no blinding in the protocols.

The current study, as designed, promises to respond to the AAN/ACNS criticisms for both EEG and rating scales in the assessment of ADHD. Important features of the experimental design are as follows. Previously validated variables and cutoffs of EEG and rating scales have been selected for replication vs. a clinical sample. The sample included all subjects over a set time period who presented to a child psychiatric clinic with suspected ADHD-like symptoms. Subjects with comorbid conditions were included. The clinical standard involved a 2- to 3-hour standard psychiatric evaluation to determine whether the symptoms were actually due to ADHD or due instead to other disorders with similar symptoms. Blinded comparisons were performed for both the EEG results and the rating scales vs. the clinical standard.

As such, the study is intended to (1) provide an experimental design that responds to criticisms from the AAP and the AAN/ACNS concerning the use of EEG in ADHD assessment, (2) offer a view of the accuracy of EEG and ratings scales when applied in clinical practice, and (3) provide direction as to whether a large sample investigation is warranted.

2. Methods

2.1. Subjects

Twenty-six subjects participated in this study. Following the psychiatric evaluation, 16 were diagnosed with ADHD and 10 were classified as non-ADHD which included no diagnosis or other disorders not ADHD. The subjects ranged in age from 6 to 21 years (mean = 10.5 years, S.D. = 3.7 years) with 20 children, 5 adolescents, and 1 young adult. Participants were identified as African American (n = 4, 15.4%), Asian American (n = 1, 3.8%), Caucasian (n = 20, 76.9%), and Middle Eastern (n = 1, 3.8%). The sample included 3 girls and 23 boys.

The sample was recruited at a child psychiatric clinic at the Louisiana State University Health Sciences Center from April 8, 2004 to December 22, 2004. Subjects were included in the study if they presented to the clinic because a parent and/or school official suspected the child/adolescent might have ADHD. All those subjects suspected of having ADHD with/without associated
disorders or co-morbidities (such as anxiety, depression, oppositional defiant disorder or conduct disorder) were included in the study. Participants were not on any psychiatric medications at the time of the study nor had they taken any psychiatric medications in the 6 months prior to the study. Exclusion criteria included a history of seizure disorder or EEG abnormalities, known serious medical problems, previous psychiatric hospital admissions, metal plate or metal device in the head, or suicidal or homicidal ideation. After receiving a complete description of the study, parents signed a consent form and subjects signed an assent form as minors or a consent form when older than 16 years of age. The protocol was approved by the Institutional Review Board of the Louisiana State University Health Sciences Center.

2.2. Psychiatric evaluation

The psychiatric evaluation consisted of a semi-structured interview (Schedule for Affective Disorders and Schizophrenia–Lifetime Version (K-SADS-PL) and K-SADS-PL Supplement for Behavioral Disorders), Clinical Global Assessment Scale (CGAS), and the Clinical Global Impression–Severity subscale (CGI-S). The K-SADS-PL Supplements of Anxiety Disorders and Affective Disorders were completed when specified by the results of the K-SADS-PL Screen. The K-SADS-PL and Supplements were conducted by a board-certified psychiatrist and psychiatric research fellows interviewing both the parent and the child/adolescent. Blinded to the EEG results, the lead psychiatrist compiled the information of the psychiatric evaluation and performed the diagnosis of each subject, designating ADHD or non-ADHD, as well as a complete differential diagnosis for the presence of other disorders.

