Should the EEG–Based Theta to Beta Ratio Be Used to Diagnose ADHD?

Sandra K. Loo, Ph.D., and Martijn Arns, Ph.D.

Electroencephalography (EEG) has a long history in child psychology and psychiatry research; it was the first methodology used for examining human cortical brain activity among children. The use of modern EEG technology was first reported by Hans Berger in the 1920s (Berger, 1929), who provided extensive description of methodology and initial recordings made on his son. Subsequently, Jasper, Solomon, & Bradley (1938) reported increased slow-wave activity in fronto-central regions, a putative indicator of abnormal brain function in a group of “behavior problem children”—described as hyperactive, impulsive, and highly variable. EEG has been widely used in child psychopathology studies, likely due to its many advantages for functional brain imaging—including flexibility, portability, non-invasive nature, wide subject acceptance, excellent temporal resolution and relatively low cost. Yet, 75 years later clinical applications of EEG in psychiatry continue to be controversial, with the primary question being whether the knowledge gained from EEG has any clinical diagnostic value (Loo & Makeig, 2012). This review is focused on a particular EEG measure, the theta to beta ratio, which is being used to aid assessment of ADHD diagnosis since its Food and Drug Administration (FDA) de novo device approval in 2013. First, a summary of the scientific literature on the association between the theta to beta ratio (TBR) and ADHD is presented. Next, a brief review of the scientific study (Snyder, Rugino, Hornig, & Stein, 2015) that was used to gain FDA device approval, including critical flaws, is given. Finally, future directions for biomarkers (EEG and others) are discussed.

WHAT IS THE EEG THETA TO BETA RATIO?

One of the primary ways to characterize EEG data is through an estimate of the power expressed within a given frequency (reported in hertz [Hz], the number of waveform cycles per second). While definitions may vary, EEG power variations are typically dominated by distinct changes in power in a few frequency bands. The standard terminology for these frequency bands is: delta (<4 Hz), theta (4–8 Hz), alpha (8–13 Hz), and beta (13–25 Hz). In a resting state, (slower frequency) theta band activity can reflect drowsiness or “cortical slowing.” Alpha band activity is typically observed during eyes closed rest, particularly in posterior regions, and is negatively associated with central nervous system arousal. Beta band activity, by contrast, generally accompanies mental activity and concentration. The theta to beta ratio (TBR) is a ratio of power in the slow-wave theta frequency band relative to power in the faster, beta frequency band, often measured at a single electrode in the middle of the head (Cz) during resting conditions. The TBR has been proposed to be a better way to capture the relative levels of the offsetting brain activation patterns evident in the EEG, rather than focusing on one particular frequency band, such as slow-wave theta band power (Monastra, Lubar, & Linden, 2001).

FUNCTIONAL SIGNIFICANCE OF THE TBR

In the 1970s, Satterfield conducted a series of EEG studies of children with ADHD and found EEG abnormalities, including excess slow-wave activity and increased epileptiform spike and wave activity (Satterfield, Cantwell, & Satterfield, 1974). These findings were thought to suggest under-arousal and maturational delay as underlying pathophysiologies in ADHD. Further, children with ADHD who had greater excess slow-wave activity were more likely to have a positive response to stimulant medication (Satterfield, Cantwell, Saul, Lesser, & Podosin, 1973), a finding that fit well with cortical under-arousal theories. Because increases in slow wave activity contribute heavily to the relationship between the TBR and ADHD, this ratio was also believed to represent hypoarousal (Lubar, 1991; Mann, Lubar, Zimmerman, Miller, & Muensch, 1992). Two recent attempts, however, to explicitly test the association between TBR and arousal—Barry and colleagues and Clarke and colleagues—have reported no significant relationship between TBR and a measure of arousal, skin conductance level (SCL) (Barry, Clarke, Johnstone, McCarthy, & Selikowitz, 2009; Clarke et al., 2013). Instead, they replicated their previously reported associations between SCL and power in the “alpha” (8–14 Hz) frequency range (Barry et al., 2004). Thus, the TBR does not appear to represent hypoarousal, at least as measured by SCL. In other studies, the TBR has been negatively correlated with mean reaction time in adults both with and without ADHD (van Dongen-Boomsma et al., 2010), but it was not significantly correlated with ADHD symptoms or cognitive performance on a sustained attention task in a child sample (Ogrim, Kropotov, & Hestad, 2012). Based on these findings, the TBR has been proposed to reflect task-related cortical activation, but more research is needed to identify the range of conditions under which these differences appear and to understand the functional significance of these effects in terms of the underlying cortical processes that produce them.

