Biomarker-based treatment stratification in ADHD: An out-of-sample validation

Helena Voetterl
Research Institute Brainclinics, Brainclinics Foundation
Guido van Wingen
Amsterdam UMC, University of Amsterdam https://orcid.org/0000-0003-3076-5891
Giorgia Michelini
Department of Psychiatry & Biobehavioral Sciences, Devid Geffen School of Medicine, Semel Institute for Neuroscience & Human Behavior, University of California Los Angeles (UCLA)
Kristi Griffiths
Brain Dynamics Centre, Westmead Institute for Medical Research, The University of Sydney, Westmead https://orcid.org/0000-0002-7108-2272
Evian Gordon
Brain Resource Ltd, San Francisco
Roger deBeus
The University of North Carolina at Asheville https://orcid.org/0000-0001-9426-2690
Mayuresh Korgaonkar
Westmead Institute for Medical Research https://orcid.org/0000-0002-1339-2221
Sandra Loo
University of California Los Angeles
Donna Palmer
Total Brain, Sydney
Rien Breteler
Department of Clinical Psychology, Radboud University Nijmegen https://orcid.org/0000-0002-6698-3842
Damiaan Denys
Department of Psychiatry, Amsterdam UMC, University of Amsterdam
Eugene Arnold
Ohio State University https://orcid.org/0000-0002-0886-0692
Paul du Jour
Neuroscan, Dordrecht
Rosalinde van Ruth
NeuroCare Group Netherlands
Jeanine Jansen
Eindhovens Psychologisch Instituut

Hanneke van Dijk
Research Institute Brainclinics, Brainclinics Foundation
Martijn Arns (martijn@brainclinics.com)
Research Institute Brainclinics, Brainclinics Foundation https://orcid.org/0000-0002-0610-7613

Article

Keywords: ADHD, Biomarker, QEEG, Stratified Psychiatry

Posted Date: August 26th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-736917/v1

Abstract

Attention-deficit/hyperactivity disorder (ADHD) is characterized by neurobiological heterogeneity, possibly explaining why not all patients benefit from a given treatment. As a means to select the right treatment (stratification), biomarkers may aid in personalizing treatment prescription, thereby increasing remission rates.

The present study introduces a clinically interpretable and actionable, age- and sex-standardized biomarker based on individual alpha peak frequency (iAPF) assessed during resting-state electroencephalography (EEG). The biomarker was developed in a heterogeneous sample (N=4249), and stratifies patients with a higher iAPF to Methylphenidate (MPH; N=336) and those with a lower iAPF to Neurofeedback (NFB; N=136), resulting in a predicted gain in normalized remission of 17-30%. Blinded out-of-sample validation studies for MPH (N=58) and NFB (N=96) corroborated these findings, yielding a predicted gain in stratified normalized remission of 36% and 29%, respectively.

These findings suggest that acknowledging neurobiological heterogeneity can inform stratification of patients to their individual best treatment and enhance remission rates.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is arguably the most common neurodevelopmental disorder and is characterized by highly heterogeneous impairment profiles and etiology [1,2]. Due to this heterogeneity and differential modes of treatment action (e.g., psychostimulant vs non-stimulant medication vs non-pharmacological treatments such as multimodal neurofeedback), even the most common treatments are only effective in part of the ADHD population [3,4], with real-life remission rates of 31-57% [5]. Therefore, individualized treatment recommendation based on biomarkers that predict clinical response to specific therapeutic interventions is desirable, one example being specific activity patterns measured by electroencephalography (EEG) [6].

Ideally, treatment should be individually adapted to a given patient as envisioned in precision psychiatry. However, the multidimensionality of psychiatric disorders, in contrast to such clearly delineated problems as tumor tissue, complicates tailoring treatment to a single person [7]. An implementable intermediate step is *treatment stratification*, which aims to select a treatment from a range of effective treatments for a given disorder, informed by a biomarker.

As an example, EEG biomarker studies for treatment prediction in major depressive disorder (MDD) have shown that specific EEG patterns or abnormalities are differentially associated with *drug-specific* or *drug-class* specific antidepressant treatment effects [8–10], as well as rTMS outcome [8–10]. Such studies have also demonstrated sex differences in topographic distribution of EEG activity and yielded sex-specific predictors of MDD treatment response [9,11,12], as well as of Methylphenidate (MPH) response in ADHD [13]. Treatment stratification has already been implemented in the treatment of different cancer types [14–16]

and recently also MDD, where stratification to different antidepressant medications was informed by pre-treatment EEG biomarkers, resulting in improved remission rates relative to treatment-as-usual [17].

EEG is one of the most cost-effective and easily deployable methods to measure brain activity and is, thus, suitable for broad usage in clinical practice. Although several EEG patterns have been proposed for predicting treatment success in different mental disorders [7,18], in ADHD most biomarker studies have focused on diagnostic biomarkers, while studies investigating prognostic ADHD biomarkers are still scarce [19,20].

The individual alpha peak frequency (iAPF) is the modal frequency at which an individual's alpha activity oscillates and is known to index brain maturation [21,22]. This EEG pattern has been extensively investigated and shows promise in predicting outcome to various treatments across different disorders [10,23]. A higher mean frequency or a faster alpha peak is often associated with better cognitive performance, possibly reflective of faster information processing in thalamocortical pathways [19,20,24,25] Conversely, many mental disorders such as Alzheimer's disease, mild cognitive impairment [26], psychosis and schizophrenia [27,28] and ADHD [29] are characterized by a slowed iAPF, potentially reflective of reduced or slowed information flow between the thalamus and the cortex [19]. Furthermore, slow iAPF has been associated with worse clinical outcome to different treatments such as psychostimulants in ADHD [13,30] and most antidepressant medication in MDD [31], whereas it was found to be related to *better* clinical outcome to multimodal neurofeedback (NFB) treatment in ADHD [32] and sertraline in MDD [8].

The current study therefore investigated whether iAPF is able to differentially predict clinical outcome to two effective ADHD treatments, MPH and NFB. Given the opposite implications reported for these treatments, we hypothesized that iAPF can help subdivide a heterogeneous population into more homogeneous subpopulations with relevance to clinical outcome and thus serve as a biomarker informing treatment stratification between medications (e.g. MPH) and NFB.

Across the EEG literature, EEG (pre-) processing, EEG montages and frequency-band definitions vary considerably, which diminishes comparability and reproducibility that might at worst result in different findings [33] (see supplement S1 for more details). We therefore first initiated a Biomarker Discovery Phase, where the most precise iAPF algorithm, i.e. the algorithm yielding the most biologically plausible iAPF, was determined. This algorithm was validated against a ground truth scenario, in this case relying on the well-established finding that iAPF indexes brain-maturation [21,22]. This standardization was conducted in the large heterogeneous psychiatric patient dataset TD-BRAIN+ (N= 4249) since the goal was to explain variance in clinical data. To enhance biomarker clinical actionability and interpretability for clinicians, the resulting standardized iAPF EEG processing pipeline was used to develop an optimized iAPF-based biomarker that is both age- and sex-standardized, in line with reported sex differences in biomarkers [9,34] and effects of brain maturation on iAPF [21,22], and divided into deciles for enhanced interpretability.

