Psychiatry Research 300 (2021) 113879

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

# Biomarker support for ADHD diagnosis based on Event Related Potentials and scores from an attention test

L.A. Häger [a,b,*], J. Åsberg Johnels [b,c], J.D. Kropotov [d,e], B. Weidle [f,g], S. Hollup [h], P. G. Zehentbauer [i], C. Gillberg [b], E. Billstedt [b], G. Ogrim [a,b,h,*]

[a] Neuropsychiatric Team, Åsebråten Clinic, Østfold Hospital Trust, Fredrikstad, Norway
[b] Gillberg Neuropsychiatry Centre, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden
[c] Speech and Language Pathology Unit, Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden
[d] N.P. Bechtereva Institute of the Human Brain, Russian Academy of Science, St. Petersburg, Russian Federation
[e] Department of Neuropsychology, Andrzej Frycz Modrzewski Krakow University, Poland
[f] St. Olavs University Hospital, Trondheim, Norway
[g] Regional Centre for Child and Youth Mental Health and Child Welfare, Central Norway
[h] Institute of Psychology, Norwegian Institute of Science and Technology (NTNU), Trondheim, Norway
[i] Nordfjord BUP, Helse Førde, Norway

ARTICLE INFO

*Keywords:*
ADHD
Event related potentials
Biomarkers
Children
Win EEG
Electrophysiology
Diagnostic index

ABSTRACT

ADHD is a heterogeneous neurodevelopmental disorder associated with dysfunctions in several brain systems. Objective markers of brain dysfunction for clinical assessment are lacking. Many studies applying electroencephalography (EEG) and neuropsychological tests find significant differences between ADHD and controls, but the effect sizes (ES) are often too small for diagnostic purposes. This study aimed to compute a diagnostic index for ADHD by combining behavioral test scores from a cued visual go/no-go task and Event Related Potentials (ERPs).

Sixty-one children (age 9–12 years) diagnosed with ADHD and 69 age- and gender-matched typically developing children (TDC) underwent EEG-recording while tested on a go/no-go task. Based on comparisons of ERP group-means and task-performance, variables that differed significantly between the groups with at least moderate ES were converted to a five points percentile scale and multiplied by the ES of the variable. The sum-scores of the variables constituted the diagnostic index.

The index discriminated significantly between patients and TDC with a large ES. This index was applied to an independent sample (20 ADHD, 21 TDC), distinguishing the groups with an even larger ES.

The diagnostic index described has the potential to support assessment. Further research establishing diagnostic indexes for differential diagnoses is needed.

## 1. Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental disorder with a global average prevalence of around 5% (Krieger and Amador-Campos, 2017). The main symptoms are problems related to attention, hyperactivity and impulsivity (American Psychiatric, 2013). It is the rule rather than the exception that individuals with ADHD display comorbid conditions (Kadesjo and Gillberg, 2001), and the expression of the disorder varies both in individuals and in different developmental stages. This extended heterogeneity may constitute a barrier to the identification of the condition. ADHD is linked to life-long challenges in multiple areas of functioning such as in school, work, personal relationships, economy and health (Sayal et al., 2018). A recent study showed that individuals formally diagnosed with ADHD seem to experience higher work-related productivity, better quality of life and self-esteem compared with individuals without a diagnosis but who show equally high symptom load of ADHD (Pawaskar et al., 2020). Reliable identification of ADHD in childhood is therefore important to ensure proper treatment and understanding on an individual level, but also on a community level.

Even though many group studies have found deviant brain functions in ADHD (e.g., Friedman and Rapoport, 2015), no "brain-based"

* Corresponding authors. Sykehuset Østfold HF, Pb 300 1714 Grålum, Norway.
*E-mail addresses:* linda.hager@so-hf.no (L.A. Häger), geir.ogrim@so-hf.no (G. Ogrim).

https://doi.org/10.1016/j.psychres.2021.113879
Received 4 January 2021; Accepted 14 March 2021
Available online 19 March 2021

methods are currently required – or indeed available – to diagnose the disorder in clinical settings. The diagnosis – like other psychiatric and neurodevelopmental conditions – is based on observed and reported behavioral symptoms, exclusion of alternative medical or psychiatric disorders, developmental history, and significant impairment of daily functioning. Thus, the current "gold standard" relies on informants' and clinicians' interpretations, and reliable diagnostic decisions are therefore time consuming and requires considerable experience and expertise (Ewen et al., 2019; Spitzer, 1983).

Many attempts have been made to discover valid biomarkers for psychiatric disorders (Ritsner and Strous, 2010), including ADHD (Ewen et al., 2019; Helgadottir et al., 2015; Kropotov, 2016a; McLoughlin et al., 2014; Muller et al., 2019). A biomarker is a characteristic or indication of a medical state that can be measured objectively, be reliably reproduced, and that potentially could aid an (early) diagnosis (Kropotov, 2016b). In 2012, a consensus report of the World Federation of ADHD (Thome et al., 2012) concluded that a biomarker for ADHD must have 1) diagnostic sensitivity ≥80% for detecting ADHD and 2) a diagnostic specificity ≥80% for distinguishing ADHD from other disorders with ADHD-like symptoms. In addition, the biomarker must be 3) reliable, 4) affordable, and 5) it should be confirmed by at least two independent studies. According to Ewen et al. (2019), the work of finding biomarkers for neurodevelopmental disorders like ADHD is far from complete. For instance, the effects of age, gender, comorbidities, and different etiological mechanisms behind similar symptoms are not fully understood.

Neuropsychological tests can be considered as indirect measures of brain function (Thome et al., 2012). Many studies report that most patients with ADHD score lower than controls on tests of executive function although sensitivity and specificity of such impairments are moderate at the individual level (Willcutt et al., 2005). For instance, continuous performance tests (CPTs) are widely used to assess variables of attention demanding executive control. Such testing is computerized and typically generates scores of omission and commission errors, reaction time and reaction time variability. In particular, omissions and reaction time variability have been associated with ADHD (Egeland and Kovalik-Gran, 2010), however, the results are controversial among experts in the field (Russell A Barkley and Eme, 2019). In a review of the clinical utility of CPTs in pediatric ADHD (Hall et al., 2016) the authors report mixed findings, underscoring that CPTs cannot be used alone to diagnose ADHD. Such tests may however inform on aspects of the disorder not captured by rating scales and interviews.

Suggestions have been made that more direct measures of neural activity might improve the diagnostic accuracy (e.g., Kropotov, 2016b). Many studies applying EEG in children with ADHD-like symptoms have been published, starting in 1938 with a study on "behavior problem children" (Jasper et al., 1938). Two comprehensive reviews of research on quantitative EEG in ADHD have been published (Barry et al., 2003; Clarke et al., 2020). Several studies have aspired to identify event related potentials (ERPs) as biomarkers for ADHD (Lenartowicz and Loo, 2014; Szuromi et al., 2011). ERPs are cerebral generated electrical voltages recorded with EEG-equipment on the scalp in response to specific time-locked stimuli or responses, for example events in a CPT (Kropotov, 2016b). The ERPs are recorded with a high time resolution (Luck, 2014) and may therefore be able to give information about the neural activity related to behavioral parameters in tasks (Lenartowicz and Loo, 2014). Numerous research findings indicate that the ERP components most consistently associated with ADHD in cued visual go/no-go tasks are: the Contingent Negative variation (CNV), reflecting preparation of response, N2 and P3 no-go, associated with conflict monitoring, impulse control and allocation of attentional resources, P2 and P3b /P3 go (selective attention) and cue P3 as a measure of target identification (Johnstone et al., 2013; Kropotov, 2016b). In a recent meta-analysis (Kaiser et al., 2020) the authors report that ADHD patients have smaller Cue P3 amplitudes and longer latencies, longer P3 go latencies, smaller amplitudes, and longer latencies in P3 no-go, and smaller CNV amplitudes.