THE TBR AND ADHD DIAGNOSIS:
BRIDGING THE GAP BETWEEN PRE- AND POST-2009 FINDINGS

Early studies reported a significantly higher TBR among children and adolescents with ADHD compared to non-ADHD controls (Barry, Clarke, Johnstone, & Brown, 2009; Bresnahan, Anderson, & Barry, 1999; Clarke, Barry, McCarthy, & Selikowitz, 2001; Clarke, Barry, McCarthy, & Brown, 2009; Clarke, Clarke, & Hestad, 2012). Based on these findings, the TBR has been proposed to reflect task-related cortical activation, but more research is needed to identify the range of conditions under which these differences appear and to understand the functional significance of these effects in terms of the underlying cortical processes that produce them.
thy, Selikowitz, & Brown, 2002; Monasta et al., 2001; Monasta et al., 1999). In their meta-analysis of nine studies reporting TBR data \((n = 1,498)\), Snyder and Hall (2006) reported an effect size (ES) of 3.08, suggesting that high TBR is commonly observed in ADHD relative to controls. This estimate is largely consistent with their pilot study (Quintana, Snyder, Purnell, Aponte, & Sita, 2007) as well as a large, multi-site, prospective study of the discriminant validity of the theta/beta ratio in ADHD from this same group. In 2008, Snyder and colleagues (2008) reported that the overall accuracy rate in identifying ADHD was 89%, suggesting that an abnormally high TBR identifies almost all of the children who are subsequently given a diagnosis of ADHD. Further, the TBR purportedly demonstrated similarly high rates of diagnostic accuracy across demographic groups that varied according to age, gender, and ethnic background (range 87–95%) as well as in the presence or absence of comorbid conditions (range 87–96%) (Snyder et al., 2008). Collectively, early studies of the TBR suggested it may have strong clinical utility for ADHD diagnosis.

Starting in 2010, however, several independent research groups published results that fail to support the early results on the TBR and ADHD diagnosis (Buyck & Wiersema, 2014; Koehler et al., 2009; Lansbergen, Arns, van Dongen-Boomsma, Spronk, & Buitelaar, 2011; Liechti et al., 2013; Loo et al., 2013; Nazari, Walois, Ardalan, & Berquin, 2011; Ogrim et al., 2012; Sohn et al., 2010). For example, in the largest study of the TBR to date, no significant differences were found between 562 children, adolescents, and adults with ADHD compared to 309 non-ADHD controls, with modest heterogeneity attributed to ADHD subtype and psychiatric comorbidity (Loo et al., 2013). Further, Liechti and colleagues (2013) found 81% accuracy in predicting age based on TBR, but only 53% accuracy in predicting ADHD diagnosis. Similarly, Buyck and Wiersma (2014) found a 97% successful identification of age group (e.g., child, adult) but only 55% discrimination rate for the TBR and ADHD diagnosis, with ADHD subtype playing a significant mediating role. These null results were reflected in a recent meta-analysis by Arns, Conners, & Kraemer (2013), who reported a diminishing TBR effect size that was significantly associated with year of study publication \((r = -0.93, \ p = 0.007)\). Notably the TBR for the ADHD group has remained fairly stable across years; however, the control group TBR has steadily risen over the years between 1999 and 2013 (Arns et al., 2013). The reason for this dramatic increase in the TBR among controls is unknown and requires further research. Collectively, the recent literature suggests that the TBR is not reliable in discriminating between individuals with and without ADHD.

Several potential factors account for this marked change in empirical support for the TBR.

1. A larger number of independent research groups without a conflict of interest (with commercial products, U.S. patents, or delivery of clinical services based on the TBR) have published recent data on the association between the TBR and ADHD.