The board-certified psychiatrist followed a set procedure for determining diagnoses from the psychiatric evaluation materials. The CGAS and CGI-S were included to provide information on impairment, functioning, and severity of the disorder. CGAS/CGI-S were reviewed to confirm the presence of impaired functioning per DSM-IV requirements. A CGAS rating score of less than 60 indicated impairment in daily functioning. A CGI-S score greater than or equal to 3 indicated moderate illness. The K-SADS-PL Unstructured Interview assisted in confirming impairment in functioning and the severity of illness. This interview provided information such as signs, symptoms and medical/psychiatric history. For instance with an ADHD diagnosis per DSM-IV criteria, the K-SADS-PL Unstructured Interview was examined to ensure that onset of illness was before 7 years of age and the duration of illness was at least 6 months. Furthermore, the presence of ADHD had to be the primary diagnosis, necessitating that ADHD represented the main problem requiring the focus of treatment, and that ADHD onset preceded the presence of any other disorder. A review of the K-SADS-PL screen and K-SADS-PL Behavioral Disorders Supplement identified the category of the patient’s problems for use towards differential diagnosis. The review, based on DSM-IV diagnostic criteria, determined the presence or absence of a disruptive disorder. The K-SADS-PL Screen had to be scored as 3 (meeting threshold-irritable, angry daily, or almost daily, at least 50% of the time). For ADHD diagnosis with DSM-IV criteria, the K-SADS-PL Behavioral Disorders Supplement had to display six (or more) items in the inattention category, and/or six (or more) items in the hyperactivity or impulsivity category. Both the K-SADS-PL Screen and the Behavioral Disorders Supplement had to agree, in order to confirm the presence of ADHD. Under similar considerations, the K-SADS-PL Anxiety Disorder Supplement was used to identify whether anxiety symptoms were present, and the Affective Disorder Supplement was used to identify whether affective symptoms were present. In 4 of the 26 cases, a complete diagnosis required a further evaluative conference of the psychiatrist and research fellows following the psychiatrist’s review of the collected materials.

2.3. Rating scales

The Attention-Deficit/Hyperactivity Disorder Rating Scale, Version-IV (ADHD-IV) is a narrow band rating scale that assists with the measurement of inattentive, hyperactive, and impulsive symptoms of ADHD. In a recent 10-year review of behavior rating scales, the ADHD-IV was the only scale covered that offered criterion validity results for ADHD assessment based on accepted experimental designs and statistical techniques (Collett et al., 2003). Therefore the ADHD-IV was selected as the representative rating scale for the current study so that the previous ADHD vs. control criterion validity results could be compared to the ADHD differential diagnosis validation of the current study. The ADHD-IV: Home Version was distributed to parents at the initial visit. The presence or absence of ADHD based on rating scales data was determined using procedures previously described (DuPaul et al., 1998). A 1.5 standard deviation cutoff was used with an age and gender matched normative database. Note that one subject was older than the validated age range for ADHD-IV and was therefore not evaluated by the rating scales.
2.4. EEG

EEG data collection and analysis were performed blinded to the results of the psychiatric evaluation and the rating scales. About 3 to 7 days after the psychiatric evaluation, EEG was recorded using a Digital Cortical Scan apparatus (Lexicor Medical Technology, Augusta, GA). Electrodes were placed using a 19-channel electrode cap in accordance with the International 10–20 system. Impedance for all electrodes was less than or equal to 10 kΩ. The cap ground electrode was at Fz. A linked-ears reference was used. Horizontal eye movement was monitored by electrooculography (EOG) with electrodes placed next to the right eye and the left eye. Vertical eye movement was monitored by EOG with electrodes placed above and below the right eye. Patients were seated in a fixed, straight-back chair and EEG was recorded during two conditions for 3 min each: (1) eyes closed and (2) eyes open with fixed attention on a point on the wall at eye level. Digitized EEG data was collected at a sample rate of 128 Hz with 32,000 amplification, a 60 Hz notch filter, and band pass down 3 dB at 0.5 and 36 Hz. Artifact removal was performed by visual assessment of an EEG technician trained through a 6-month certification program. The EEG technician determined acceptance or rejection of 2-s epochs based on artifact presence. At least 15 epochs (30 s) of data with minimal artifact were required from each of the readings in order to perform the EEG analysis. Spectral analysis was performed on artifact-free epochs using Fast Fourier Transform with a frequency resolution of 0.5 Hz. A two-variable EEG test was utilized for the prediction of ADHD vs. non-ADHD. Variables of frontal beta power and theta/beta ratio were compared against normative database values in order to produce the ADHD prediction at standard deviation cutoffs of 2.0 and 1.5, respectively.

Frontal beta power was examined in a manner previously suggested (Clarke et al., 2001d) utilizing a 2.0 standard deviation cutoff to identify a subgroup of ADHD against a normative sample. In the current study, normative comparisons were performed using a version of the Lifespan Normative Database (Applied Neuroscience, Reddington Shores, FL) that had been subjected to frequency response transformation and resampling in order to conform to the filtering characteristics and sampling format of the Digital Cortical Scan apparatus. The database required eyes closed, EEG data with the beta frequency band defined at 13.0–21.5 Hz. The database is matched by age groups with resolution ranging from 0.2 to 1.3 years.