However, the results in ERP studies comparing ADHD and typically developing controls are also not fully consistent. Some studies find significant differences between ADHD and controls (Overtoom et al., 1998; Szuromi et al., 2011), while other studies do not (Banaschewski et al., 2004; Lau-Zhu et al., 2019). Some of the inconsistent results may be due to large age spans. Hager et al. (2020) showed that ERPs seem to capture different aspects of functioning at different ages in pediatric ADHD and for this reason age needs to be considered in interpretations of ERP results. The authors used age groups of 9–12 and 13–17 and concluded that in particular the P3 no-go showed age-dependent correlations. This finding motivated the choice of age group in this study.

Moreover, most studies attempting to evaluate ERP indices as diagnostic biomarkers for ADHD have relied on a single ERP component. As underscored by (Lenartowicz and Loo, 2014), a heterogeneous disorder like ADHD is probably not captured by a single neurophysiological variable. Therefore, the authors propose that combining several EEG based measures might increase the accuracy of the biomarker. Mueller et al. (2011), applying the same paradigm as we do, combined ERP variables to support the diagnostic process in adult ADHD. A large number of ERP variables were fed into a vector machine learning algorithm to calculate the combination of variables that best discriminated between adult ADHD and healthy controls. An accuracy of 91% in a sample of 150 adults was found. This combination index was later applied to a new independent sample with similar classification accuracy.

Above, we have referred studies applying CPTs to help diagnose ADHD, and studies using ERPs for diagnostic purposes. The aim of the present study was to develop a diagnostic index for ADHD in children age 9–12 years by combining behavioral variables from a visual cued go/no-go task (VCPT) and ERPs registered during the task. We reasoned that such an index, as compared with a single ERP or CPT variable, would be better able to identify children with this heterogeneous disorder. Specifically, we hypothesized that this index discriminates between ADHD and TDC controls with an ES of Cohen's $d > 0.80$ considered as the minimum for clinical application. We also hypothesized that the index applied to an independent sample would confirm the finding of the main study.

## 2. Methods

### 2.1. Participants and diagnostic procedures

*Patients:* The 61 ADHD patients aged 9–12 years had been diagnosed at three different child psychiatry outpatient clinics in Norway in accordance with DSM 5 criteria. The patients included were the total number of available cases who fulfilled all inclusion criteria. The neuropsychiatric team, Åsebråten, Fredrikstad was the main clinic. Some patients had participated in earlier studies applying DSM IV. Specialists in clinical psychology or psychiatry were responsible for diagnostic conclusions. The patients were screened medically, i.e. a general examination of physical health and exclusion of somatic conditions as alternative explanations of the symptoms. Medical and psychosocial background information was recorded.

Exclusion criteria were IQs <70, a diagnosed brain injury / neurological disorder, and/or autism spectrum disorder (ASD). ASD was excluded for two reasons: DSM IV did not allow diagnosing ASD along with ADHD, and we plan a study comparing ADHD and ASD. Patients with common comorbidities such as learning disabilities, language disorders, Tourette syndrome, behavioral- and emotional disorders were not excluded. The incidences of such comorbidities were not systematically reported from the participating clinics and are not included in Table 1. (In previous research projects from then main clinic, with a partly overlapping (about 30%) sample, comorbid conditions were seen in most cases (Ogrim et al., 2012), which is in accordance with the literature).

**Table 1**
Characteristics of the ADHD and the TDC samples*.

| | Age (SD) | SexMale: *N*,% | ADHD-C(*N*,%) | ADHD-I(*N*,%) | VIQ (SD)*N* = 48 | PIQ (SD)*N* = 48 | Tot IQ (SD)*N* = 48 |
|---|---|---|---|---|---|---|---|
| ADHD (*N* = 61) | 10.52 (1.2) | Male: *N* = 37 (61%) | *N* = 41<br>79% | *N* = 11<br>21% | 94 (13) | 98 (20) | 96 (13) |
| TDC (*N* = 69) | 10.58 (1.2) | Male: *N* = 42 (61%) | | | | | |
| Difference | *P* = 0.79<br>Ns. | Chi-square: sig.=1.00 Ns. | | | | | |

TDC: Control subjects with no diagnosed neurodevelopmental disorders. ADHD-C: Combined presentation. ADHD-I: Inattentive presentation. VIQ: Verbal IQ. PIQ: Performance IQ. Tot IQ: Total IQ. Ns.: Non-significant. SD: Standard deviation.
NOTE: Classification of presentation available for 52 of 61 patients. IQs available for 48 patients.
*: The prevalence of common comorbidities in the ADHD group not available – see text.

None of the patients were on ADHD medication when tested, and did not use alcohol, tobacco or illegal drugs (age 9–12 years).

*Diagnostic methods and instruments:* The diagnostic instruments varied to a certain degree between clinics, for example which ADHD rating scale or clinical interview method that was used, but the DSM 5 (or IV) criteria were applied by experienced specialists who made global diagnostic conclusions from the available information.

The following instruments were applied: Broad spectrum clinical interviews: Kiddie-SADS (Kaufman et al., 1997), Development and Well Being Assessment www.dawba.com, Rating scales: Achenbach System of Empirically Based Assessment(Achenbach and Rescorla, 2007) – mapping internalizing and externalizing symptoms, Behavior Rating Inventory of Executive Function (BRIEF) (Gioia et al., 2000), Conners' 3 rating scale (ADHD and common comorbidities) (Kao and Thomas, 2010), ADHD rating scale (DuPaul et al., 1998), the 5–15 (FTF) (broad spectrum questionnaire for ages 5 to 15, eight domains, www.5–15.org; Kadesjö et al., 2004). The children were observed in clinic or school, and self-report scales and/or child interviews were used in most cases. Intelligence was assessed using either WISC-IV (Wechsler Intelligence Scale for Children, 4th edition(David Wechsler, 2012) or WASI (Wechsler Abbreviated Scale for Intelligence (D Wechsler, 1999). In a few cases WISC-V was applied (David Wechsler, 2014).

*Controls:* The comparison cases (*n* = 69) were the cases available in our database of controls when matching for gender and age. Criteria for inclusion were no formal diagnoses of learning disabilities, neurological or developmental disorders such as epilepsy, ADHD, ASD or Tourette syndrome. These controls are denoted as TDC (typically developing children) in this study. About 50% of the group was recruited at two Norwegian sites as comparison cases for this and associated projects. The remaining TDCs were selected from the Human Brain Indices (HBi, www.hbimed.com) database. They were drawn from this database based on gender and age. Their test scores were revealed *after* the inclusion had taken place. The inclusion criteria in HBi are identical to the TDC criteria described above. The majority of the HBi controls were tested in Switzerland in 2003–2006. Identical equipment was used at the recording sites. Calibration checks have shown that the latencies of ERP components recorded at the Norwegian sites were 20 ms. faster than in Switzerland 2003–2006. Latencies are not included in the present study. The amplitudes of the components did not differ between sites. The Swiss TDC and their parents were informed, and accepted, that anonymous test data would be included in a database for clinical and research purposes.

