

ARTICLE OPEN



ROC Analysis of Biomarker Combinations in Fragile X Syndrome-Specific Clinical Trials: Evaluating Treatment Efficacy via Exploratory Biomarkers

Jordan E. Norris^{1,2}, Elizabeth M. Berry-Kravis², Mark D. Harnett³, Scott A. Reines³, Melody R. Smith⁴, Emma K. Auger¹, Abigail H. Outterson², Jeremiah Furman², Mark E. Gurney³ and Lauren E. Ethridge^{1,5}✉

© The Author(s) 2025

Fragile X Syndrome (FXS) is a rare neurodevelopmental disorder caused by a trinucleotide repeat expansion on the 5' untranslated region of the *FMR1* gene. FXS is characterized by intellectual disability, anxiety, sensory hypersensitivity, and difficulties with executive function. A recent phase 2 placebo-controlled clinical trial assessing BPN14770, a first-in-class phosphodiesterase 4D allosteric inhibitor, in 30 adult males (age 18–41 years) with FXS demonstrated cognitive improvements on the NIH Toolbox Cognitive Battery in domains related to language and caregiver reports of improvement in both daily functioning and language. However, individual physiological measures from electroencephalography (EEG) demonstrated only marginal significance for trial efficacy. A secondary analysis of resting state EEG data collected as part of the phase 2 clinical trial evaluating BPN14770 was conducted using a machine learning classification algorithm to classify trial conditions (i.e., baseline, drug, placebo) via linear EEG variable combinations. The algorithm identified a composite of peak alpha frequencies (PAF) across multiple brain regions as a potential biomarker demonstrating BPN14770 efficacy. Increased PAF from baseline was associated with drug but not placebo. Given the relationship between PAF and cognitive function among typically developed adults and those with intellectual disability, as well as previously reported reductions in alpha frequency and power in FXS, PAF represents a potential physiological measure of BPN14770 efficacy.

Translational Psychiatry (2025)15:323; <https://doi.org/10.1038/s41398-025-03558-2>

INTRODUCTION

Fragile X Syndrome (FXS) is a rare X-linked, monogenic neurodevelopmental disorder caused by a trinucleotide repeat expansion of ≥ 200 CGG repeats in the 5' untranslated region of the Fragile-X messenger ribonucleoprotein 1 (*FMR1*) gene resulting in gene methylation and a subsequent full or partial reduction in Fragile X messenger ribonucleoprotein (FMRP) output [1, 2]. The relationship between molecular disruptions and externally measurable cognitive and behavioral symptoms for FXS, and whether these are mediated by abnormal physiology, is unknown [3]. Identifying the relationship between molecular mechanisms and clinical measures indexing cognition is important for developing targeted treatments that address cognitive disruptions, including impaired verbal and nonverbal intelligence, language processing and production difficulties, crystalized cognition issues, and cognitive inflexibility [4, 5]. Cognitive symptoms and difficulties with learning are among the most distressing symptoms for individuals with FXS and their families as intellectual disability (ID) limits independence in activities of daily living [6–9]. Current behavioral and pharmaceutical interventions aim to address symptoms of FXS but without specifically

targeting the underlying pathophysiology of FXS, or the ID/cognitive deficits [10, 11].

Recent findings from a successful phase 2 clinical trial assessing the novel therapeutic BPN14770 (Zatolmilast) demonstrated good safety and tolerability with improvements on measures of cognition and daily functioning, showing promise as the first pharmaceutical intervention targeting the underlying pathophysiology of cognition in FXS [11]. BPN14770 is a first-in-class phosphodiesterase-4D (PDE4D) allosteric inhibitor, which is specific for the dimeric, PKA-activated form of PDE4D that acts as a key modulator of cAMP levels relevant to important cognitive functions, such as learning and memory (Fig. 1). Reduction of FMRP in FXS affects typical cAMP metabolism such that the decreases in cAMP are often observed in FXS [12–14]. While the phase 2 trial showed a cognitive benefit for BPN14770, understanding of the impact of BPN14770 on EEG-derived physiological measures remains incomplete.