The algorithm that most closely indexes brain-maturation was then ported to the Biomarker Transfer Phase, where this biomarker was utilized to find the best way to stratify patients to MPH (N=336) or NFB (N=136) according to the previously demonstrated directionality of effects [13,32]. Next, in the Biomarker Validation Phase, both the MPH and NFB prediction were submitted to blinded out-of-sample validations, by means of blinded prediction of remission on external datasets (N=58 and N=96 respectively), with accuracy verified by a third person not involved in the EEG analysis.

Finally, in an exploratory phase to test performance of the biomarker to another commonly prescribed form of pharmacotherapy for ADHD (i.e., noradrenergic medications), the predictive value of the biomarker to Atomoxetine (ATX; n=47) and Guanfacine (GUAN; n=55) was examined.

In order to maximize clinical utility of this stratification biomarker, we focused on remission as primary outcome, it representing the most clinically relevant measure [35,36], and conducted biomarker analyses separately for males and females in accordance with previous reports of sex differences [9,34].

Results

Datasets

Table 1 provides a summary of the basic demographic information of all datasets.

Full Datasets	TD-BRAIN+	iSPOT-A	NFB	MPH/GUAN *Dataset*	ICAN	ACTION
Sample size (N)	4249	336	136	141	142	56
Age range, years	6-88	6-18	6-68	7-15	7-11	6-16
Included in analysis						
Sample size (N)	4126	257	50	113	96	47
Males (%)	2528 (60)	184 (71)	38 (76)	77 (68)	76 (79)	39 (83)
Mean Age (SD), years	29.3 (18.3)	11.8 (3.2)	11.1 (3.3)	10.2 (2.1)	8.5 (1.2)	11.6 (2.6)
Treatment	NA	MPH	NFB multimodal treatment*	MPH / GUAN	NFB / multimodal treatment	ATX

Full datasets sample size reflects N of people who were enrolled. Sample size included in analysis reflects N of people with complete baseline data who finished treatment (except for TD-BRAIN+, where only baseline data but no clinical data was used). In the TD-BRAIN+ dataset the full age range was used for

age-standardization while an age range of 6-18 years was used for the correlation analyses. Sample size of this age range was 1715 (1253 male); mean age was 11.8 (SD: 3.1).

** NFB treatment augmented with advice on sleep hygiene & coaching*

MPH = Methylphenidate, NFB = Neurofeedback, GUAN = Guanfacine, ATX = Atomoxetine, SD = Standard Deviation, NA = not applicable, since no treatment effects were assessed in the discovery dataset

Biomarker Discovery Phase

Figure 1 visualizes the individual steps of the biomarker development.

In short, a total of 108 algorithm permutations were tested (Fig. 1.1.). The resulting best permutation (linked-mastoid reference/eyes closed/5s segments) was selected for further prospective testing of the biomarker (Fig. 1.2.). Age-standardized of the full TD-BRAIN+ dataset was conducted in GraphPad Prism (GraphPad Software, La Jolla California USA, www.graphpad.com) for males and females separately (Fig.1.3.). A linear regression of the resulting age-standardized values (divergence values) yielded a model with a slope of 0 (β = .000), demonstrating that the curve fitting procedure successfully removed the age effect seen before (e.g. Fz: R2= .000). For better clinical interpretability, divergence values were split into deciles.

For an overview of all correlation and secondary analyses, see supplementary material S3.

Biomarker Transfer Phase: Stratification with Brainmarker-1 results in higher likelihood of remission

In line with the analyses from the biomarker discovery phase, divergence values (i.e. the age-and sex-standardized iAPF values) were calculated for both transfer datasets. Their distribution across deciles can be found in supplementary figure S3. The primary analyses pertained to young males (6-18 yrs.) based on previous findings relating to males only [13,32] and low sample sizes in females.

Figure 2 summarizes the outcome of the transfer phase. The direction of stratification was informed by the previously reported directionality of effects (higher iAPF indicating stratification to MPH [13], lower iAPF indicating stratification to NFB [32]) and was based on the Fz electrode as primary site based on prior literature [13] (see supplement S5 for a post-hoc analysis examining stratification based on Fz and Oz). A decile cut-off point of 1-5 for NFB and 6-10 for MPH was chosen a priori, stratifying approximately half of the patients to each treatment. To test this a priori decision, positive predictive values (PPVs), indicating remission rates in the patient subsample that would have been stratified according to our biomarker were determined for different decile cut-off points. The chosen cut-off point of decile 5 indeed led to the highest combined PPV (supplementary table S1). Therefore, the presented biomarker (Brainmarker-1) was based on this cut-off point, recommending NFB treatment to boys with a relatively lower iAPF in the decile range 1-5 and MPH to boys with a relatively higher iAPF in deciles 6-10. For additional accuracy measures of predictions in all datasets, the reader is referred to supplementary table S2.

As observed remission rates strongly differed between treatments and PPVs are affected by prevalence (here: remission rate), we normalized the PPV by dividing it by the observed remission rate in each non-stratified treatment group and subtracting one. The normalized PPV indicated a predicted increase in remission rate of 17% compared to the observed remission rate if patients had received MPH (PPV= 41%) and of 30% if patients had received NFB (PPV= 62%) as treatment recommendation based on Brainmarker-1.

In a post-hoc analysis predicting remission with Brainmarker-1 calculated at the occipital site (Oz), no improvement could be seen for MPH (normalized PPV = +1.7%), however, for NFB the PPV increased to 71.4% (normalized PPV = +51% as compared to +30% in Fz). Despite this improvement for NFB treatment, Fz remained the primary stratification site, as prediction for MPH is only possible with the iAPF recorded at this location. For the results of stratification based on both Fz and Oz locations, we direct the reader to supplement S5.

'Out-of-sample' Validation Phase: Stratification biomarker predicts remission in prospective validation analysis

Next, the biomarker was validated by predicting remission to MPH and multimodal treatment including NFB (ICAN) study in two independent datasets [37,38], blinded to clinical outcomes and clinical data, based solely on the subjects' age, sex and baseline iAPF. Accuracy was verified by a third person not involved in the EEG analysis (for MPH: authors GM and SKL; and for NFB: author MA). Results are visualized in Figure 2.

In line with the previous analyses, we normalized PPVs to improve comparability with the transfer datasets. The normalized PPV predicted an increase in remission rate of 36% (PPV = 50%) compared to the observed remission rate if patients had received MPH and of 29% (PPV= 29%) if patients had received the multimodal treatment based on Brainmarker-1.

Biomarker exploration phase

In a last step, we explored the predictive potential of Brainmarker-1 for ATX and GUAN treatment. When testing different decile ranges for ATX, a cut-off point of £ decile 6 resulted in the highest normalized PPV of +27% (PPV= 40%). This seems to point to a similar directionality of effect as was observed for NFB treatment, while using the cut-off point that was also used for MPH (deciles ³ 6) results in a decline in remission rate (improvement: -8%). However, when the same decision process as for NFB was applied, i.e. predicting remission to ATX in individuals with decile scores £5, the resulting improvement was marginal (PPV = 33%, improvement= +6%).