### 2.2. Assessment of ERPs and scores on the go/no-go task

EEG was recorded using a Mitsar 201 19-channel EEG system (www.mitsar-medical.com). The registration consisted of 3 min eyes-closed condition, 3 min eyes-opened, and 20 min during a cued go/no-go task.

Pictures of animals (*a*), plants (*p*) and humans (*h*) were used. The trials consisted of paired stimuli with inter-stimulus intervals of 1000 ms and inter-trial intervals of 3000 ms. Four trial categories were used: *a-a, a-p, p-p, and p-h.* (Fig. 1). Subjects were instructed to respond (press mouse button) to *a-a* trials, and to be accurate, but also fast. The overall luminance and the image sizes of animals and plants were about equal.

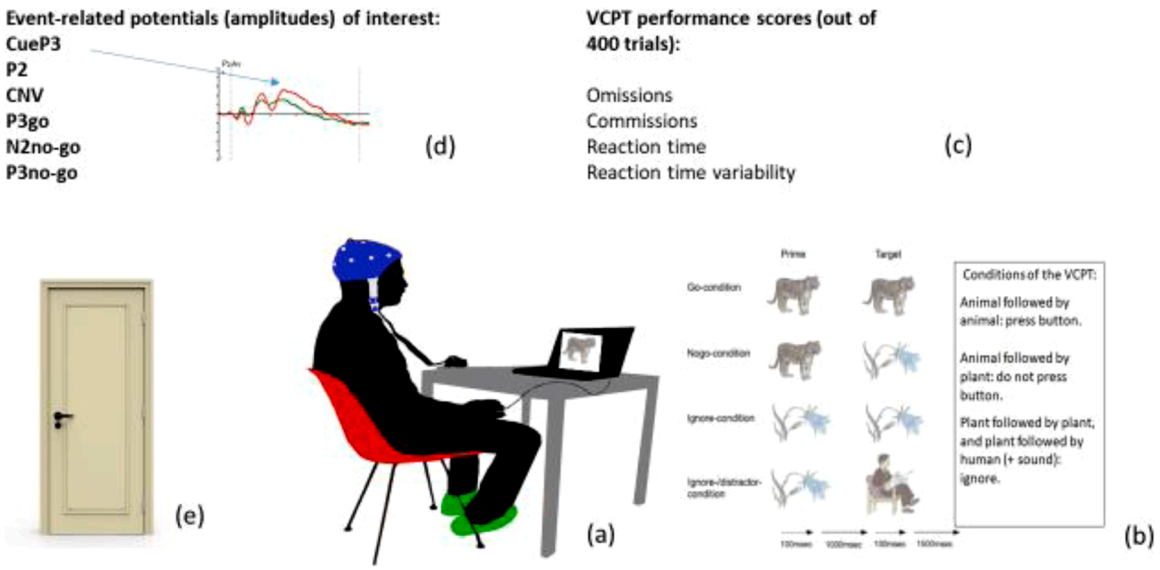


**Fig. 1.** Conditions of the VCPT. (a) is an illustration of the test situation. (b) is an illustration of the test stimuli in the VCPT. 400 trials. (c) details the neuropsychological performance scores obtained in the VCPT. (d) shows the ERP component Cue P3 in a patient (green line) and the same component for the age-matched control group. (e) door to test-room. .

To avoid habituation, 20 different animal, plant, and human images are used in the test. In the *p-h* trials, novel sounds were presented along with human images. These novel sounds produced an orientation reaction, confirmed by elicitation of the novelty ERP wave.

The go/no-go task generates scores of omissions (number of go-responses omitted), commissions (number of keystrokes in no-go condition), reaction time in milliseconds in the go condition, and reaction time variability (standard deviation of reaction time in go-condition).

The four blocks consist of a pseudo-random presentation of 400 trials with 100 unique trials in each category: Animal-animal (a-a, go-condition) – animal-plant (a-p, no-go condition) – plant-plant (p-p, ignore condition) and plan-human + sound (p-h, ignore / distractor condition). Participants practiced the task before the recording started. They sat upright in a comfortable chair looking at a 17-inch CRT computer screen positioned 1.5 m in front of them. Pressing the button to *a-a* pairs within 200–1000 ms. after presentation of the second stimulus was registered as a correct response. Failure to respond to *a-a* pairs within this time interval was considered an omission error. Impulsive responses to *a-p* pairs were considered commission errors. Two short breaks, after 150 and 300 trials, were given.

Input signals were referenced to earlobe electrodes, filtered between 0.5 Hz and 50 Hz, and digitized at a sampling rate of 250 Hz, with impedance kept below 5 kΩ for all 19 electrodes. An electrode cap with tin electrodes (Electro-cap International, Eaton OH, USA) was used. The electrodes were placed in accordance with the international 10–20 system.

EEG data were re-referenced offline to the common average montage prior to data processing. Eye-blink artifacts were corrected by zeroing the activation curves of individual independent components extracted by Independent Component Analysis (Infomax algorithm) and corresponded to eye-blink topographies (Jung et al., 2000; Vigario, 1997). EEG epochs with excessive amplitude (100 μV) and/or excessively fast (35 μV in 20–35 Hz band) and slow (50 μV in 0–1 Hz band) frequency activities were automatically excluded from analysis.

Local peak amplitudes of the ERP components were measured individually at the electrodes where the components were observed to be strongest in the grand- average (GAF) ERPs of the total group. Local peak amplitude refers to the point within the defined time window for the component of interest with the largest amplitude, which is surrounded on both sides by lower voltages, thereby avoiding measuring the offset of preceding or onset of following components (Luck, 2014).

The ERP waves, sites and time intervals after stimulus 1 were: P2 at O2 (280–450 ms.), cue-P3 at P3 (400–600 ms.), CNV at Cz or Pz (1000–1100 ms.) (The site with the strongest CNV amplitude of the participant was used).

The ERP waves, sites, and time intervals after stimulus 2 were: P3go at Pz (200–400 ms.), N2no-go at Fz (230–400 ms.), P3no-go at Cz, (330–500 ms.).

### 2.3. Statistical methods and calculation of the diagnostic index

In WinEEG program Grand Average ERP Files (GAFs) were computed for patients and controls. The comparison option automatically estimates statistical significance of differences between the grand average ERPs. As described above, the following ERPs were selected: P2, Cue-P3, CNV, P3go, N2no-go and P3no-go. These components were individually scored, as described above, and exported to Statistical Package for the Social Sciences (SPSS) for further analyses.

The SPSS files were checked for outliers (≥ 3 SD above/below mean). There were few outliers; from 0 to 4 for the variables included, with no difference in the number of outliers between patients and controls. All outliers were "moved to the closest neighbor" (Pallant, 2013). Table 2 shows the means, SDs, *p*-values of the differences, and effect sizes (*d*s) for the variables included. Nine variables were significantly different between patients and controls with $d \geq .5$ and were used in further analyses. None of these variables correlated with the others at coefficients higher (below) 0.7 (−0.7), thereby avoiding inclusion of variables with little unique contributions.