2. Some of the earlier studies used normative databases of control group EEGs rather than collecting contemporaneous EEG data from non-ADHD individuals. This approach has the potential for confounding results merely because EEG data were collected on ADHD and control samples who were ascertained under different methodological considerations and obtained in different recording environments. This may have produced spurious large effects in the results.

3. Because there appears to be a subgroup of youth with ADHD with elevated TBR, ascertainment bias may have resulted in disproportionately large numbers of this subgroup within the ADHD samples of early studies.

4. The increasing TBR ratio in the control group over time may be a real phenomenon that is driven by environmental factors such as more prevalent use of electronic devices and subsequent neurobiological effects or sleep deprivation for other reasons. This would serve to decrease the effect size of the TBR over time (Arns et al., 2013).

Regardless of the reason for the dramatic shift in empirical results on the association of the TBR and ADHD, there also appears to be a gap in knowledge where the earlier results are well known and widely disseminated whereas the more recent negative results are not as strongly emphasized. This may simply be a time effect where it takes a certain amount of time for scientific results to make their way into the popular consciousness. In addition, the ADHD literature is voluminous, and it is certainly understandable that professionals, the lay public, and the media have not been able to stay abreast of the findings. This makes it all the more crucial to fully inform as many professionals and consumers as possible to avoid the adoption of EEG measures and devices that will likely fail to accurately identify those with ADHD.

**UTILITY OF THE NEBA DEVICE FOR ADHD DIAGNOSIS**

In July 2013, the FDA approved the marketing of the “first brain wave test to help assess children and teens for ADHD” commercialized by Neuropsychiatric EEG-Based Assessment Aid (NEBA) Health (http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm360811.htm). Recently, the details behind the clinical study were published by the Vice President of Research and Development of NEBA Health, S. Snyder and colleagues (Snyder et al., 2015). Review of this study allows (partial) evaluation behind the claim that dominated the news and social media when it was announced.

The NEBA device relies on the TBR and is presented as an “assessment aid” rather than a diagnostic test, with the goal of helping to “improve certainty with [DSM-5] criterion E’ (i.e., whether

---

1. DSM-5, ADHD diagnostic criteria, criterion E: The symptoms do not occur exclusively during the course of schizophrenia or another psychotic disorder and are not better explained by another mental disorder (e.g., mood disorder, anxiety disorder, dissociative disorder, personality disorder, substance intoxication, or withdrawal).
symptoms are better explained by another condition...” (Snyder et al., 2015, p. 1). To this aim, a method was used where clinicians at each site made a differential diagnosis for every case. The reference or gold standard used was a consensus diagnosis by an independent off-site multidisciplinary team deemed to evaluate criterion E in a complex clinical population. Comparisons were then made between the clinician’s decision (1 = ADHD present or 2 = uncertain), the NEBA test classification (based on a low, moderate, or high TBR), and the multidisciplinary team designation (1 = ADHD present; 2 = need further testing with ADHD as focus; 3 = need further testing with other conditions as focus; and 4 = ADHD absent). While this large amount of data makes integration and categorization more complex, the decision points and methods for grouping categories constitute a fundamental weakness of the study and subsequently undermine the strength of the study conclusions, as elaborated below.

CONCLUSIONS THAT ARE SUPPORTED BY THE SNYDER AND COLLEAGUES (2015) STUDY

The data presented in the Snyder and colleagues’ (2015) study demonstrate that the TBR does seem to correlate with complicating medical and psychiatric conditions that can make ADHD diagnosis more difficult (i.e., a low to moderate TBR in such conditions). Because the analyses grouped many different conditions together and did not report TBR results for the absence of ADHD in the presence of these complicating conditions, it remains unclear if there are specific conditions that drive this relationship more strongly than others. Of note, these data are in contrast with the Snyder and colleagues’ (2008) study where they state, “EEG accuracy was consistent for ADHD in the presence or absence of comorbidities” (p. 352). If further testing supports a relationship between the NEBA test and specific complicating medical and psychiatric conditions, this may perhaps be a future use of the device.