The theta/beta ratio investigation followed suggestions of two previous studies (Monastre et al., 2001, 1999) utilizing a 1.5 standard deviation cutoff for ADHD against samples of normative controls that were grouped by age (6–11, 12–15, 16–20, and 21–30 years). While the previous studies utilized both task intervention and resting baseline EEG, there were no significant differences between task and baseline theta/beta ratio except for one slight difference for which ordering effect was not ruled out. Therefore, the current study examined these cutoffs applied to theta/beta ratio with eyes open, baseline EEG. Requirements of the cutoffs included EEG measurements at Cz with analyses of a ratio of relative theta and beta powers using a theta frequency band at 4.0–7.5 Hz and a beta frequency band at 13.0–20.5 Hz. Evaluation of frequency magnitude response between EEG equipment of the current and previous studies determined that the data fell within the range of the passband demonstrating that effects of filter differences were negligible and did not require frequency response transformation. In other words, the cutoffs were verified to be applicable to EEG data collected within the format of the Digital Cortical Scan apparatus.

2.5. Statistical analyses

We hypothesized that the EEG and rating scale predictions of ADHD would agree with the ADHD diagnostic results of the standard psychiatric evaluation. The agreement of each predictor variable (EEG test or rating scales) was tested relative to the grouping variable (psychiatric evaluation). This produced sensitivity and specificity results, as well as overall classification accuracy. Sensitivity and specificity are ratio equations requiring identification of true positives, true negatives, false positives, and false negatives. “True” and “false” were determined by the reference standard, in this case the psychiatric evaluation including the KSADS-PL and supplements, the CGI-S, and CGAS. A board-certified psychiatrist and fully trained staff compiled the information collected in the psychiatric evaluation to make the diagnosis of ADHD or non-ADHD. The EEG test and the rating scales were each used to make separate predictions of ADHD or non-ADHD. Testing these predictions against the psychiatric evaluation provided designations of true positives, true negatives, false positives, and false negatives. Sensitivity refers to the probability that ADHD will be detected when present. Specificity refers to the probability
that the absence of ADHD (non-ADHD) will be detected when the disorder is not present.

3. Results

3.1. Prevalence of ADHD and comorbidities in the clinical sample

Table 1 displays the presence or absence of various disorders in this sample of patients. Of these 26 subjects who were suspected of expressing ADHD symptoms, 16 were diagnosed with ADHD (62%). The remaining 10 subjects did not have ADHD but rather numerous other disorders or no diagnosis (38%). The rate of comorbidity in subjects positive for ADHD by the psychiatric evaluation was 31% with at least one other condition; specifically 13% with ODD, 7% with CD, and 7% anxiety/depression. Rates were similar with those ADHD positive on the EEG test: 25% (at least one), 13% (ODD), 7% (CD), and 7% (anxiety/depression). In comparison, those subjects positive for ADHD on the rating scales demonstrated comorbidity rates of 60% with at least one other disorder, 25% with ODD, 10% with CD, 25% with anxiety/depression, and 11% with a learning disorder (reading disorder).

There was no significant difference in terms of age between the ADHD and non-ADHD groups. There were also no between-group differences in impairment, functioning, and severity according to CGAS or CGI-S scores.

3.2. Accuracy of rating scales

The results of the behavior rating scales seem to reflect the opinions of the informants (parents) with most of the subjects testing positive on the scales for ADHD. Of the subjects confirmed to have ADHD by full standard psychiatric evaluation, 13 out of the 16 were identified by the rating scales as having ADHD (i.e. 13 true positives and 3 false negatives). Of the non-ADHD subjects determined by the standard psychiatric evaluation to have either no diagnosis or other disorders not ADHD, 7 of 9 were falsely identified by the rating scales as having ADHD (i.e. 7 false positives, 2 true negatives, and 1 excluded because outside of age range for the scales). In other words, rating...
scales were likely to classify attention, impulsivity, and/or hyperactivity symptoms as being due to ADHD, regardless of the actual underlying disorder, leading to a sensitivity of 81% and a specificity of 22% for the rating scales when applied to a clinical sample. The overall classification accuracy of the rating scales was 60%.