**Table 2**
Variables discriminating significantly between patients and controls with effect sizes ≥0.5.

| | ADHD *N* = 61*M (SD)* | TDC *N* = 69*M (SD)* | Differencesig. | Effect size (Cohen's *d*) |
|---|---|---|---|---|
| **Omissions** | 12.11 (9.71) | 4.54 (4.81) | 7.57 $p<0.001$ | 1.00 |
| **Commissions** | 4.57 (5.45) | 2.39 (2.96) | 2.18 $p = 0.005$ | 0.50 |
| **Reaction time variability** | 14.84 (4.22) | 12.55 (3.43) | 2.29 $p = 0.001$ | 0.60 |
| **Amplitude of P2 at site O2** | 8.087 (6.42) | 11.706 (6.27) | −3.619 $p = 0.001$ | 0.57 |
| **Amplitude of CueP3** | 2.802 (2.16) | 4.933 (2.95) | −2.131 $p<0.001$ | 0.82 |
| **Amplitude of CNV** | −2.765 (1.79) | −4.006 (2.02) | 1.241 $p<0.001$ | 0.65 |
| **Amplitude of P3go** | 9.659 (5.02) | 12.770 (4.45) | −3.111 $p<0.001$ | 0.66 |
| **Amplitude of N2no-go** | −9.413 (4.80) | −12.679 (4.26) | 3.266 $p<0.001$ | 0.72 |
| **Amplitude of P3no-go** | 7.164 (5.50) | 10.159 (7.15) | −2.995 $p = 0.009$ | 0.52 |
| **Log10 ADHD Index** | **1.0713 (0.12615)** | **0.8990 (0.10712)** | **0.17232 $p<0.0001$** | **1.47** |
| **ADHD Index (no Log10 correction)** | 12.276 (3.52) | 8.181 (2.23) | 4.095 $p<0.0001$ | 1.39 |
| **VCPT behavior Log10 Index** | 0.5676 (0.199) | 0.4226 (0.156) | $P<0.0001$ | 0.81 |

M: Mean. SD: Standard deviation. P2, Cue P3, CNV, P3go, N2no-go, P3no-go: ERP components described in Introduction. VCPT behavior Log10 Index: An index based on the three behavioral VCPT variables (omissions, commissions, reaction time variability).
Variables checked but not included (not significant).
Reaction time: ADHD:423 ms. (SD=71). TDC:433 ms. (SD=91). NS (sig. =0.50).
ERP component N1 at site O2: ADHD=15.27 mv. TDC=16.67 mv. $p = 0.26$ (WinEEG calculations).

Percentile cut-off scores, *based on the control group*, were computed for these nine variables. Score 1 was set for all scores within the interval 1–80 percentile. Score 2: 80–90 percentile; score 3: 90–95 percentile; score 4: 95–98 percentile and score 5: 98–100 percentile. Percentiles are sometimes used when scores are not normally distributed (see for example www.5–15.org manual). A five-point scale like this is also clinically meaningful highlighting the deviant scores. The individual's final score on a scale was the percentile (P) score multiplied by the effect size (*d*) of that scale. (If the P score was 4 and $d = 0.8$ the final score was 3.2). These final scores were summed to the diagnostic index score of each individual.

The nine variables (omissions, commissions, and RT-variability in VCPT test and six ERPs) have all been associated with ADHD / cognitive control in the literature. On the other hand, comparing nine variables in a sample that consists of 61 patients and 69 controls increases the risk of type 1 error. Although not necessary when selection of files is based on *d,* we applied Bonferroni adjustment (dividing the significance level 0.05 by 9). All variables were still significant.

The diagnostic index was not normally distributed. To apply statistical methods based on normal distributions we applied Log10 correction (Pallant, 2013) resulting in a normal distribution.

The clinical utility of a diagnostic predictor depends on the representativity of the sample used. We therefore applied this diagnostic index, based on the controls in the main study, to a new sample. Twenty ADHD patients from the main clinic (Fredrikstad) who had participated in previous research were selected to match the study sample in age and gender. Their test scores were revealed *after* the selection. The control group consisted of 21 healthy controls from the HBi database, matched for age and gender. As for the patients, the test scores were recorded

**Table 3**
Percentile cut-off scores for the 9 significant variables based on the control group.

| | <80P sc.1 | 80–90P sc.2 | 90–95P sc.3 | 95–98P sc.4 | >98P sc.5 | >20P sc.1 | 10P sc.2 | 5P sc.3 | 2P sc.4 | <2P sc.5 |
|---|---|---|---|---|---|---|---|---|---|---|
| Omis-sions | <7 | 7–13 | 13–18 | <18 | <18 | | | | | |
| Com-missions | <6 | 6–8 | <8 | <8 | <8 | | | | | |
| RT-var | <15.6 | 15.6- 17.3 | 17.3- 19.6 | >19.6 | >19.6 | | | | | |
| P2O2 amp | >6.68 | 6.68–4.09 | 4.09–2.54 | 2.54–2.75 | <−2.75 | | | | | |
| CueP3 amp | >3.02 | 3.02–1.77 | 1.77–0.14 | 0.14–1.33 | <−1.33 | | | | | |
| P3go amp | >9.84 | 9.84–8.08 | 8.08–6.22 | 6.22–3.40 | <3.40 | | | | | |
| P3no-go amp | >4.40 | 4.40–2.49 | 2.49–1.27 | 1.27–3.47 | <−3.47 | | | | | |
| CNV amp | | | | | | <−2.35 | −2.35–1.77 | −1.77–1.41 | −1.41–0.51 | >−0.51 |
| N2no-go amp | | | | | | <−10.48 | −10.48–8.08 | −8.08–4.20 | −4.20–2.14 | >−2.14 |

P: Percentile. *Sc*.: Score. RT-var: Reaction Time variability. O2P2 amp: Amplitude of the ERP component P2 at site O2. Cue P3 amp: Amplitude of ERP component Cue P3. CNV amp: Amplitude of ERP component Contingent Negative Variation. P3go amp: Amplitude of the ERP component P3go. N2no-go amp: Amplitude of ERP component N2no-go. P3no-go amp: Amplitude of ERP component P3no-go.
NOTE: For the first 7 variables a high score reflects deviance from normality. For CNV amp and N2No-go a high score (close to or above 0) reflects deviance.

*after* the selection. (The criteria for inclusion in the HBi database are described in paragraph "Participants and diagnostic procedures").

We report means, SDs, differences, *p*-values, and effect sizes (*ds*) for all variables, including the diagnostic index. For this index we also applied the receiver operating characteristic (ROC) analysis and report area under curve (AUC). A scale discriminating between groups with AUC = 0.9 to 1.0 is considered excellent, 0.8 to 0.9 is good, 0.7 to 0.8 is fair, 0.6 to 0.7 is poor, and ≤0.5 indicates failure. Statistical analyses were performed using SPSS Version 21 (http://www.spss.com), with significance = 5%. Cohen's *d* effect sizes were calculated with correction for different sample sizes.

The project was approved by the regional committee for medical and health research ethics (REK). All parents and children/adolescents received oral and written information about this research project and signed a written consent to participate. (REK 2016/1453). As described in paragraph "Participants and diagnostic procedures" the controls from the Hbi database were tested in 2003–2006 and gave informed consent that their anonymous test scores can be used for clinical and research purposes.