METHODOLOGICAL PROBLEMS PRECLUDING STRONG CONCLUSIONS REGARDING THE NEBA DEVICE

Problem 1: Unclear Definition of ADHD Present or Absent. Snyder and colleagues (2015) state on page 4, “ADHD was listed as ‘positive’ for the analysis if the clinician’s primary diagnosis was ADHD (combined, hyperactive/impulsive, or inattentive subtypes) with definite certainty. ADHD was considered as ‘negative’ for the analysis if the clinician’s primary diagnosis was for another condition, and the clinician listed ADHD as absent or secondary.” According to this, “ADHD negative” could mean that there is no ADHD, but it could also mean that ADHD is present but secondary to another condition that is primary. Therefore, if ADHD was secondary to a mood, anxiety, or disruptive behavior disorder, the case was considered to be ADHD negative. This definition-al uncertainty makes it impossible to know whether any given participant has ADHD or not, regardless of the primary condition. Additionally, most conditions listed in Snyder and colleague’s Table 4, such as medical mimics, anger, and medication issues do not preclude an ADHD diagnosis, which would likely be a focus of treatment even if it is not the primary condition due to the impairing nature of ADHD symptoms. This definitional uncertainty makes the rest of the study impossible to interpret since ADHD positive and ADHD negative are used as final outcomes, and yet ADHD negative may in fact still indicate the presence of an ADHD diagnosis. How can the usefulness of the NEBA device for detecting ADHD be evaluated without knowing with certainty whether the person has ADHD or not?

Problem 2: DSM Criterion E Certainty Is Not the Same as ADHD Present or Absent. Even if ADHD negative was clearly defined, an additional problem is the use of criterion E certainty (DSM-5; American Psychiatric Association, 2013) to mean the same thing as no diagnosis of ADHD. This problem is compounded since the authors then calculate sensitivity and specificity estimates using less likely to meet criterion E as its own category or outcome. For example, Snyder and colleagues (2015) state on page 3, “The biomarker outcome is used to separate patients with clinician designations of ‘uncertain’ and ‘positive’ (ADHD) into EEG–based subgroups with recommendations regarding: (1) ADHD confirmation, or (2) criterion E certainty (with a suggestion for resolution by further clinical testing).” This suggests that the final NEBA outcome can only be used to confirm a clinician’s ADHD diagnosis derived through standard clinical procedures and cannot be used to definitively rule out ADHD, only support further testing focusing on ADHD or other conditions. If NEBA is unnecessary for making an ADHD diagnosis and cannot rule out an ADHD diagnosis, the added clinical value of the NEBA device is unclear.

Problem 3: Discrepancy Between the Clinician and Multidisciplinary Team Is Interpreted as ADHD Over-Diagnosis. Again, using criterion E certainty to mean that there is no ADHD diagnosis, Snyder and colleagues (2015) state on page 7, “Of the 209 patients meeting ADHD criteria according to an individual clinician’s judgment, 93 were separately found by the multidisciplinary team to be less likely to meet criterion E, implying possible over-diagnosis by individual clinicians in 34% of the total clinical sample (93/275).” The reference to “over-diagnosis” of ADHD thus seems to apply more to the initial ADHD screening in this study, which is being compared to criterion E and a “No ADHD” designation that may in fact still indicate that ADHD is present (but secondary). These latter two standards are contrary to good clinical practice, incorporate nonstandard diagnostic criteria for ADHD (such as a good medication response without side effects), and are clearly at odds with clinical and neurobiological evidence. In addition, the multidisciplinary team used an additional decision category (4 categories: 1. ADHD; 2. need additional testing with ADHD as focus; 3. need additional testing with other conditions as focus; and 4. not ADHD) that was not available to the initial clinicians (3 categories: 5 = ADHD positive; 6 = uncertain; and 7 = ADHD negative). For sensitivity/specificity analyses, clinician ADHD uncertain and ADHD negative...
(6 & 7 above) were grouped together while the multidisciplinary team ratings of “ADHD” and “need additional testing with ADHD as focus” (1 & 2 above) were grouped together to mean “ADHD positive”; “need additional testing with other conditions as focus” and “not ADHD” (3 & 4 above) were grouped together to constitute “ADHD negative.” Data for each individual category were not presented, making it impossible to evaluate whether clinician and multidisciplinary team designations of ADHD (1 & 5) were consistent with each other. This arbitrary grouping and use of the “need further testing” categories is used to imply that clinicians are over-diagnosing ADHD, when in fact they were not instructed to differentiate between further testing focused on ADHD or other conditions; this makes it impossible to know whether the NEBA test is really necessary at all and does not justify the claim of ADHD over-diagnosis.