For GUAN treatment, a prediction of remission for deciles 6-10, the same that was used for MPH prediction, resulted in the highest PPV (53%) and normalized PPV (+26%).

Discussion

In the present study, an iAPF algorithm indexing brain maturation was developed in the Biomarker Discovery Phase in a large clinical sample (a subset of this EEG database as well as the EEG processing code is freely available for download at https://brainclinics.com/resources/). Subsequently, this iAPF was employed to develop an iAPF-based, age- and sex-standardized treatment stratification biomarker (Brainmarker-1), which was found to be capable of differentially informing stratification to MPH and NFB treatment. The results from the Biomarker Transfer Phase indicate that a neurobiologically heterogeneous sample of ADHD patients can be successfully divided into two more homogeneous sub-samples characterized by a relatively faster or slower iAPF and a differential response to MPH and NFB.

Given both MPH and multimodal treatment that includes NFB can be considered effective interventions for the treatment of ADHD, with remission rates between 31-51% [3,39], employing EEG to stratify to one of these treatments effectively increases predicted remission rates in the stratified group by 17-30% compared to non-stratified remission rates. Crucially, the Biomarker Validation Phase substantiated Brainmarker-1 through a blinded out-of-sample prediction of remission in two external datasets, based solely on age, sex and baseline iAPF. Since Brainmarker-1 is based only on basic demographic data and the resting-state EEG, it can easily be implemented in clinical practice, using an algorithm which calculates age-and sex-standardized iAPF into deciles and yields a treatment recommendation.

Our results suggest that we can successfully stratify between MPH and NFB treatments, while ATX and GUAN findings are promising but require future replication. Most importantly, the directionality of iAPF and its association with remission to MPH/GUAN is opposite that of NFB/ATX. This is imperative for the concept of treatment stratification, as its aim is to use a biomarker to inform the best treatment option for each patient choosing from a range of effective treatments for that disorder, instead of merely discouraging a particular intervention.

This differential association of iAPF with remission in response to different treatments might be related to the branches of the autonomous nervous system (ANS). ADHD has been associated with hypoarousal of the ANS or a hyperactivity of the parasympathetic nervous system (PNS) [40,41] which is supported by the finding that heart rate (HR) is generally lower in children with ADHD, suggestive of higher vagal tone [30]. However, there have also been studies that found an elevated sympathetic nervous system (SNS) response [34,42] or a hyperactivation of both PNS and SNS [43], pointing to a general ANS imbalance. Similarly, iAPF has been hypothesized to index fight or flight response, with the iAPF acutely speeding-up in the presence of an acute threat, such as pain [44], or slowing down with chronic stress such as chronic pain [45,46] or burnout syndrome [47], possibly reflecting a thalamocortical gating mechanism, counter-regulating the surplus of pain- or stress-induced innervation [44,45]. Moreover, it has been shown that people with PTSD, a disorder characterized by an overactive SNS, have a generally faster iAPF [48]. A slower iAPF could thus point to a hyperactive PNS while a faster iAPF could reflect relatively normal PNS or increased SNS activation.

While MPH acts on noradrenaline to some extent, its main working mechanism seems to be an increase of synaptic dopamine by inhibiting dopamine re-uptake through inhibition of the dopamine transporter

(DAT). It might, thus, be possible that the mechanism of action of MPH is relatively unrelated to ANS imbalances and instead brings about its effect by acting on a number of different neurotransmitters simultaneously [49]. This is in line with a recent meta-analysis that reports null effects of ANS imbalances in ADHD as the most common finding [40], suggesting a more diverse pathophysiology that goes beyond ANS abnormalities. A possible explanation for the relationship between MPH and iAPF comes from a study in healthy subjects, investigating the relationship between iAPF and a functional polymorphism in the gene encoding catechol-O-methyltransferase (COMT), which plays a role in dopamine metabolism [50]. They found that Valine (Val) allele homozygotes who show increased COMT activity and reduced dopamine in the prefrontal cortex (PFC) had a significantly lower iAPF compared to subjects with the methionine allele who have normal PFC dopamine signaling. This could indicate that genetically induced PFC dopamine receptor functioning might play a role in MPH treatment response. However, a later study could not replicate this association [51].

Since noradrenaline is the major neurotransmitter in the SNS, ATX, a selective noradrenaline reuptake inhibitor (SNRI) might act by normalizing PNS hyperactivity in people with a slower iAPF. Interestingly, the effect for GUAN seems to be opposite that of ATX although it also acts on noradrenaline, however as an alpha2a adrenergic receptor agonist. Similar differential biomarker findings for *within-drug* classes, alluding to drug-specific effects, have been reported before. For example, the selective serotonin reuptake inhibitors Escitalopram and Sertraline for treatment of depression that are thought to primarily act on the serotonergic system, demonstrated differential effects with regard to iAPF and treatment outcome, with only sertraline responders showing a lower iAPF. Our biomarker findings thus, suggest that there might be relevant functional differences between ATX and GUAN, requiring further study.

The precise working mechanism of NFB is unknown at present. However, speculatively, it has been hypothesized that NFB might affect sleep-regulating mechanisms [52–54]. Since ADHD has been associated with increased daytime sleepiness [55] and sleepiness is correlated with increased parasympathetic activity [56], NFB might work by improving sleep and thereby normalizing parasympathetic activity. On the other hand, Pimenta and colleagues recently emphasized the multimodal nature of this treatment [5], also evident from the absence of group effects in the double-bind placebo controlled ICAN study [38] that was used here in the validation phase. Long-term effects of up-to one year follow-up in the ICAN study demonstrated clinical benefits – on the group level – similar to the MPH arm of the NIMH-MTA study [38]. This further suggests that the multimodal approach including frequent reinforcement as well as sleep hygiene coaching are important factors, pointing again to the possibility that the relationship between slow iAPF and remission to NFB could be driven by sleep regulation.

While we demonstrated the prognostic value of Brainmarker-1 in two independent and blinded out-of-sample validations, the present study also had some limitations. Brainmarker-1 presently only pertains to

males and ages 6-18 years. The reason for this is limited sample size for females in the treatment studies and clear qualitative sex-specific effects [13], as well as a lack of adults for most of the datasets, which prevented us from investigating stratification for these groups (see supplement S4 for a discussion of results for females). Findings in females might be particularly important since they are usually underrepresented in ADHD research [57]. Likewise, investigating treatment stratification in adults with ADHD would be valuable.

Since the present study examined multiple treatment datasets from different test locations with different designs, rating scales, methods, and EEG methodology, testing was not standardized. However, the fact that the out-of-sample validation was successful demonstrates the strength of the developed biomarker in spite of those differences.

Moreover, the transfer NFB sample received NFB treatment augmented with sleep hygiene management and coaching while the NFB validation dataset received a multimodal NFB or control treatment and sleep and nutrition counselling. Findings might, therefore, not be directly comparable to standard NFB monotherapy [32].