All of the ADHD inattentive subtype subjects were correctly identified by the rating scales, and the three false negatives were of the ADHD combined subtype. The presence of comorbidities did not appear to affect the outcome of the rating scales in detecting the presence of ADHD. The rating scales appeared to function well in identifying symptoms due to ADHD (sensitivity = 81%) but not so well in differentiating as to whether these symptoms were due to ADHD or due to other disorders (positive predictive power = 65%). The false positives of the rating scales were due to other disorders, including oppositional defiant disorder, anxiety disorder, conduct disorder, reading disorder, and dyslexia.

3.3. Accuracy of EEG

The age-matched EEG pattern for ADHD was observed to be present in 15 of 16 subjects diagnosed by the standard psychiatric evaluation as having ADHD (sensitivity = 94%). A standardized increase in the theta/beta ratio was the marker of the EEG pattern observed in 15 of 16 ADHD subjects, while no ADHD subjects with the excess frontal beta power marker were present in this sample. The EEG pattern was present in all of the subjects with inattentive or combined subtypes. There was only one subject of the hyperactive/impulsive subtype, and this person was the one false negative of the EEG test. The false negative subject was also comorbid for conduct disorder. The presence of a comorbid disorder did not affect the EEG outcomes in the identification of ADHD inattentive and combined subtypes.

The non-ADHD group contained subjects with numerous different disorders that included oppositional defiant disorder (4), anxiety disorder (4), dysthymic disorder (1), conduct disorder (1), adjustment disorder (1), reading disorder (2), and dyslexia (1). In addition, there was one patient with no diagnosis. Regardless of the presence of ADHD-like symptoms, the EEG pattern for ADHD did not occur in any of these non-ADHD patients (specificity = 100%). The overall classification accuracy of the EEG test was 96%.

3.4. EEG differences between groups

In order to further explore the nature of the effects behind the EEG test, analyses of covariance were performed for the effects of group (ADHD vs. non-ADHD) on relative power values and the theta/beta ratio at Cz controlling for age as a covariate with an interaction included when significant in the random effects (small sample) model. There was a significant age × group interaction for beta relative power ($F_{1,22} = 22.1$, $P<0.001$) and theta/beta ratio ($F_{1,22} = 4.6$, $P<0.05$). There was a group effect for theta relative power ($F_{1,23} = 9.2$, $P=0.006$) and for theta/beta ratio ($F_{1,22} = 12.6$, $P=0.002$). There was an age effect for beta relative power ($F_{1,22} = 94.7$, $P<0.001$) and theta/beta ratio ($F_{1,22} = 15.8$, $P=0.001$). Direct group comparisons were provided by one-way analysis of variance; ADHD showed a significant increase in theta relative power ($F_{1,24} = 11.5$, $P=0.002$) and theta/beta ratio ($F_{1,24} = 19.4$, $P<0.001$) and decrease in beta relative power ($F_{1,24} = 26.2$, $P<0.001$) vs. other disorders. There were no significant differences by group for relative power results in the other frequency bands (delta1 (1.0–1.5 Hz), delta2 (2.0–3.5 Hz), alpha (8.0–12.5 Hz), and beta2 (21.0–31.5 Hz).

4. Discussion

All subjects included in the study had presented to a clinic due to suspicion of ADHD-like symptoms, yet only 62% were diagnosed with ADHD, and the remaining 38% had other disorders or no diagnosis. Rating scales characterized this clinical sample with an overall accuracy of 60%, which is within the range reported by previous studies (55–79%). The EEG results identified ADHD vs. other disorders with an overall accuracy of 96%, also within the range reported by previous studies (77–96%). The study’s design directly followed recommendations of the AAP and the AAN/ACNS, using a blinded protocol and a clinical standard in the investigation of an applicable clinical sample. The accuracy of the results and the consistency with previous studies support that a large sample study should be pursued for further verification. These issues are discussed in detail in the sections that follow.