## 3. Results

Table 2 shows means, standard deviations, significance of differences between patients and controls. For behavioral variables, the following effect sizes (Cohen's *d*) were obtained when comparing the groups (details in Table 2): Omissions (1.00), commissions (0.50), Reaction time (RT) (n.s.), RT variability (0.60). For the ERP-based measures, the following effects were seen: P1 (ns), P2 (0.57), Cue P3 (0.82), CNV (0.65), P3 go (0.66), N2 no-go (0.72), P3 no-go (0.52). As described in the Methods section percentile cut-off scores for the nine variables, based on the control group, were calculated.

The global diagnostic index was transformed with Log10 correction (see methods section). The calculated diagnostic index discriminated significantly between patients and controls with a large ES ($d = 1.47$). We also report the index without the log10 correction, which was slightly lower; $d = 1.39$. An effect size based on the three behavioral VCPT variables (omissions, commissions, and variability) was also calculated, applying the percentile /Log10 method described. The effect size *d* was 0.81.

For the global diagnostic ADHD index, we also applied the receiver operating characteristic (ROC) analysis and report area under curve (AUC), Fig. 2. The accuracy of the log10 Index was 84.4% which is considered as good (details in Methods). (When the ROC analysis was applied to the Index without Log10 correction the accuracy was identical).

### 3.1. The diagnostic index applied to a new sample

The diagnostic index, based on the controls in the main study, was applied to a new independent sample of 20 children with ADHD and 21 controls, matched for age and gender. On the nine variables each participant received a score based on his/her raw score which was converted to a percentile score (1–5) and multiplied with the *d* of the variable in focus. The index was the sum of these scores. The same percentile / Log10 procedure was applied. ES ($d = 3.03$) and accuracy (97.7%) was very high (see Table 4).

The ROC analysis was applied to this independent test sample (Fig. 2).

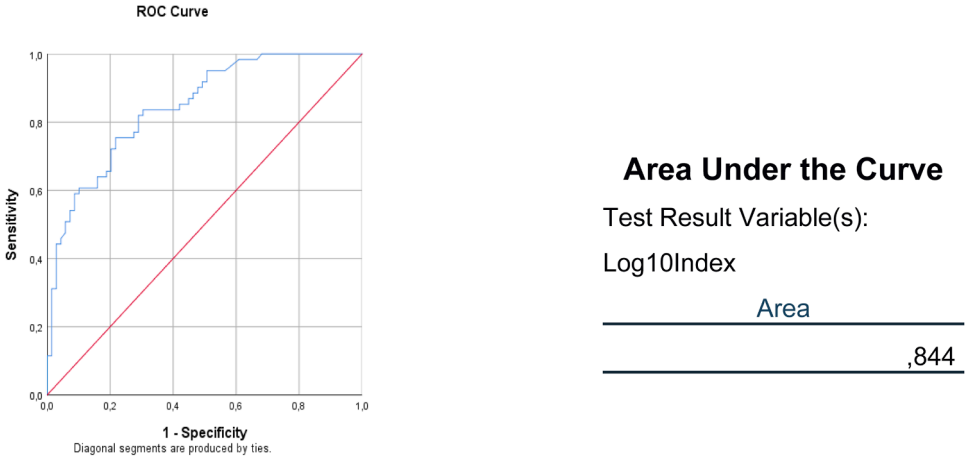


**Fig. 2.** ROC curve for the diagnostic index. Fig. 2 shows the accuracy of the diagnostic index; 84.4%.

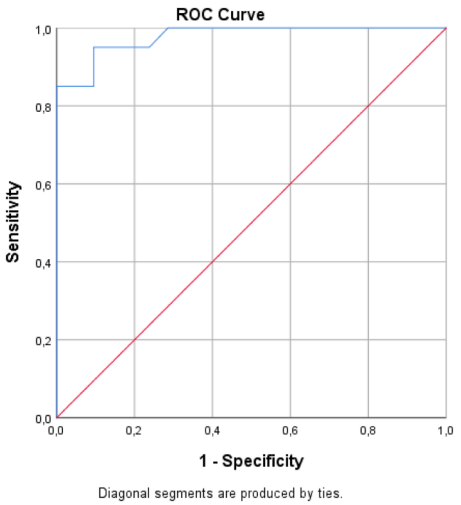


**Fig 3.** ROC curve for the diagnostic index in the independent test sample.

**Table 4**
Differences between ADHD and TDC in an independent test sample.

| | ADHD $N = 20$M (SD) | TDC $N = 21$M (SD) | Effect size (Cohen's $d$) |
|---|---|---|---|
| **Omissions** | 23.90 (16.6) | 2.29 (1.8) | 1.89 |
| **Commissions** | 12.35 (8.8) | 0.67 (0.9) | 1.83 |
| **RT-var** | 16.00 (5.2) | 11.95 (3.2) | 0.97 |
| **P2O2 amp** | 10.40 (4.2) | 10.48 (5.6) | 0.01 |
| **CueP3 amp** | 3.15 (2.7) | 5.84 (2.4) | 1.06 |
| **CNV amp** | −3.24 (1.4) | −4.57 (1.3) | 1.40 |
| **P3go amp** | 10.22 (5.0) | 14.06 (3.1) | 0.92 |
| **N2no-go amp** | −9.51 (3.5) | −12.70 (4.8) | 1.17 |
| **P3no-go amp** | 6.36 (5.0) | 9.69 (6.3) | 0.52 |
| **Log10 Index** | **1.0958 (0.102)** | **0.8388 (0.063)** | **3.03** |

The abbreviations are the same as in Table 2. In Table 4 the p-values of the differences are not reported because of small sample sizes.

### 3.2. Figures of ERP differences between ADHD and TDC

The ERP differences between ADHD and TDC are shown in Figs. 4,5 and 6. In these figures "Test" (left side) refers to the main study ($N = 61$ ADHD and 69 TDC) and "Retest" (right side) refers to the independent retest sample ($N = 20$ ADHD and 21 TDC). Fig. 4 shows the group differences after stimulus 1 (CueP3 and CNV). Fig. 5 shows the differences in go condition (after stimulus 2). Fig. 6 shows the differences in no-go condition; N2no-go and P3no-go.

### 3.3. Follow up analyses of possible effects of age, sex, and presentation

Because EEG, ERPs and CPT performance are age-related, we analyzed correlations between the 9 variables and age. Two variables correlated significantly, but weakly, with age: Omissions: $r = -0.263^{**}$ ($p = 0.003$) and commissions: $r = -0.177^{*}$ ($p = 0.044$).

An exploratory analysis showed that only the P3 no-go amplitude was significantly different in boys and girls (larger in boys). The diagnostic index was not significantly different between boys and girls ($p = 0.684$).