While this study already successfully validated MPH and multimodal NFB prediction by means of Brainmarker-1, a prospective validation study would be valuable that prospectively stratifies patients between the interventions based on baseline iAPF, similar to the feasibility study of van der Vinne [17] that successfully tested the biomarker-based stratification approach for MDD patients. Since the relationship between iAPF and MDD treatment outcome has already been established [8,58,59], a next step will involve expanding Brainmarker-1 that was applied to ADHD samples in the current study into a transdiagnostic marker, incorporating different pharmacological and non-pharmacological interventions for MDD.

In summary, the present study introduces a clinically implementable and interpretable treatment stratification biomarker for young males with ADHD and corroborates it in two blinded out-of-sample validations.

Methods

Datasets – Biomarker Discovery Phase

The large TD-BRAIN+ dataset, comprising patients with various psychiatric disorders was utilized to determine the optimal parameters of iAPF calculation. Full details of the open access TD-BRAIN dataset (N=1274) have been published in van Dijk et al. (under review; www.brainclinics.com/resources) with all data recorded at Research Institute Brainclinics (Brainclinics Foundation, Nijmegen). In the TD-BRAIN+ dataset this was complemented with data from additional clinics (EPI-PIT clinics (Eindhoven & Tilburg; author JJ), EEG resource (Nijmegen; author RB), Neuroscan (Dordrecht; author PdJ), neuroCare clinics (Hengelo; Groningen; Munich; Sydney, author RvR)), while EEG caps, amplifiers, instructions and other details were identical to van Dijk et al. (under review).

Datasets - Biomarker Transfer Phase

For iAPF prospective analyses we relied on the following datasets (see Table 1 for overview): Methylphenidate (iSPOT-A: N=257 [13]) and Neurofeedback (N=50) [32].

Datasets - Biomarker Validation Phase

For independent out-of-sample replication analysis, we predicted remission in the MPH/GUAN dataset [37] (Table 1) and the International Collaborative ADHD Neurofeedback (ICAN) study [38].

In the former trial, subjects were blindly randomized to either MPH (N=58) or GUAN (N=55) treatment, with dosing adjusted based on weight. In the ICAN study (N=96), subjects were blindly randomized to a multimodal treatment of sleep and nutrition counselling and either theta/beta ratio NFB or a control NFB treatment (NFB administered based on the pre-recorded EEG).

Datasets - Biomarker Exploration Phase

Explorative analyses were conducted in Atomoxetine (ACTION [60]) and in GUAN from the MPH/GUAN dataset used in the validation phase for MPH replication [37].

All participants (or their parents or care-takers) gave written informed consent prior to testing.

Data availability

The TD-BRAIN EEG data is freely available for download at https://brainclinics.com/resources/. Other data is available from the corresponding author on reasonable request.

Code availability

The Python code used for processing the EEG and calculating the iAPF is freely available for download at https://brainclinics.com/resources/.

EEG data collection and preprocessing

All EEGs were recorded in a standardized manner as developed by Brain Resource Ltd. (for more details see [9]) apart from the independent MPH/GUAN validation dataset [37].

In short, EEGs were recorded from 26 channels according to the 10-20 electrode international system (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz, O2; Quikcap, NuAmps). Measurements consisted of 2-minute Eyes Open (EO) and 2-minute Eyes Closed (EC) recordings. During EO recordings, participants were asked to fixate a dot in the middle of the computer screen.

Data was recorded with the ground at AFz, and a sampling rate of 500 Hz and a low-pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization. Horizontal eye-

movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3 mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was <10 kW for all electrodes.

Automatic artifact detection and removal were performed using a custom-built Python package [61–64] and were in accordance with deartifacting as described in [9] and van Dijk et al. (under review), with full code available online (www.brainclinics.com/resources).

For the MPH/GUAN validation dataset [37], eyes-closed EEGs were recorded from 40 channels (AF3, AF4, AFz, C3, C4, CPz, Cz, F10, F3, F4, F7, F8, F9, FCz, FP1, FP2, FPz, FT10, FT7, FT8, FT9, Fz, Iz, O1, O2, Oz, P10, P3, P4, P7, P8, P9, POz, Pz, T7, T8, TP10, TP7, TP8, TP9) for 5 minutes with a sampling rate of 256 Hz and referenced to linked ears (for further details, see [37,65]). Recordings were subsequently matched to our data, i.e., the 40 channels were reduced to 22 channels matching TD-BRAIN+ set-up (with FC3, FC4, CP3 and CP4 missing). Artifact rejection for the independent validation dataset was performed in BrainVision Analyzer Version 2.2.0 (Brain Products GmbH, Gilching, Germany) by semi-automatic removal of epochs with signal amplitudes >150mV.

iAPF determination

The individual alpha peak frequency was determined by computing the FFT of the preprocessed, artefact-free data. Frequency resolution varied depending on the segmentation of data (e.g., 0.2 Hz for 5s segments). Subsequently each individual's iAPF was determined by identifying the highest peak within the frequency range of 7 to 13 Hz.

Biomarker Discovery Phase

Biomarker discovery a priori focused on males and females separately due to previously reported qualitative sex differences [9,34].

First, datasets with LVA were identified and excluded from further analysis since in cases of absent alpha, no reliable peak can be determined which would decrease the signal-to-noise ratio for iAPF and weaken treatment prediction on the group-level. In short, subjects whose alpha power fell below a z-score of -1.96 of the log-transformed average spectral power distribution were discarded (for more details, see supplement S2).

In order to optimize EEG processing, iAPFs determined with different processing parameters were correlated with age based on the well-known notion that iAPF indexes brain-maturation, thus validating against the biologically most plausible alpha peak that will explain most of the variance (i.e., the highest correlations with age). An upper age threshold of 18 years was chosen a priori, i.e., the age at which iAPF is assumed to plateau, based on early literature [21] and more recent work that showed the iAPF maturation effect in a sample aged 6 to 18 years [13]. Parameters tested were: a) EEG segment lengths from 2-7s, b) reference to an AR or a LM montage and c) site (locations Fz, Pz, Oz). All combinations of these

parameters were applied to determine iAPFs from the EC and EO recordings separately, as well as from the difference between EC and EO power spectra.

A decision on segment length was made based on 1) the strength of the correlation and 2) the number of subjects retained for each segment length and averaged across reference (LM and AR), and conditions (EC, EO, Diff) for all 3 electrode locations (Fz, Pz, Oz) separately. The choice of reference montage was based on the highest iAPF age correlation for the age range of interest, i.e., subjects below the age of 18.

Subsequently, the iAPF-age effect was eliminated and resulting values divided into 10 equal-sized bins (deciles) to improve interpretability.

In order to validate the use of a clinical instead of a normative dataset, the full curve fitting procedure in GraphPad prism, specified above, was repeated in a normative dataset [66]. Subsequently, in a comparison of fit both the normative and the clinical curve fit were applied to both the normative and clinical data separately and the fit was compared.