4.1. Rating scales

If the rating scales had been effective, then only the ADHD patients would have received positive results from the rating scales leaving the non-ADHD subjects with negative results despite the presence of their various symptoms. However the rating scales produced false positives for most of the non-ADHD subjects. The rating scales seemed to be vulnerable to the tendency to identify any inattentive, hyperactive, and impulsive symptoms as indicative of ADHD, regardless of the actual underlying disorder.
In previous studies using optimal cutoffs, the ADHD Rating Scale-IV produced a sensitivity of 84% and a specificity of 49% for ADHD vs. normal controls (DuPaul et al., 1998). While the sensitivity (81%) in the current study is close to the expected accuracy, the specificity (22%) was lower than predicted by the ADHD vs. control study. One explanation is that in the classification of ADHD vs. asymptomatic controls, lay informants (parents) were essentially differentiating between the presence or absence of symptoms. In the current study, the lay informants (parents) were differentiating between ADHD and other disorders with similar symptoms. The rating scales did not demonstrate the ability to guide the lay informant in differentiating between symptoms due to ADHD vs. similar symptoms due to other disorders. Therefore the rating scales resulted in many disorders being falsely labeled as ADHD leading to a lower specificity that is more applicable to the clinical setting.

4.2. EEG

In contrast, brain electrical activity in the baseline condition may provide a marker that is specific to ADHD within the clinical sample of subjects presenting with ADHD-like symptoms. Previous validation studies of the EEG marker (theta/beta ratio) for ADHD have demonstrated sensitivity in the range of 86–90% and specificity of 94–98% (Monastra et al., 2001, 1999). The current study observed 94% sensitivity and 100% specificity. Given that previous studies examined ADHD vs. asymptomatic controls and the current study examined a diverse clinical sample, a slight improvement in diagnostic accuracy was not necessarily expected. One explanation may be because the previous studies examined larger sample sizes of 482 and 129, which would have allowed for more precise resolution of validity results. When considering the current sample of 26 in terms of the previous validity results, only 2 false positives and/or false negatives are to be expected. Therefore it is within the realm of reasonable probability to have observed only 1 false positive due to the limited sampling size of the current study.

4.3. EEG and frontal beta power

As much as 15–20% of ADHD combined subtype subjects can be expected to demonstrate excess frontal beta power rather than an increase in theta/beta ratio (Chabot and Serfontein, 1996; Clarke et al., 1998, 2001b, c,d, 2002b). The current study included five subjects diagnosed as ADHD combined type, therefore the above prevalence implies that one subject with excess frontal beta power would be expected to be found within this sample size. The fact that none were observed suggests that a study with a larger sample size may be necessary to determine contributions of frontal beta power to the EEG assessment scheme. It also remains a possibility that the samples observed with excess frontal beta power in previous studies were not necessarily ADHD, but instead another disorder with similar symptoms. The excess frontal beta power subtype has been shown to have not only symptoms of inattention, hyperactivity, and impulsivity, but also proneness to tantrums and moodiness (Clarke et al., 2001d). Alternative diagnoses might include major depressive disorder, which not only may be responsible for such behavioral symptoms, but also has been observed to demonstrate excess frontal beta power (Matousek, 1991; Pollock and Schneider, 1990).

4.4. Prevalence of ADHD and comorbidities

The comorbid rates of subjects positive for ADHD on the EEG or the psychiatric evaluation 25–31% (at least one other condition), 13% (ODD), 7% (CD), and 7% (anxiety depression) fall short of previously reported values; 67% (at least one), 27% (ODD), 10% (CD), and 21% (anxiety depression) (Cantwell, 1996; Wolraich et al., 1998). One possible explanation is a sampling difference for this clinic relative to other ADHD clinics, limiting the generalizability. A solution would be to apply the current protocol to a multi-site study.

The comorbid rates of subjects positive for ADHD on the ratings scales are quite similar to that of previous studies with the current values at 60% (at least one), 25% (ODD), 10% (CD), and 25% (anxiety depression). In the current study, the main reason the comorbid rates for the rating scales are at this higher level is because of the rate of false positives in the rating scales. Subjects diagnosed per the psychiatric evaluation as ‘non-ADHD with other disorders’ were often diagnosed as ‘ADHD with comorbidities’ using the rating scales. Therefore the appearance of ADHD false positives has a strong effect on observed comorbid rates.