To determine the impact of BPN14770 on potential EEG biomarkers associated with clinical change, EEG was obtained in the phase 2 clinical trial. Specifically, N1 event-related potential (ERP) amplitude and habituation to repeated auditory stimulation

¹Department of Psychology, University of Oklahoma, Norman, OK, USA. ²Department of Pediatrics, Neurological Sciences, and Biochemistry, Rush University Medical Center, Chicago, IL, USA. ³Tetra Therapeutics, A Shionogi Company, Grand Rapids, MI, USA. ⁴Department of Anesthesiology, Perioperative and Pain Medicine, Stanford University, Stanford, CA, USA. ⁵Department of Pediatrics, Section on Developmental and Behavioral Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA. ✉email: ethri@ou.edu

Received: 16 March 2024 Revised: 21 July 2025 Accepted: 18 August 2025

Published online: 28 August 2025

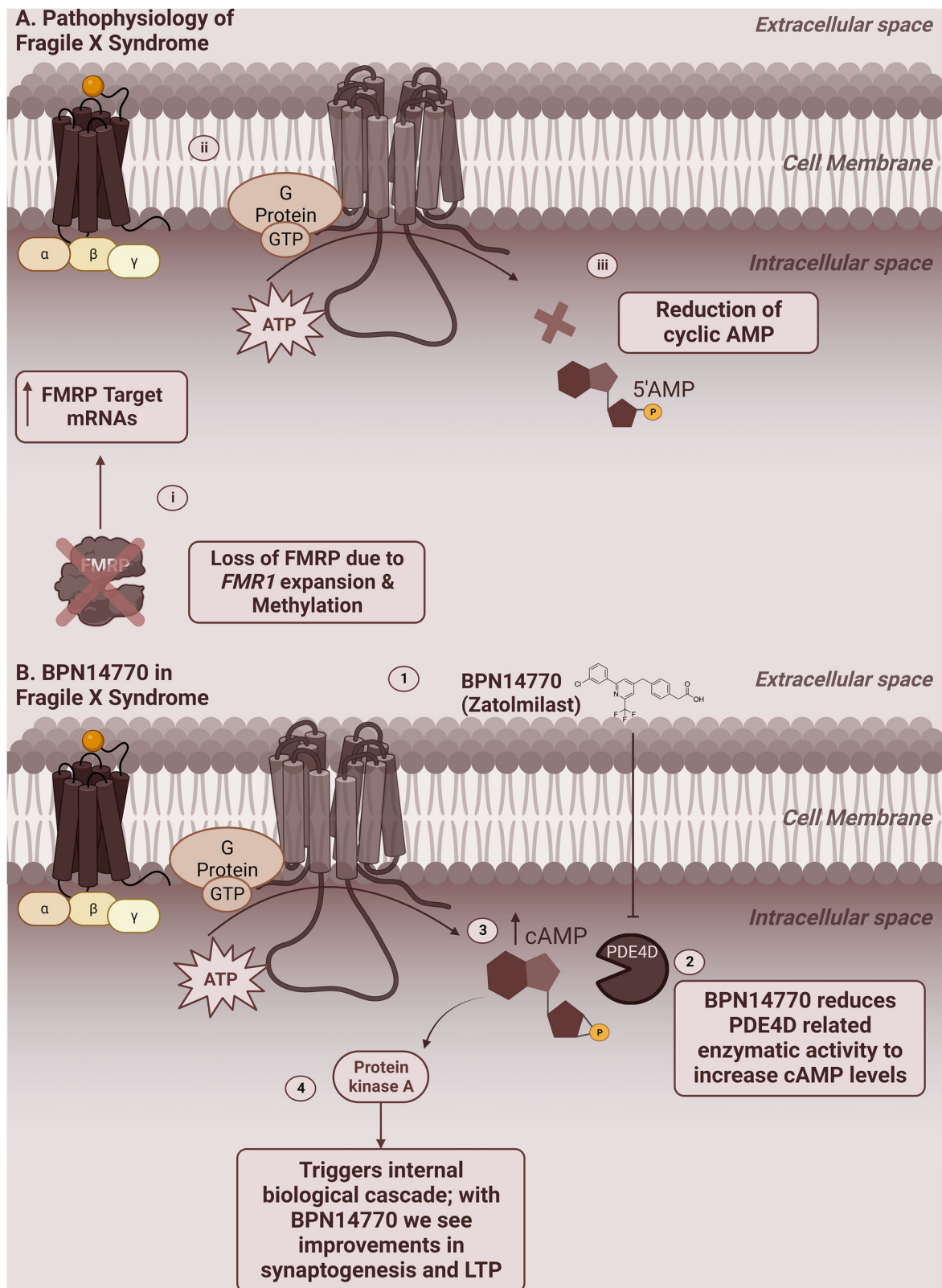


Fig. 1 Pathophysiology of FXS and its Relation to the Mechanism of Action for BPN14770. **A)** The pathophysiology of FXS: i-iv.) Reduced FMRP results in aberrant patterns of protein translation, internalization of AMPA receptors and reduced levels of cAMP. within neurons **B)** The mechanistic action of BPN14770: 1-4 BPN14770 inhibits activity of phosphodiesterase 4D or PDE4D leading to increased cAMP circulation within neurons resulting in improvements in synaptogenesis and LTP. Made with BioRender.

were selected a priori for assessment based on previous findings of increased N1 amplitudes and decreased N1 habituation in FXS [15]. Differences between treatment and placebo were marginally significant ($p = .06$), and N1 amplitude was correlated with serum levels of drug suggesting some reduction in abnormally elevated N1 amplitude with BPN14770 [11, 16]. However, comparisons were underpowered due to data loss in the ERP, which is a longer and thus relatively more difficult-to-collect measure in FXS. Exploratory resting state EEG (rsEEG) was also collected to evaluate improvements in frequency bands most affected in FXS. rsEEG is a rich source of information about intrinsic neural activity and analyses of rsEEG data are generally more robust against data loss compared to the ERP. In previous rsEEG studies, individuals with FXS typically exhibit increased power in gamma (30 – 90 Hz) and theta (4 – 7 Hz) with notable decreases in alpha (8 – 13 Hz) bands [15, 17, 18], as well as reduced dynamic utilization of alpha oscillations, and a downshifted peak alpha frequency (PAF) which is known to correlate with cognitive performance in typically developed individuals [17–23]. Re-examination of EEG with a focus on rsEEG data may reveal underlying shifts associated with BPN14770 in biomarkers that index FXS pathophysiology.