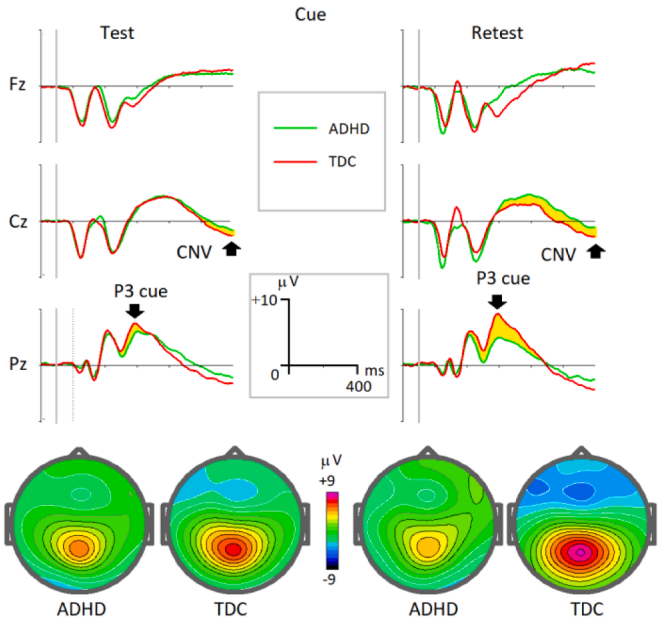


**Fig. 4.** Grand averaged event-related potentials (ERPs) in the cued go/no-go task in response to cue computed for two groups of typically developing children (TDC) and two groups of children with Attention Deficit Hyperactivity Disorder (ADHD) for the test study on the left (N1 (TDC) =69, N2(ADHD) = 61) and for the retest study on the right (N1 (TDC)= 21, N2(ADHD)=20). *Top:* ERPs for TDC (red) and ADHD (green) groups with the ERP components CueP3 and CNV marked by arrows on the waveforms. *Bottom:* Maps of the P3cue components taken at the maximum (around 500 ms).

Three of the nine variables were significantly different when comparing ADHD-C and ADHD-I. The combined type had more commissions, a faster reaction time, and a larger P3 no-go amplitude.

## 4. Discussion

This study aimed to compute a diagnostic index for ADHD by combining behavioral test scores from a cued visual go/no-go task and Event Related Potentials (ERPs). Initially, nine variables were identified that discriminated significantly between the groups with effect sizes (Cohen's $d$) $\geq 0.5$. Then, percentile cut-off scores (1–5) based on the typically developing child (TDC) group, were computed for these nine variables. An individual diagnostic index score was then calculated

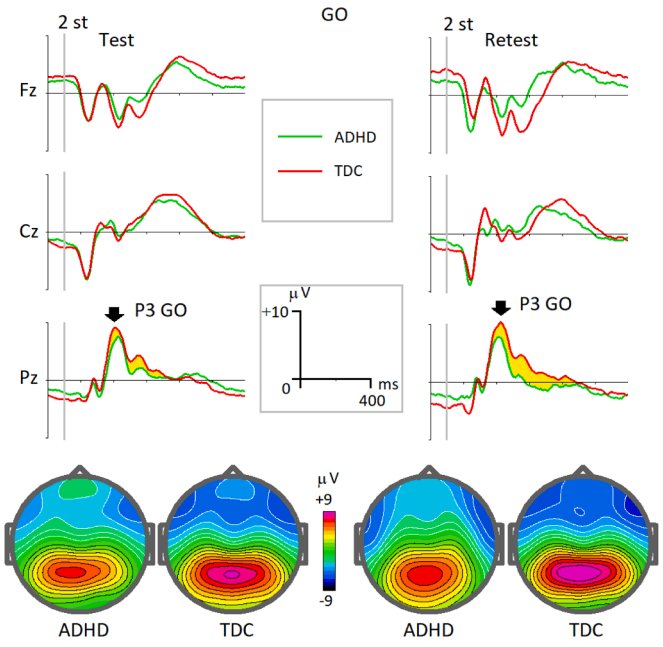


**Fig. 5.** Grand averaged event-related potentials (ERPs) in the cued go/no-go task in response to go stimuli computed for two groups of traditionally developed children (TDC) and two groups of children with Attention Deficit Hyperactivity Disorder (ADHD) for the test study on the left (N1 (TDC) =69, N2 (ADHD) = 61) and for the retest study on the right (N1 (TDC)= 21, N2 (ADHD)=20). *Top:* ERPs for TDC (red) and ADHD (green) groups with the ERP component P3go marked by arrows on the waveforms. *Bottom:* Maps of the P3go component taken at the maximum (around 320 ms).

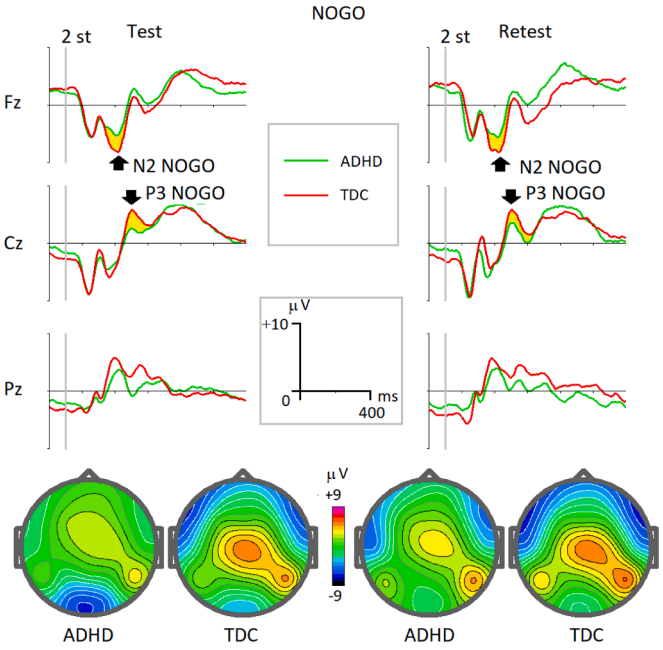


**Fig. 6.** Grand averaged event-related potentials (ERPs) in the cued go/n o-go task in response to no-go stimuli computed for two groups of traditionally developed children (TDC) and two groups of children with Attention Deficit Hyperactivity Disorder (ADHD) for the test study on the left (N1 (TDC) =69, N2 (ADHD) = 61) and for the retest study on the right (N1 (TDC)= 21, N2 (ADHD)=20). *Top:* ERPs for TDC (red) and ADHD (green) groups with the ERP component P3no-go marked by arrows on the waveforms. *Bottom:* Maps of the P3no-go component taken at the maximum (400 ms).

based on the sum of weighted scores on these variables. We found that the index discriminated significantly between patients and controls with a large effect size ($d = 1.47$) and an accuracy of 84.4%. This index, based on the 69 TDC, was then applied to a smaller independent sample of 20 ADHD children and 21 TDC, matched for gender and age. In this replication sample, the effect size was $d = 3.07$ and the accuracy was 97%.

Most studies on diagnostic biomarkers for ADHD have focused on the discriminatory power of a single variable, not a profile of variables as was explored in the current study. We are not aware of any prior studies combining several event-related potentials (ERPs) and behavioral variables from a CPT to differentiate ADHD from a control group of typically developing children (TDC), although some previous authors suggest combining electrophysiological variables to improve accuracy (Clarke et al., 2020; Lenartowicz and Loo, 2014). In view of the clinical heterogeneity of ADHD both in presentations and etiology, we were guided by the idea that considerable benefits potentially can be gained from multivariable analyses (Lenartowicz and Loo, 2014).

CPTs are today widely used in clinical settings to examine attentional functions in patients assessed for possible ADHD. In the present study we found that omissions, commissions and reaction time variability significantly discriminated between patients and controls. We also computed an index based om the three CPT variables that were significantly different in the two groups. ES of this index was $d = 0.81$. Importantly, our results show that the inclusion of ERP variables improve classification considerably compared with CPT (behavioral) data only.

With regards to specific ERP variables, several previous studies have reported smaller amplitudes in ADHD compared with TDC for components such as Cue P3, CNV, P3 go, N2 no-go and P3 no-go (e.g., Brandeis et al., 2002; Groom et al., 2008; Johnstone et al., 2013; Kropotov, 2016b). Overall, our data replicate these prior findings.