When comparing rating scales vs. both EEG and the psychiatric evaluation, the key difference is the presence or absence of informant bias. EEG is a quantitative, physiological measure compared against previously set cutoffs, which offers no informant bias. At most, there could be a processing bias in the removal of artifact from the EEG data, however a blinded protocol such as provided by the current study minimizes if not eliminates such an effect.

The rating scales have no overt control for informant bias which has been reported to have a significant effect.
on the outcome (Collett et al., 2003; Doyle et al., 1997; Eiraldi et al., 2000). The commonly observed disagreement in rating scale outcome between informants such as parents vs. teachers underscores this effect (AAP, 2000). In contrast, the psychiatric evaluation is designed to control bias of the informants (child/adolescent and parent). For instance, with the K-SADS-PL component there are not only two layers of screens (K-SADS-PL and various supplements for specific disorders), but there is an unstructured interview to crosscheck the structured interview. Further the structured questions are designed to be rephrased and revisited to check the consistency of the informant. The inclusion of both the child/adolescent and parent as informants in the K-SADS-PL should be ideal for identifying externalizing disorders (for which parents are considered to be the most effective informants) and internalizing disorders (for which the children/adolescents are thought to be the best informants). And finally, most potential childhood and adolescent disorders are addressed which allows consideration of similar symptoms for numerous disorders, as opposed to the rating scales which addressed simply ADHD.

4.5. Responding to previous criticisms

Previous reviews have addressed specific criticisms against EEG in the identification of ADHD. Claims of inconsistent results between studies, insufficient sensitivity and specificity, and inadequate experimental designs are the key points of concern (AAP, 2000; Levy and Ward, 1995; Nuwer, 1997). In direct response to these criticisms, the current study utilized the recommended experimental design of a blinded protocol with a clinical standard to examine a diverse clinical sample of patients in the application of EEG and rating scales in the identification of ADHD. With an eye towards testing consistency between studies, cutoffs and variables were examined as replicated from previous studies for EEG (Clarke et al., 2001d; Monastra et al., 2001, 1999) and rating scales (DuPaul et al., 1998). The rating scale results were fairly consistent with previous results. And, the accuracy results were favorable for EEG (94% sensitivity and 100% specificity) as well as consistent with the previous results (86–90% sensitivity and 94–98% specificity). Further, these accuracy results were also consistent with those of a range of related methods of the field for both EEG (Chabot et al., 1996; Chabot and Serfontein, 1996; Clarke et al., 2002b; Kovatchev et al., 2001; Mann et al., 1992; Monastra et al., 2001, 1999) and rating scales (Bussing et al., 1998; Doyle et al., 1997; DuPaul et al., 1998; Eiraldi et al., 2000; Luk and Leung, 1989; McCann et al., 2000; Rucklidge and Tannock, 2002; Sprafkin et al., 2002; Vaughn et al., 1997). The underlying EEG changes of increased theta relative power and theta/beta ratio and decreased beta relative power were consistent with the results of studies of EEG and DSM-IV ADHD (Bresnahan et al., 1999; Clarke et al., 1998, 2001a,b,c,d, 2002a,b; El-Sayed et al., 2002; Lazzaro et al., 1999, 1998; Monastra et al., 2001, 1999), and these changes are consistent with the model of cortical hypoarousal as described in a comprehensive review of EEG and ADHD (Barry et al., 2003). Therefore, the prior criticisms of EEG are due for reconsideration in light of the consistency and accuracy of the current results.

4.6. Limitations

There are several limitations to this study that are important to recognize. First, a small sample was examined in this preliminary study. With 16 ADHD subjects, sensitivity could be distinguished in increments of 6%. With 10 non-ADHD subjects, specificity could be determined in increments of 10%. To characterize the validity with more precise resolution, a study with a larger sample size will be required. A second is the use of a single psychiatric clinical site for the current study which limits the generalizability of the results. The generalizability of the results would be strengthened by including other clinical sites such as primary care and pediatrics settings. Also a more representative sample would be provided with the inclusion of further geographic locations. The difference in comorbidity rates between the current study and previous reports underscores the need for a multi-site study in order to provide a more representative patient sample.