A secondary analysis, comparing the curve fit of the clinical TD-BRAIN+ dataset with the curve fit specific to a normative dataset [66] in GraphPad prism, indicated that the parameters of the clinical dataset generalized significantly better (p=0.03) to the normative data than the other way around (p=.21), suggesting that the clinical data is capable of explaining higher variance.

Biomarker transfer phase

We first aimed to align previous findings which differed with regard to primary outcome measure (response vs remission) and subsample (boys aged 6-18 vs boys aged 12-18) [13,32]. To increase comparability and clinical impact, we focused our analyses on males in the age range of 6-18 years and on remission – defined as an item mean of ≤1.00 on the ADHD-RS-IV - as primary clinical outcome [35].

Biomarker validation phase

Finally, Brainmarker-1 was prospectively validated for MPH and multimodal NFB treatment by a blinded prediction of remission status, solely based on age, sex and baseline EEG in two independent dataset.

Biomarker exploration phase

Analyses for the exploration phase were similar to those in the transfer phase but without a guided hypothesis. The focus of analysis was also an age range of 6-18 years and remission as clinical outcome.

Statistics

First, Spearman correlations between the various iAPFs resulting from different EEG processing combinations (see supplement S3 and supplementary figure S2) and age in subjects below 18 years (N=1671) were calculated. To determine standardized iAPF values independent of age, we derived non-

linear regression models based on the full TD-BRAIN+ dataset that most closely fit the given data for each electrode (Fz, Pz, Oz). Different mathematical models following the developmental trajectory of the iAPF (such as a Log gaussian model, in line with [67]) were contrasted against a linear model (null hypothesis) and individually adjusted for females and males and for each site (channel). Divergence values representing where the individual's iAPF lies in relation to other people's iAPFs, were calculated from the resulting models by subtracting the model-derived average iAPF for each subject's age from the person's actual iAPF. Correlations between divergence values and age were conducted to confirm that the age effect had been eliminated from the data. The resulting divergence values were ranked from low to high and divided into 10 equal-sized bins (deciles) to increase interpretability by clinicians.

The final stratification outcome for the transfer phase and stratification decision for the exploration phase were based on the positive predictive values (PPVs) at different decile cut-off points, indicating remission rate within the subsample of patients that the Brainmarker-1 would have stratified to the respective treatment. Since PPVs are dependent on prevalence (here: observed remission) and remission rates differed between treatment datasets, we normalized PPVs for better comparability across datasets by dividing the PPV by the observed remission and subtracting 1.

Curve fitting models were developed in GraphPad Prism version 8.4.0 for MacOS. Spearman correlations were conducted with Python modules scipy, and numpy.

All other statistical analyses were performed in IBM SPSS Statistics for Macintosh, Version 27.0.

Declarations

Disclosures

KRG and MSK have received research funding from Takeda Pharmaceutical Company (Japan) and Shire (Australia) for work unrelated to that presented in this manuscript.

EG is founder and receives income as Chief Executive Officer and Chairman for Brain Resource Ltd. He has stock options in Brain Resource Ltd. RD has received research funding from NIMH, has served on the Board of Directors for the International Society for Neurofeedback and Research, and has a clinic in NC where he performs neurofeedback among other clinical services

DP has received income and stock options with the role of science and data processing manager as an employee with Brain Resource Ltd. RB is owner of EEG resource, a neurofeedback/psychology practice. LEA has received research funding from Curemark, Forest, Lilly, Neuropharm, Novartis, Noven, Otsuka, Roche/Genentech, Shire, Supernus, and YoungLiving (as well as NIH and Autism Speaks), has consulted with CHADD, Neuropharm, Organon, Pfizer, Sigma Tau, Shire, Tris Pharma, and Waypoint, and been on advisory boards for Arbor, Ironshore, Novartis, Noven, Otsuka, Pfizer, Roche, Seaside Therapeutics, Sigma Tau, Shire. RvR holds stock in neuroCare Group AG. MA is unpaid chairman of the non-profit Brainclinics Foundation, a minority shareholder in neuroCare Group (Munich, Germany), and a co-inventor on 4 patent

applications related to EEG, neuromodulation and psychophysiology, but receives no royalties related to these patents; Research Institute Brainclinics received research and consultancy funding from neuroCare Group (Munich, Germany), Neuroscience Software (UK), Urgotech (France) and equipment support from Deymed, neuroConn and Magventure.

HV, GvW, GM, SKL, DD, PdJ, JJ, and HvD have nothing to disclose.

Acknowledgements

We would like to thank Mark Koppenberg for the design and editing of figures and tables.

Author contributions

HV was responsible for data collection, analyzed the data, and wrote the manuscript. GvW provided supervision, and reviewed the manuscript. GM participated in analyzing the data, and reviewed the manuscript. KRG, EG, RD, MSK, SKL, DP, RB, DD, LEA, PdJ, RvR, JJ were involved in collecting the data and reviewing the manuscript. HvD was involved in designing the study and analyzing the data, participated in writing the manuscript, and supervised all phases of the study. MA designed the study, was involved in data collection, participated in analyzing the data and writing the manuscript, and supervised all phases of the study.

Materials & Correspondence

Correspondence and material requests should be addressed to the corresponding author, MA.

References

1. Luo, Y., Weibman, D., Halperin, J. M. & Li, X. A Review of Heterogeneity in Attention Deficit/Hyperactivity Disorder (ADHD). *Front Hum Neurosci* **13**, 42 (2019).

2. Banaschewski, T. *et al.* Attention-Deficit/Hyperactivity Disorder: A Current Overview. *Deutsches Aerzteblatt Online* **114**, 149–159 (2017).

3. Cortese, S. *et al.* Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis. *Lancet Psychiatry* **5**, 727–738 (2018).

4. Molina, B. S. G. *et al.* The MTA at 8 Years. *J Am Acad Child Adolesc Psychiatry* **48**, 484–500 (2009).

5. Pimenta, M. G., Brown, T., Arns, M. & Enriquez-Geppert, S. Treatment Efficacy and Clinical Effectiveness of EEG Neurofeedback as a Personalized and Multimodal Treatment in ADHD: A Critical Review. *Neuropsych Dis Treat* **Volume 17**, 637–648 (2021).

6. Atkinson, A. *et al.* Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clin Pharmacol Ther* **69**, 89–95 (2001).

7. Olbrich, S., Dinteren, R. van & Arns, M. Personalized Medicine: Review and Perspectives of Promising Baseline EEG Biomarkers in Major Depressive Disorder and Attention Deficit Hyperactivity Disorder. *Neuropsychobiology* **72**, 229–240 (2016).

8. Arns, M., Gordon, E. & Boutros, N. N. EEG Abnormalities Are Associated With Poorer Depressive Symptom Outcomes With Escitalopram and Venlafaxine-XR, but Not Sertraline: Results From the Multicenter Randomized iSPOT-D Study. *Clin EEG Neurosci* **48**, 33–40 (2017).