Problem 4: Circular ADHD Definition. Many of the exclusion conditions (Snyder et al., 2015, Table 4) do not match the corresponding DSM criterion E, suggesting that the multidisciplinary team aimed at a partly circular ADHD definition. For example, the inclusion of side effects or poor response to ADHD medication results in a circular definition of ADHD (e.g., using a good response to ADHD medication without side effects as an indicator that an individual has ADHD). This is not in line with clinical guidelines and with the established heterogeneity of ADHD. Since EEG information was not collected systematically on treatment response, it can also not support the potentially more promising use of biomarkers as measures of prognosis or prediction of treatment response rather than diagnosis (Arns & Gordon, 2014).

Problem 5: Difficulties Reconciling Previous Literature on the TBR in ADHD with Current Study. It is notable that the Snyder and colleagues’ (2015) study was unable to replicate their earlier 2008 study in using the TBR as a sensitive and specific diagnostic marker of ADHD. In discussing their results, Snyder and colleagues explain the decreasing effect size (ES) for TBR across the years to be the result of over-diagnosis of ADHD or “rapid increase in ADHD prevalence, possibly related to a less stringent application of criterion E” (Snyder et al., 2015, p. 13). However, in their analysis to justify this (by correlating the ES of the TBR to the ADHD prevalence rate), there is no mention of the fact that this decrease in ES is caused by an increased TBR for the control groups, and not by a decreasing TBR for the ADHD groups. Their own data presented in their Table 5 support recent studies and meta-analysis suggesting that the TBR should not be used for diagnostic purposes in ADHD, even as a diagnostic aid. In this light, the consistently low TBR of participants who initially screened positive for ADHD but met some of these various nonstandard exclusion criteria is also an unexpected finding. As such, it requires independent replication, given the lack of replicated systematic TBR differences compared to normal controls, and the lack of a simple plausible mechanism underlying TBR.

Indeed, the authors’ stated goal for the NEBA and the current study is to demonstrate that the “integration of the biomarker may increase specificity in ADHD diagnosis.” However, to do this, the data must be based on actual, verified cases with no ADHD diagnosis, not just 1) ADHD as a secondary diagnosis, 2) the increased likelihood of meeting criterion E, or 3) a high correlation with other conditions that mimic ADHD. Until this has been done, the clinical utility of the EEG biomarker in increasing specificity in ADHD diagnosis has not been proven.

FUTURE DIRECTIONS FOR ADHD BIOMARKERS

Currently, no single diagnostic test for ADHD exists. Because inattention is pathognomic to nearly all childhood psychiatric disorders, it is often difficult to make differential diagnoses between ADHD and other disorders that can have a similar presentation, including autism spectrum disorders, mood and anxiety disorders, and learning disabilities. Thus, a biologically based diagnostic test or biological marker (i.e., biomarker) that is sensitive and specific to ADHD would be of great assistance. Unfortunately, as the past 75 years will attest, finding a simple diagnostic measure for ADHD (i.e., behavioral, cognitive, etiologic, neurophysiologic, neurobiologic, or genetic) has not been possible. Given the clinical heterogeneity present in nearly all psychiatric disorders, it seems unlikely that a single biomarker measure of any type can capture all of the variance within a diagnostic category. Instead, finding biomarkers for neural circuits associated with functional domains seems like a potentially more fruitful approach. Organizations such as the U.S. National Institute of Mental Health (NIMH) have proposed a move away from characterizing psychopathology based on psychiatric diagnoses but rather according to translational dimensions of functioning, such as cognitive processes, positive and negative valence, arousal, and self-regulation (Research Domain Criteria; RDOC). Each of these dimensions can be characterized by a multi-level pathway (i.e., genes, neural circuitry, brain dynamics, and behavior) that can inform searches for both underlying neural mechanisms and potential treatments.