Taking a more data-driven approach to evaluating the rsEEG may help reduce sample size concerns when exploring physiological correlates of BPN14770 efficacy. Neural biomarker development via EEG and translation of biomarkers to clinical-based interventions in FXS research supports efforts to identify novel therapeutic interventions specific to cognitive outcomes in FXS. However, data loss in populations with FXS presents a challenge to determining therapeutic efficacy via decreases in statistical power, particularly in phase 1 and 2 studies which typically recruit minimally necessary samples to establish safety and tolerability. Despite a focus on establishing safety and tolerability in phase 2 studies, determination of efficacy and target engagement is still required to move into phase 3. Efficacy challenges faced by clinical trials can often reflect methodological difficulties rather than a definitive lack of target engagement. Approximately 60% of clinical trials involving novel therapeutics that fail in clinical development, fail due to inadequate evidence for efficacy where inability to demonstrate efficacy may occur due to misspecification of the best endpoint, making secondary analysis an important follow-up step [20]. Given the demonstration of efficacy in cognitive measures in the current phase 2 clinical trial assessing BPN14770, concomitant physiological changes were likely present but not as robust in EEG due to a mismatch in the chosen endpoint (N1 amplitude), statistical underpowering due to data loss in the ERP task, or by focus on singular rather than composite EEG measures [11]. Composite measures, which integrate patterns

across multiple variables, may provide optimal protection from chance fluctuation in single variables, thus increasing signal-to-noise allowing better capture of physiological change in small samples [21]. While biomarkers are excellent tools for improving diagnostic specificity, measuring therapeutic efficacy, and mechanistically understanding biological processes, utilizing single biomarkers may not fully characterize physiological dynamics relevant to detection of effects in the smaller samples typical of phase 2 clinical trials. Recent work in FXS and other populations has demonstrated that combining biomarkers to index nuanced physiological patterns may allow for more robust measurement at the individual level that will enhance relationships to clinically relevant outcomes (i.e., changes in observable behaviors) and be robust against consequences of data loss for individual measures [24–27]. However, introducing multiple combinations of measures for statistical comparison can lead to multiple comparisons concerns and increase chance of Type I error, necessitating some data reduction strategy to select the most likely variable combinations in advance this step. Machine learning classifiers, most frequently used to identify and separate diagnostic groups via receiver operating curve (ROC) analyses [28], may also be utilized to identify and separate treatment conditions, and can serve as an important screening step for composite biomarker identification.