In terms of clinical implications, the current procedure for diagnosing ADHD (and psychiatric disorders in general) are based on interviews, history, rating scales and observations. We believe that a research-based supplementary biomarker could potentially increase diagnostic validity and perhaps reduce time needed for assessment and intervention conclusions. Importantly, the index calculated in this study is based on scores from the test procedure and can be directly applied to clinical cases using a percentile table. Thus, once standardized, there is no need for extensive technical skills.

To improve diagnostics in child- and adolescent psychiatry using a biomarker approach, the proposed biomarkers need to be evaluated carefully according to established criteria, e.g., the criteria for biomarker validation of the World Federation of ADHD (Ewen et al., 2019). Although the current findings provide initial support for the CPT/ERP approach in biomarker development in ADHD, it is important to underscore that the reported index is not necessarily helpful in discriminating ADHD from other clinical categories such as learning disabilities, autism spectrum or anxiety disorders. Relatedly, most ADHD patients, including the ones participating in our study, also have other diagnoses that need to be assessed to secure treatment that meet the needs of the patients (Gillberg, 2010). Common comorbidities were not excluded in this study to best reflect the clinical reality of ADHD patients. Taken together, an important avenue for future research is to examine whether the reported index is specifically tapping (pure) ADHD as compared to a broader clinical cluster of neurodevelopmental traits and problems (cf., Gillberg, 2010).

### 4.1. Limitations

Besides those mentioned above, there are some further noteworthy limitations of the current study. First, ADHD is a heterogenous diagnosis. Our sample consists of patients diagnosed with the combined and the inattentive presentations. The diagnostic index is based on the whole group to secure the necessary statistical power. When we screened for differences between the presentations, some variables were significantly different across groups with different presentations. In a larger sample

separate indexes for the two presentations could be calculated. Second, the TDC group is a mixture of Norwegian youngsters tested within the last five years and controls from the HBi database tested in 2003–2006. Although the inclusion criteria and the equipment and test procedures are identical, we do not know the results of a study with controls that are demographically better matched with the patient group. Finally, the data used in this study was based on only one type of test, a VCPT. Future research is needed to determine the possible added value of including other types of executive tests, e.g., of planning or shifting.

### 4.2. Conclusion

The identified diagnostic index based on ERPs and a go/no-go task discriminated significantly, with a large effect size, between ADHD patients and children without diagnosed neurodevelopmental or psychiatric disorders (TDC). While acknowledging caveats and challenges, the results of the current study are highly promising, and further evaluations of this supplementary biomarker for ADHD are warranted. Ultimately, a robust, accurate and user-friendly ADHD index might not only improve diagnostic decision-making but can potentially also increase acceptance of the ADHD diagnosis in society.

## Funding

This study was funded by Sykehuset Østfold HF.

## Declaration of Competing Interest

The authors declare there are no conflicts of interest.

## References

Achenbach, T.M., & Rescorla, L. (2007). Achenbach system of empirically-based assessment. Retrieved from Mental Measurements Yearbook via EBSCOhost.

Diagnostic and Statistical Manual of Mental disorders: DSM-5, 2013. 5th edition. ed. American Psychiatric Publishing.

Banaschewski, T., Brandeis, D., Heinrich, H., Albrecht, B., Brunner, E., Rothenberger, A., 2004. Questioning inhibitory control as the specific deficit of ADHD–evidence from brain electrical activity. J Neural Transm (Vienna) 111 (7), 841–864. https://doi.org/10.1007/s00702-003-0040-8.

Barkley, R.A., Eme, R., 2019. Is Neuropsychological Testing Useful for Any Reason in the Evaluation of ADHD? A Rejoinder to Mapou. The ADHD Report 27 (5), 1–8.

Barry, R.J., Clarke, A.R., Johnstone, S.J., 2003. A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. Clin Neurophysiol 114 (2), 171–183. https://doi.org/10.1016/s1388-2457(02)00362-0.

Brandeis, D., Banaschewski, T., Baving, L., Georgiewa, P., Blanz, B., Warnke, A., Scheuerpflug, P., 2002. Multicenter P300 brain mapping of impaired attention to cues in hyperkinetic children. J Am Acad Child Adolesc Psychiatry 41 (8), 990–998. https://doi.org/10.1097/00004583-200208000-00018.

Clarke, A.R., Barry, R.J., Johnstone, S., 2020. Resting state EEG power research in Attention-Deficit/Hyperactivity Disorder: a review update. Clin Neurophysiol 131 (7), 1463–1479. https://doi.org/10.1016/j.clinph.2020.03.029.

DuPaul, G.J., Power, T.J., Anastopoulos, A.D., Reid, R., 1998. ADHD Rating Scale—IV: Checklists, norms, and Clinical Interpretation. Guilford Press.

Egeland, J., Kovalik-Gran, I., 2010. Measuring several aspects of attention in one test: the factor structure of Conners's Continuous Performance Test. J Atten Disord 13 (4), 339–346.

Ewen, J.B., Sweeney, J.A., Potter, W.Z., 2019. Conceptual, Regulatory and Strategic Imperatives in the Early Days of EEG-Based Biomarker Validation for Neurodevelopmental Disabilities. Front Integr Neurosci 13, 45. https://doi.org/10.3389/fnint.2019.00045.

Friedman, L.A., Rapoport, J.L., 2015. Brain development in ADHD. Curr Opin Neurobiol 30, 106–111. https://doi.org/10.1016/j.conb.2014.11.007.

Gillberg, C., 2010. The ESSENCE in child psychiatry: early Symptomatic Syndromes Eliciting Neurodevelopmental Clinical Examinations. Res Dev Disabil 31 (6), 1543–1551. https://doi.org/10.1016/j.ridd.2010.06.002.

Gioia, G., Isquith, P.K., Guy, S.C., Kenworthy, L., 2000. Behavior Rating Inventory of Executive Function-Professional Manual. Pyychological Assessment Resources, Inc, Lutz.

Groom, M.J., Bates, A.T., Jackson, G.M., Calton, T.G., Liddle, P.F., Hollis, C., 2008. Event-related potentials in adolescents with schizophrenia and their siblings: a comparison with attention-deficit/hyperactivity disorder. Biol Psychiatry 63 (8), 784–792. https://doi.org/10.1016/j.biopsych.2007.09.018.

Hager, L.A., Ogrim, G., Danielsen, M., Billstedt, E., Gillberg, C., Asberg Johnels, J., 2020. Indexing Executive Functions with Test Scores, Parent Ratings and ERPs: how Do the Measures Relate in Children versus Adolescents with ADHD? Neuropsychiatr Dis Treat 16, 465–477. https://doi.org/10.2147/ndt.s230163.

Hall, C.L., Valentine, A.Z., Groom, M.J., Walker, G.M., Sayal, K., Daley, D., Hollis, C., 2016. The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: a systematic review. Eur Child Adolesc Psychiatry 25 (7), 677–699. https://doi.org/10.1007/s00787-015-0798-x.

Helgadottir, H., Gudmundsson, O.O., Baldursson, G., Magnusson, P., Blin, N., Brynjolfsdottir, B., Johnsen, K., 2015. Electroencephalography as a clinical tool for diagnosing and monitoring attention deficit hyperactivity disorder: a cross-sectional study. BMJ Open 5 (1), e005500. https://doi.org/10.1136/bmjopen-2014-005500.