In addition, potential biomarkers will likely need to encompass multiple measures in a multivariate fashion to capture various aspects of clinical heterogeneity. For example, we now understand that the genetic underpinnings of psychiatric disorders include the effects of hundreds, possibly thousands, of genes (Schizophrenia Working Group of the Psychiatric Genomics, 2014). Further, there appear to be genetic loci that are shared between psychiatric disorders (Cross-Disorder Group of the Psychiatric Genomics, 2013), such that similarities in brain function and cognitive processing across disorders are more likely. If that is the case, the likelihood of finding single genetic, brain, or biological measures that uniquely identify a particular psychiatric disorder is significantly smaller. More likely, there will be multivariate patterns of genes and/or brain measures that are associated and organized around more fundamental biological factors.

Another reason multivariate biomarkers may be necessary is that recent studies have reported that heterogeneity is present among individuals with psychiatric disorders such as ADHD,
but also within typically developing populations. Within several domains, including neuropsychological functions (Fair, Bathula, Nikolas, & Nigg, 2012), temperament (Karalunas et al., 2014), and brain function (Clarke et al., 2011; Costa Dias et al., 2015), similar subgroups have been detected in both populations. This again suggests that there will likely be no single measure that will differentiate between individuals with and without psychiatric disorder but instead a multivariate pattern of findings that indicate increased risk of disorder (Lenartowicz & Loo, 2014). An example of this can be found in the recent international ADHD-200 competition where the goal was to distinguish between children with and without ADHD using any combination of structural and functional magnetic resonance image measures recorded on over 700 children (491 ADHD, 285 typically developing) (ADHD Consortium, 2012). Results ranged in accuracy from 37 to 61%, with the highest prediction accuracy correctly classifying 94% of typically developing children but only 21% of children with ADHD. These results suggest that measures at different levels of phenotypes (genes + brain + behavior) may be necessary to differentiate those at greatest risk for psychiatric disorder. Collectively, this broader focus on etiology and symptomatic heterogeneity within psychiatric disorders and new focus on examination of domains of function across disorders suggests that efforts to develop more robust, highly cost-effective EEG-based biomarkers for psychiatric disorders will need to increase in complexity and span multiple levels of measurement.

CONCLUSION

Presented here is a summary of the brief, recent history of the NEBA device, which is based on an outdated literature that previously supported a strong relationship between the theta to beta ratio (TBR) and ADHD diagnosis. Although the TBR in older studies might have well dissociated ADHD patients from controls, all published studies after 2010—including the Snyder and colleagues’ (2015) study—have been unable to replicate this finding. This lack of differentiation in the post-2010 studies is related to an increasing TBR for control groups across the years, whereas the TBR for ADHD groups has remained stable across years.

This review also clarifies what the Snyder and colleagues’ (2015) paper does and does not demonstrate with their data. Methodological problems that preclude a strong conclusion regarding the clinical utility of the NEBA system as a diagnostic aid have been delineated. In conclusion, the proposed use of the TBR as a diagnostic aid is unclear in its definitions, exploits ambiguous decision categories, is partly circular, and does not integrate with good clinical practice of ADHD diagnosis. The study does not establish TBR as a biomarker that is diagnostically or prognostically useful beyond the standard clinical ADHD evaluation and provides neither independent validation nor a plausible biological mechanism for low TBR in the subgroup recommended for further evaluation. The NEBA test should not be used, even as a diagnostic aid, for ADHD diagnosis.

Sandra K. Loo, Ph.D., is Associate Professor-in-Residence in Psychiatry and Bi behavioral Sciences at the Brain Research Institute of the David Geffen School of Medicine, University of California, Los Angeles. Martijn Arns, Ph.D., is with the Research Institute Brainclinics, Nijmegen, Utrecht University, Department of Experimental Psychology, and neuroCare Group, Nijmegen, all in The Netherlands. Dr. Loo can be contacted by e-mail at sloo@mednet.ucla.edu

REFERENCES


Lenartowicz, A., & Loo, S. K. (2014). Use of EEG to diagnose ADHD. *Current Psychiatry Reports*, 16(11), 498. doi:10.1007/s11920-014-0498-0