9. Arns, M. *et al.* EEG alpha asymmetry as a gender-specific predictor of outcome to acute treatment with different antidepressant medications in the randomized iSPOT-D study. *Clin Neurophysiol* **127**, 509–19 (2016).

10. Olbrich, S. & Arns, M. EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response. *Int Rev Psychiatr* **25**, 604–618 (2013).

11. Iseger, T. A. *et al.* EEG connectivity between the subgenual anterior cingulate and prefrontal cortices in response to antidepressant medication. *Eur Neuropsychopharmacol* **27**, 301–312 (2017).

12. Dinteren, R. van *et al.* Utility of event-related potentials in predicting antidepressant treatment response: an iSPOT-D report. *Eur Neuropsychopharmacol* **25**, 1981–1990 (2015).

13. Arns, M. *et al.* Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *Eur Neuropsychopharm* **28**, 881–891 (2018).

14. Orr, K. E. & McHugh, K. The new international neuroblastoma response criteria. *Pediatr Radiol* **49**, 1433–1440 (2019).

15. Deng, K., Li, H. & Guan, Y. Treatment Stratification of Patients with Metastatic Castration-Resistant Prostate Cancer by Machine Learning. *Iscience* **23**, 100804 (2020).

16. Kato, M. & Manabe, A. Treatment and biology of pediatric acute lymphoblastic leukemia. *Pediatr Int* **60**, 4–12 (2018).

17. Vinne, N. van der *et al.* EEG biomarker informed prescription of antidepressants in MDD: a feasibility trial. *Eur Neuropsychopharm* (2021) doi:10.1016/j.euroneuro.2020.12.005.

18. Keizer, A. W. Standardization and Personalized Medicine Using Quantitative EEG in Clinical Settings. *Clin Eeg Neurosci* 1550059419874945 (2019) doi:10.1177/1550059419874945.

19. Clark, C. R. *et al.* Spontaneous alpha peak frequency predicts working memory performance across the age span. *Int J Psychophysiol* **53**, 1–9 (2004).

20. Pahor, A. & Jaušovec, N. Making Brains run Faster: are they Becoming Smarter? *Span J Psychology* **19**, E88 (2016).

21. Lindsley, D. B. BRAIN POTENTIALS IN CHILDREN AND ADULTS. *Science* **84**, 354–354 (1936).

22. Smith, J. R. The Electroencephalogram During Normal Infancy and Childhood: II. The Nature of the Growth of the Alpha Waves. *Pedagogical Seminary J Genetic Psychology* **53**, 455–469 (1938).

23. Arns, M. EEG-Based Personalized Medicine in ADHD: Individual Alpha Peak Frequency as an Endophenotype Associated with Nonresponse. *J Neurother* **16**, 123–141 (2012).

24. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Rev* **29**, 169–195 (1999).

25. Grandy, T. H. *et al.* Individual alpha peak frequency is related to latent factors of general cognitive abilities. *Neuroimage* **79**, 10–18 (2013).

26. Rodriguez, G., Copello, F., Vitali, P., Perego, G. & Nobili, F. EEG spectral profile to stage Alzheimer's disease. *Clin Neurophysiol* **110**, 1831–1837 (1999).

27. Murphy, M. & Öngür, D. Decreased peak alpha frequency and impaired visual evoked potentials in first episode psychosis. *Neuroimage Clin* **22**, 101693 (2019).

28. Yeum, T.-S. & Kang, U. G. Reduction in Alpha Peak Frequency and Coherence on Quantitative Electroencephalography in Patients with Schizophrenia. *J Korean Med Sci* **33**, (2018).

29. Bazanova, O. M., Auer, T. & Sapina, E. A. On the Efficiency of Individualized Theta/Beta Ratio Neurofeedback Combined with Forehead EMG Training in ADHD Children. *Front Hum Neurosci* **12**, 3 (2018).

30. Arns, M., Gunkelman, J., Breteler, M. & Spronk, D. EEG Phenotypes predict treatment outcome to stimulants in children with ADHD. *J Integr Neurosci* **07**, 421–438 (2008).

31. Ulrich, G., Renfordt, E., Zeller, G. & Frick, K. Interrelation between Changes in the EEG and Psychopathology under Pharmacotherapy for Endogenous Depression. A contribution to the predictor question. *Pharmacopsychiatry* **17**, 178–183 (1984).

32. Krepel, N. *et al.* A multicenter effectiveness trial of QEEG-informed neurofeedback in ADHD: Replication and treatment prediction. *Neuroimage Clin* 102399 (2020) doi:10.1016/j.nicl.2020.102399.

33. Yao, D. *et al.* Which Reference Should We Use for EEG and ERP practice? *Brain Topogr* **32**, 530–549 (2019).

34. Hermens, D. F., Kohn, M. R., Clarke, S. D., Gordon, E. & Williams, L. M. Sex differences in adolescent ADHD: findings from concurrent EEG and EDA. *Clin Neurophysiol* (2005)

doi:10.1016/j.clinph.2005.02.012.

35. Steele, M., Jensen, P. S. & Quinn, D. M. P. Remission versus response as the goal of therapy in ADHD: A new standard for the field? *Clin Ther* **28**, 1892–1908 (2006).

36. SWANSON, J. M. *et al.* Clinical Relevance of the Primary Findings of the MTA: Success Rates Based on Severity of ADHD and ODD Symptoms at the End of Treatment. *J Am Acad Child Adolesc Psychiatry* **40**, 168–179 (2001).

37. Loo, S. K. *et al.* Effects of d-Methylphenidate, Guanfacine, and Their Combination on Electroencephalogram Resting State Spectral Power in Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry* **55**, 674-682.e1 (2016).

38. Group, T. N. C. *et al.* Double-blind placebo-controlled randomized clinical trial of neurofeedback for attention-deficit/hyperactivity disorder with 13 month follow-up. *J Am Acad Child Adolesc Psychiatry* (2020) doi:10.1016/j.jaac.2020.07.906.

39. Arns, M. *et al.* Neurofeedback and Attention-Deficit/Hyperactivity-Disorder (ADHD) in Children: Rating the Evidence and Proposed Guidelines. *Appl Psychophys Biof* **45**, 39–48 (2020).

40. Bellato, A., Arora, I., Hollis, C. & Groom, M. J. Is autonomic nervous system function atypical in attention deficit hyperactivity disorder (ADHD)? A systematic review of the evidence. *Neurosci Biobehav Rev* **108**, 182–206 (2020).

41. Musser, E. D. *et al.* Emotion Regulation via the Autonomic Nervous System in Children with Attention-Deficit/Hyperactivity Disorder (ADHD). *J Abnorm Child Psych* **39**, 841–852 (2011).

42. Leikauf, J. E. *et al.* Identification of biotypes in Attention-Deficit/Hyperactivity Disorder, a report from a randomized, controlled trial. *Personalized Medicine Psychiatry* **3**, 8–17 (2017).

43. Morris, S. S. J. *et al.* Emotion Regulation via the Autonomic Nervous System in Children with Attention-Deficit/Hyperactivity Disorder (ADHD): Replication and Extension. *J Abnorm Child Psych* **48**, 361–373 (2020).