Jasper, H.H., Solomon, P., Bradley, C., 1938. Electroencephalographic analyses of behavior problem children. American Journal of Psychiatry 95 (3), 641–658.

Johnstone, S.J., Barry, R.J., Clarke, A.R., 2013. Ten years on: a follow-up review of ERP research in attention-deficit/hyperactivity disorder. Clin Neurophysiol 124 (4), 644–657. https://doi.org/10.1016/j.clinph.2012.09.006.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin Neurophysiol 111 (10), 1745–1758. https://doi.org/10.1016/s1388-2457(00)00386-2.

Kadesjo, B., Gillberg, C., 2001. The comorbidity of ADHD in the general population of Swedish school-age children. J Child Psychol Psychiatry 42 (4), 487–492.

Kadesjö, B., Janols, L.O., Korkman, M., Mickelsson, K., Strand, G., Trillingsgaard, A., Gillberg, C., 2004. The FTF (Five to Fifteen): the development of a parent questionnaire for the assessment of ADHD and comorbid conditions. Eur Child Adolesc Psychiatry 13 (3), iii3–iii13.

Kaiser, A., Aggensteiner, P.M., Baumeister, S., Holz, N.E., Banaschewski, T., Brandeis, D., 2020. Earlier versus later cognitive event-related potentials (ERPs) in attention-deficit/hyperactivity disorder (ADHD): a meta-analysis. Neurosci Biobehav Rev 112, 117–134. https://doi.org/10.1016/j.neubiorev.2020.01.019.

Kao, G.S., Thomas, H.M., 2010. Test review: C. Keith Conners Conners 3rd edition Toronto, Ontario, Canada: multi-Health systems, 2008. J Psychoeduc Assess 28 (6), 598–602.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., Ryan, N., 1997. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36 (7), 980–988. https://doi.org/10.1097/00004583-199707000-00021.

Krieger, V., Amador-Campos, J.A., 2017. Assessment of executive function in ADHD adolescents: contribution of performance tests and rating scales. Child Neuropsychol 1–25. https://doi.org/10.1080/09297049.2017.1386781.

Kropotov, J.D., 2016a. Functional Neuromarkers For psychiatry: Applications For Diagnosis and Treatment.

Kropotov, J.D., 2016b. Functional Neuromarkers For psychiatry: Applications for Diagnosis and Treatment. Academic Press.

Lau-Zhu, A., Tye, C., Rijsdijk, F., McLoughlin, G., 2019. No evidence of associations between ADHD and event-related brain potentials from a continuous performance task in a population-based sample of adolescent twins. PLoS ONE 14 (10), e0223460. https://doi.org/10.1371/journal.pone.0223460.

Lenartowicz, A., Loo, S.K., 2014. Use of EEG to diagnose ADHD. Curr Psychiatry Rep 16 (11), 498. https://doi.org/10.1007/s11920-014-0498-0.

Luck, S.J., 2014. An Introduction to the Event-Related Potential Technique. MIT press.

McLoughlin, G., Makeig, S., Tsuang, M.T., 2014. In search of biomarkers in psychiatry: eEG-based measures of brain function. Am J Med Genet B Neuropsychiatr Genet 165b (2), 111–121. https://doi.org/10.1002/ajmg.b.32208.

Mueller, A., Candrian, G., Grane, V.A., Kropotov, J.D., Ponomarev, V.A., Baschera, G.M., 2011. Discriminating between ADHD adults and controls using independent ERP components and a support vector machine: a validation study. Nonlinear Biomed Phys 5, 5. https://doi.org/10.1186/1753-4631-5-5.

Muller, A., Vetsch, S., Pershin, I., Candrian, G., Baschera, G.M., Kropotov, J.D., Eich, D., 2019. EEG/ERP-based biomarker/neuroalgorithms in adults with ADHD: development, reliability, and application in clinical practice. World J Biol Psychiatry 1–11. https://doi.org/10.1080/15622975.2019.1605198.

Ogrim, G., Kropotov, J., Hestad, K., 2012. The quantitative EEG theta/beta ratio in attention deficit/hyperactivity disorder and normal controls: sensitivity, specificity, and behavioral correlates. Psychiatry Res 198 (3), 482–488. https://doi.org/10.1016/j.psychres.2011.12.041.

Overtoom, C.C., Verbaten, M.N., Kemner, C., Kenemans, J.L., van Engeland, H., Buitelaar, J.K., Koelega, H.S., 1998. Associations between event-related potentials and measures of attention and inhibition in the Continuous Performance Task in children with ADHD and normal controls. J Am Acad Child Adolesc Psychiatry 37 (9), 977–985. https://doi.org/10.1097/00004583-199809000-00018.

Pallant, J., 2013. SPSS Survival Manual. McGraw-Hill Education (UK).

Pawaskar, M., Fridman, M., Grebla, R., Madhoo, M., 2020. Comparison of quality of life, productivity, functioning and self-esteem in adults diagnosed with ADHD and with symptomatic ADHD. J Atten Disord 24 (1), 136–144. https://doi.org/10.1177/1087054719841129.

Ritsner, M.S., Strous, R.D., 2010. Neurocognitive deficits in schizophrenia are associated with alterations in blood levels of neurosteroids: a multiple regression analysis of findings from a double-blind, randomized, placebo-controlled, crossover trial with DHEA. J Psychiatr Res 44 (2), 75–80. https://doi.org/10.1016/j.jpsychires.2009.07.002.

Sayal, K., Prasad, V., Daley, D., Ford, T., Coghill, D., 2018. ADHD in children and young people: prevalence, care pathways, and service provision. Lancet Psychiatry 5 (2), 175–186. https://doi.org/10.1016/s2215-0366(17)30167-0.

Spitzer, R.L., 1983. Psychiatric diagnosis: are clinicians still necessary? Compr Psychiatry 24 (5), 399–411. https://doi.org/10.1016/0010-440x(83)90032-9.

Szuromi, B., Czobor, P., Komlosi, S., Bitter, I., 2011. P300 deficits in adults with attention deficit hyperactivity disorder: a meta-analysis. Psychol Med 41 (7), 1529–1538. https://doi.org/10.1017/s0033291710001996.

Thome, J., Ehlis, A.C., Fallgatter, A.J., Krauel, K., Lange, K.W., Riederer, P., Gerlach, M., 2012. Biomarkers for attention-deficit/hyperactivity disorder (ADHD). A consensus report of the WFSBP task force on biological markers and the World Federation of ADHD. World J Biol Psychiatry 13 (5), 379–400. https://doi.org/10.3109/15622975.2012.690535.

Vigario, R.N., 1997. Extraction of ocular artefacts from EEG using independent component analysis. Electroencephalogr Clin Neurophysiol 103 (3), 395–404. https://doi.org/10.1016/s0013-4694(97)00042-8.

Wechsler, D., 1999. Manual For the Wechsler abbreviated Intelligence Scale (WASI). The Psychological Corporation, San Antonio, TX.

Wechsler, D., 2012. Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition. Psychological Corporation, San Antonio, TX.

Wechsler, D., 2014. WISC-V: Technical and Interpretive Manual. NCS Pearson, Incorporated.

Willcutt, E.G., Doyle, A.E., Nigg, J.T., Faraone, S.V., Pennington, B.F., 2005. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry 57 (11), 1336–1346. https://doi.org/10.1016/j.biopsych.2005.02.006.