In the current study, females comprised 12% of the sample. Although low, this ratio falls within the range expected for a clinical sample (10–25%) (Arnold, 1996; Gaub and Carlson, 1997). However given this ratio, a larger sample would be required to provide adequate representation of females to allow specific analyses of gender differences. The same is true of groupings by age and race. A fourth limitation relates to the blinding of the protocols. Although EEG and the standard psychiatric evaluation were performed in a blinded protocol, this process was not applied to the rating scales. The clinicians determining the diagnoses by psychiatric evaluation were not blinded to the rating scales results, and although the rating scales outcomes did not appear to be improved by this practice, a blinded protocol is required to ensure elimination of bias.

The core of the psychiatric evaluation was the K-SADS-PL which relied on the parent and child/adolescent as informants. Although teacher rating scales are
acquired as part of the regular practice of the participating clinic, they were not standardized into the current diagnostic protocol. Insufficient information from teachers can result in false positives (AAP, 2000; de Nijs et al., 2004) which must be considered a potential limitation for the study.

The present study used only one type of rating scale, the ADHD-IV. An argument could be made that other scales might have performed better, such as the Swanson, Nolan, and Pelham Rating Scale (SNAP) or the Conners’ Rating Scales–Revised (CRS-R) which have been well received in previous reviews (AAP, 2000; Collett et al., 2003; Green et al., 1999). The ADHD-IV rating scale was selected, however, because it was the only scale for which criterion validity has been previously examined using accepted experimental designs and statistical techniques (Snyder et al., 2004). When examining the validation protocols of seemingly better performing scales, it should be noted that the CRS-R and SNAP were validated using each scale as a reference for itself, which is a circular proof (Atkins et al., 1985; Conners, 1997). And other validations of the CRS-R involved the misuse of discriminant analysis by using the same sample for development as for validation (Conners, 1997; Conners et al., 1998a,b, 1997; Snyder et al., 2004). Because of the definitive effect that these previous studies have had on the practices and guidelines of the field (AAP, 2000; Green et al., 1999), it may be enlightening to include CRS-R or SNAP in a future validation study that utilizes a more rigorous experimental design.

4.7. Implications

Although parents and school officials may anticipate an ADHD diagnosis when a child presents with inattentive, impulsive, and/or hyperactive symptoms, there is a significant chance that the child is suffering from another disorder (Cantwell, 1996). In the current study, 38% (10/26) of the patients who presented to a clinic because of suspected ADHD-like symptoms did not actually have ADHD. These results support that parent and teacher bias must be handled effectively when ruling out other psychiatric disorders in the diagnosis of ADHD.

The false positive results of the rating scales when used alone would have led to the incorrect confirmation of the original suspicions of many of the parents and school officials. Without further evaluation of these patients, over-diagnosis of ADHD would have occurred. Professional guidelines and rating scale manuals do strongly advise against the use of rating scales as a stand-alone diagnostic (AAP, 2000; Bussing et al., 1998; Collett et al., 2003). Unfortunately when used in the common scenario of the limited 15-min office visit, the false positives of rating scales may have an undue influence on the final diagnosis. The implication is that rating scales, when not used properly, may contribute to the over-diagnosis of ADHD.

An ADHD diagnostic protocol might be designed to include the observation of behavioral symptoms, the recognition of impairment in two settings, the confirmation of symptoms occurring before the age of 7, and the presence of the EEG marker. Therefore it would be worthwhile to test the validity of this protocol in a multi-site study examining a large and diverse clinical sample. If validated, the EEG test may prove to be of benefit to the dilemma of the limited 15-min office visit for ADHD diagnosis. An EEG laboratory test provided separately by trained technicians would not require further office time, and would provide a novel piece of evidence in position to complement the observation of behavioral symptoms and impairments, leading potentially to an improvement in the accuracy of differential diagnosis in the clinical setting.

References


attention deficit or specific developmental learning disorders. Clinical Electroencephalography 27, 26–34.
Doyle, A., Ostrander, R., Skare, S., Crosby, R.D., August, G.J., 1997. Convergent and criterion-related validity of the Behavior Assess-