44. Nir, R.-R., Sinai, A., Raz, E., Sprecher, E. & Yarnitsky, D. Pain assessment by continuous EEG: Association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest. *Brain Res* **1344**, 77–86 (2010).

45. Boord, P. *et al.* Electroencephalographic slowing and reduced reactivity in neuropathic pain following spinal cord injury. *Spinal Cord* **46**, 118–123 (2008).

46. Sarnthein, J., Stern, J., Aufenberg, C., Rousson, V. & Jeanmonod, D. Increased EEG power and slowed dominant frequency in patients with neurogenic pain. *Brain* **129**, 55–64 (2006).

47. Luijtelaar, G. van, Verbraak, M., Bunt, M. van den, Keijsers, G. & Arns, M. EEG Findings in Burnout Patients. *J Neuropsychiatry Clin Neurosci* **22**, 208–217 (2010).

48. Wahbeh, H. & Oken, B. S. Peak High-Frequency HRV and Peak Alpha Frequency Higher in PTSD. *Appl Psychophys Biof* **38**, 57–69 (2013).

49. Challman, T. D. & Lipsky, J. J. Methylphenidate: Its Pharmacology and Uses. *Mayo Clin Proc* **75**, 711–721 (2000).

50. Bodenmann, S. *et al.* The Functional Val158Met Polymorphism of COMT Predicts Interindividual Differences in Brain α Oscillations in Young Men. *J Neurosci* **29**, 10855–10862 (2009).

51. Veth, C. P. M. *et al.* Association between COMT Val158Met genotype and EEG alpha peak frequency tested in two independent cohorts. *Psychiat Res* **219**, 221–4 (2014).

52. Sterman, M. B., Howe, R. C. & Macdonald, L. R. Facilitation of Spindle-Burst Sleep by Conditioning of Electroencephalographic Activity While Awake. *Science* **167**, 1146–1148 (1970).

53. Arns, M., Feddema, I. & Kenemans, J. L. Differential effects of theta/beta and SMR neurofeedback in ADHD on sleep onset latency. *Front Hum Neurosci* **8**, 1019 (2014).

54. Arns, M. & Kenemans, J. L. Neurofeedback in ADHD and insomnia: Vigilance stabilization through sleep spindles and circadian networks. *Neurosci Biobehav Rev* **44**, 183–194 (2014).

55. Golan, N., Shahar, E., Ravid, S. & Pillar, G. Sleep Disorders and Daytime Sleepiness in Children with Attention-Deficit/ Hyperactive Disorder. *Sleep* **27**, 261–266 (2004).

56. Pressman, M. R. & Fry, J. M. Relationship of Autonomic Nervous System Activity to Daytime Sleepiness and Prior Sleep. *Sleep* **12**, 239–245 (1989).

57. Bedard, K. & Witman, A. Family structure and the gender gap in ADHD. *Rev Econ Household* **18**, 1101–1129 (2020).

58. Corlier, J. *et al.* The relationship between individual alpha peak frequency and clinical outcome with repetitive Transcranial Magnetic Stimulation (rTMS) treatment of Major Depressive Disorder (MDD). *Brain Stimul* (2019) doi:10.1016/j.brs.2019.07.018.

59. Roelofs, C. *et al.* Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the ICON-DB consortium. *Clin Neurophysiol* (2020) doi:10.1016/j.clinph.2020.10.017.

60. Griffiths, K. R. *et al.* Response inhibition and emotional cognition improved by atomoxetine in children and adolescents with ADHD: The ACTION randomized controlled trial. *J Psychiatr Res* **102**, 57–64 (2018).

61. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput Sci Eng* **9**, 90–95 (2007).

62. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272 (2020).

63. Harris, C. R., Millman, K. J., Walt, S. J. van der & al., et. Array programming with NumPy. *Nature* **585**, 357–362 (2020).

64. team, T. pandas development. *pandas-dev/pandas*. (Zenodo, 2020).

65. McCracken, J. T. *et al.* Combined Stimulant and Guanfacine Administration in Attention-Deficit/Hyperactivity Disorder: A Controlled, Comparative Study. *J Am Acad Child Adolesc Psychiatry* **55**, 657-666.e1 (2016).

66. Gerrits, B. *et al.* Probing the “Default Network Interference Hypothesis” With EEG: An RDoC Approach Focused on Attention. *Clin EEG Neurosci* **50**, 404–412 (2019).

67. Dinteren, R. van, Arns, M., Jongsma, M. L. A. & Kessels, R. P. C. Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. *Front Aging Neurosci* **6**, 294 (2014).