The current study constitutes a secondary analysis of rsEEG using data driven methods to explore overarching shifts in neural physiology by utilizing a naïve Bayes Classifier to determine whether a priori linear variable combinations can separate participants along trial conditions (i.e., pre-dose baseline, placebo, and BPN14770) for the phase 2 trial of BPN14770 in FXS. By evaluating the feasibility of utilizing machine learning to “pre-screen” composite variable biomarkers in this clinical trial with limited sample size, we aim to further address target engagement and physiological effects of BPN14770 on neural processes of interest in FXS.

METHODS

Study Design

Participants were 30 males (full sample: age 18–41 years, $M = 31.63$, $SD = 7.32$; IQ 24.63–66.19, $M = 42.78$, $SD = 11.16$) with FXS participating in a single-center, phase 2a clinical trial assessing the efficacy and safety of BPN14770. The clinical trial was randomized, double-blinded, placebo-controlled, and utilized a two-period cross-over design without a washout period (Fig. 2; ClinicalTrials.gov identifier: NCT03569631, see [11] for full inclusion and exclusion criteria). All participants or their legal guardians signed informed consent which included consent for EEG data collection. The clinical site was Rush University Medical Center (RUMC) where all study documents, including study protocol, consent documents, recruitment

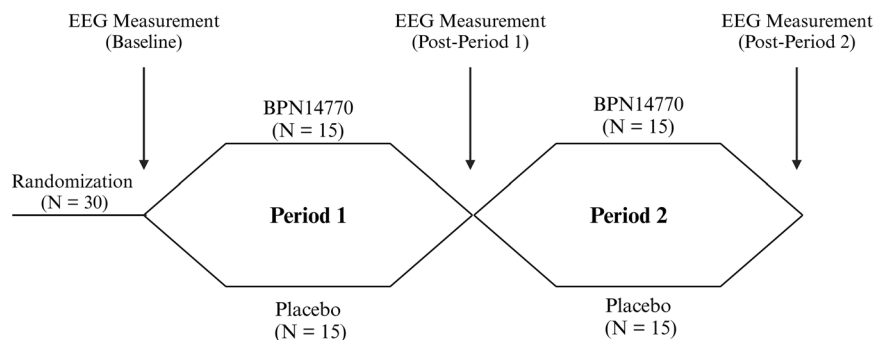


Fig. 2 Flow Diagram for Clinical Trial Design with Indication of EEG Measurement Timepoints. The phase 2a trial evaluating BPN14770 (Zatolmilast) utilized a cross-over design and enrolled 30 individuals with FXS. The 30 were then randomized into two groups and the first (baseline) EEG measurement was taken during screening. During period 1, half received BPN14770 and the other group received placebo. After period 1, another EEG was taken and was either the drug condition or placebo condition EEG measure depending on group. During period 2, the groups switched, and they received whichever intervention conditions they had not received in period 1. After period 2, the final EEG measurement was taken. Made with BioRender.