Figures

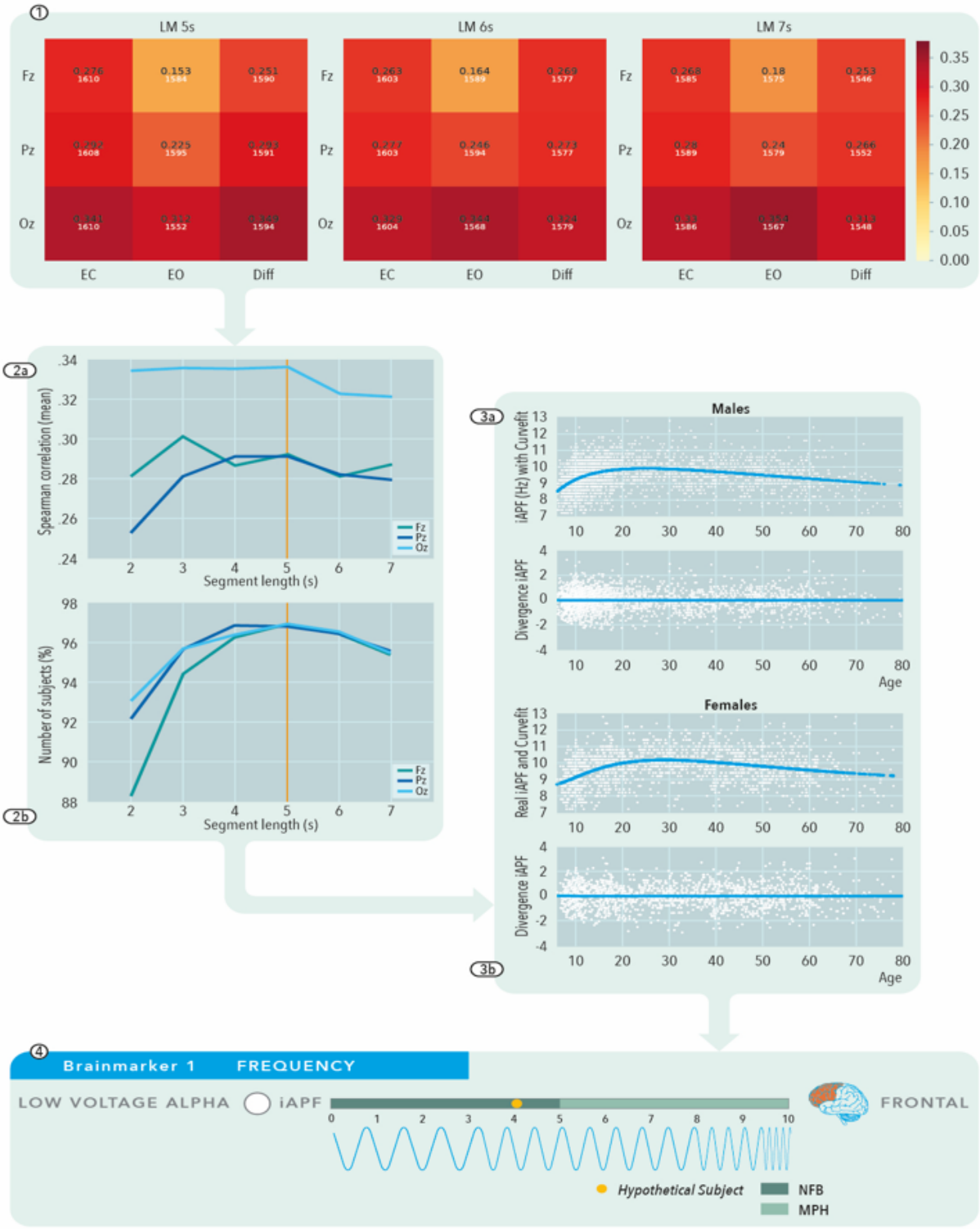

Figure 1

Biomarker Discovery Phase. (1) Excerpt from heatmaps of the total of 108 algorithm permutations (27 depicted) that were tested and selected based on the highest correlation between age and iAPF in subjects <18 years (Spearman Correlation ρ; black digits) and the highest retention of data (number of subjects N; white digits). (2) Spearman Correlation ρ between age (6-18 yrs.) and iAPF (2a) and number of subjects (2b) for each electrode and segment length (2-7s) for condition EC averaged across reference

montages (N=1715). (3) Flattening the iAPF-age curve for males (3a) and females (3b) separately at electrode location Oz. Upper subplots depict non-standardized iAPFs and the optimized Log Gaussian model fit. Lower subplots depict the age-standardized divergence values and a linear fit through the data. (4) Example of the derived biomarker (Brainmarker-1) based on the final age- and sex-standardized scores, with deciles 1-5 yielding a recommendation for NFB treatment and deciles 6-10 yielding a recommendation for MPH.

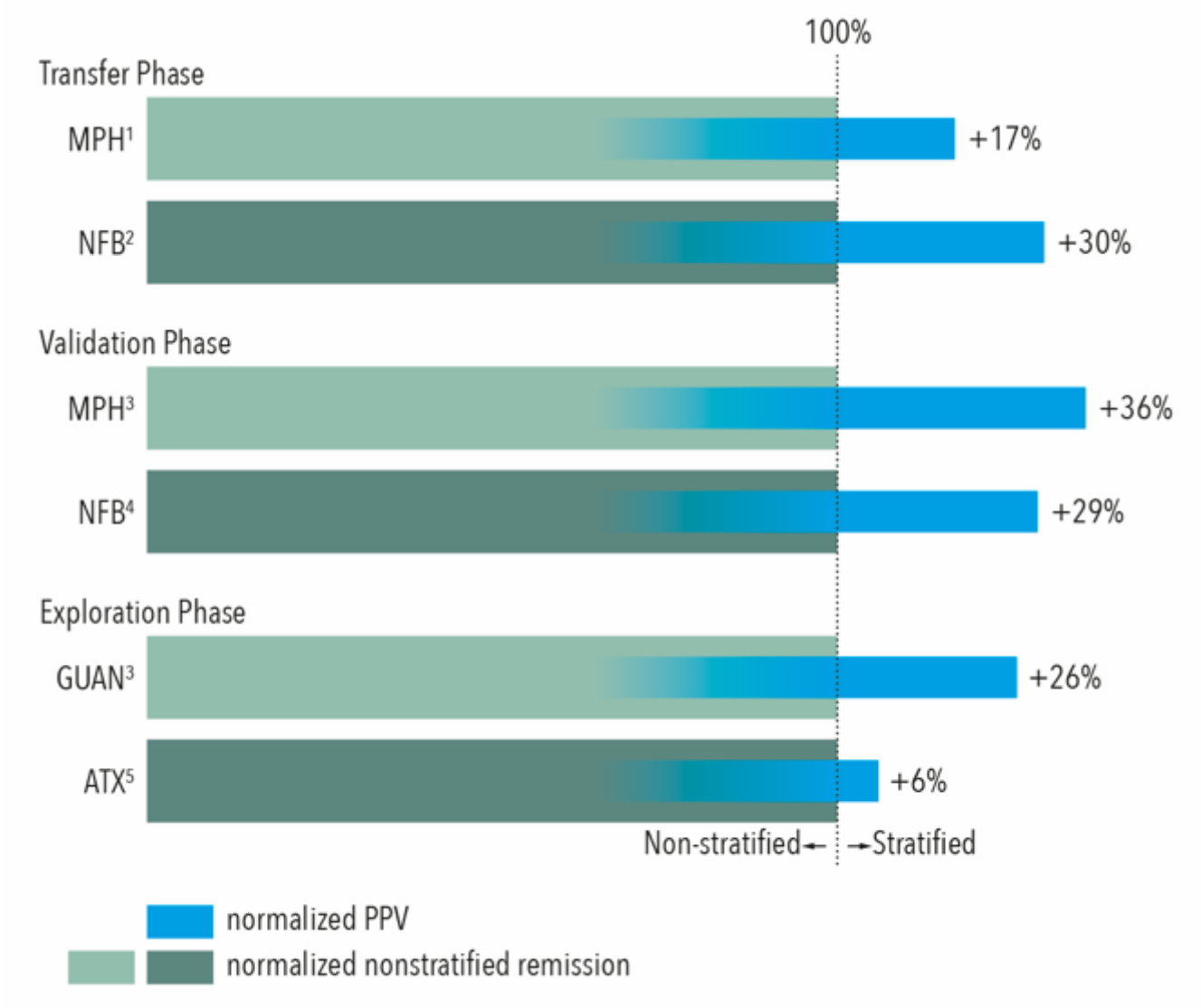

Figure 2

Predicted remission rate after stratification. Normalized PPVs (in blue) for each treatment group depict predicted gain in remission if patients had been stratified according to Brainmarker-1. Light vs. darker green implicates opposite direction for Brainmarker-1, i.e. light green indicates decile 6-10 (e.g. MPH) and dark green deciles 1-5 (e.g. NFB). Note that the predicted remission in the blinded validation is highest. 1 iSPOT-A (N=257), 2 NFB dataset (N=50), 3 MPH/GUAN dataset (MPH: N=58, GUAN: N= 55), 4 ICAN (N=96), 5 ACTION (N=47) MPH= Methylphenidate, NFB = Neurofeedback, ATX=Atomoxetine, GUAN = Guanfacine

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymaterialBiomarkerbasedtreatmentstratificationinADHDAnoutofsamplevalidation.docx