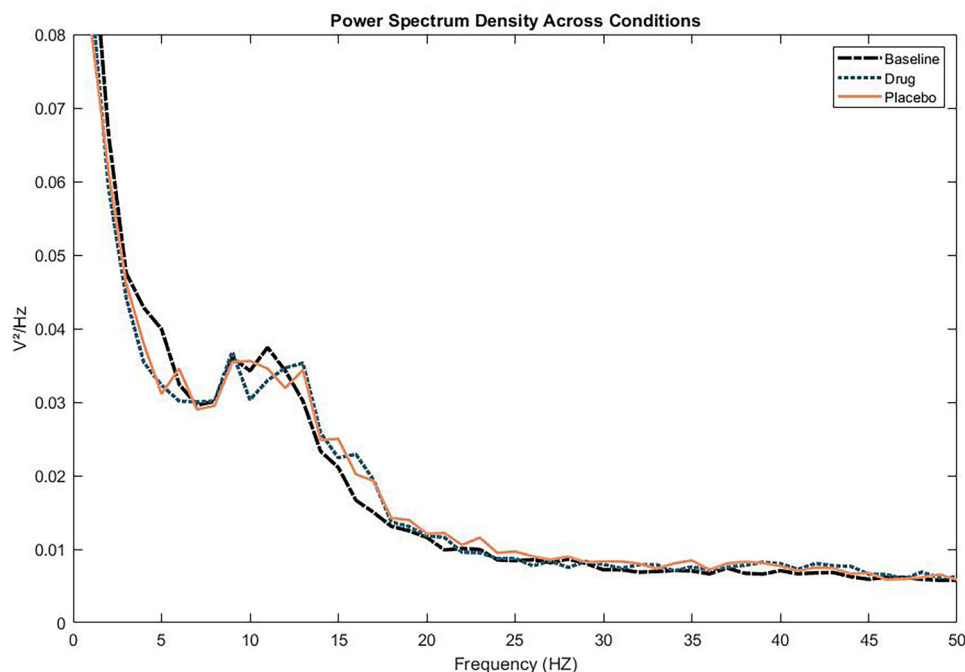


Fig. 3 Relative Power Spectrum Density Plot Across Conditions. PSD for all conditions generated and used to compute PAF values.

materials, safety information for participants, and information about study compensation were approved by the RUMC institutional review board (Approval #00000482) [11]. Participant recruitment was managed by RUMC and supplemented by FXS patient advocacy groups. All EEG procedures were approved by the University of Oklahoma institutional review board (Approval #10129). All study methods were conducted according to the ethical guidelines of the Declaration of Helsinki, seventh revision, 64th World Medical Association General Assembly Meeting and are consistent with the International Conference on Harmonization/good clinical practice, applicable regulatory requirements and the sponsor or its delegate's policy on bioethics.

The current methods cover a secondary analysis of rEEG data collected to explore biomarker outcomes for assessing target engagement. There was some initial EEG data loss in period 2 due to the study being completed partly during the Covid19 pandemic. The baseline sample consisted of 23 individuals. Of the 23, 17 provided quality EEG data across all conditions for both periods (i.e., BPN14470 and placebo). Two additional individuals were missing a drug or placebo measurement where multiple imputation was used to impute the relationship between drug and baseline for statistical comparison only (all conditions sample: $N = 19$ ages 21–41 years; $M = 32.95$, $SD = 6.58$; $IQ\ 24.63$ – 63.41 , $M = 41.20$, $SD = 10.37$). Finally, the period 1 sample consisted of 23 who provided quality placebo or BPN14770 EEG data for period 1 and were used to evaluate BPN14770 against placebo as a between subjects' comparison ($N = 23$, age 21–41 years, $M = 32.74$, $SD = 6.57$; $IQ\ 24.63$ – 63.41 , $M = 41.46$, $SD = 9.44$).

EEG Recording and Preprocessing

EEG data were continuously recorded and digitized at 512 Hz, with a 5th order Bessel anti-aliasing filter at 200 Hz, using a 32-channel BioSemi ActiveTwo system (BioSemi). All sensors were referenced to a Common Mode Sense-Driven Right Leg active reference loop which replaces traditional ground electrodes and actively corrects for electrical noise during recording. Data were inspected offline and preprocessed to remove artifacts prior to analysis. No more than ~ 5% of sensors were interpolated (interpolation limited to a max of 2 channels); data were digitally filtered offline from 0.5 to 100 Hz with a 57–63 Hz notch, resampled to a 500 Hz sampling rate, submitted to independent components analysis (ICA) via EEGLAB for artifact removal with segments of data containing large movement-related artifacts that would negatively impact the ICA decomposition removed prior to ICA [29], and re-referenced to the average of all channels (see supplemental Fig. 1). Final data lengths were all greater than 20 s for analysis (average data length in seconds: baseline: $M = 92.04$, $SD = 39.56$; BPN14770: $M = 85.83$, $SD = 41.79$; Placebo:

$M = 91.51$, $SD = 35.24$). A repeated measures ANOVA was run to assess differences in data length across conditions ($N = 19$) with no main effect of data length, $F(2,17) = .83$, $p = .45$, $ES = .09$.

Frequency Bands

Continuous absolute and relative power were calculated across all electrodes for each frequency band where spectral power density was divided into 7 bands: delta (2 – 3.5 Hz), theta (3.5 – 7.5 Hz), alpha 1 (8 – 10 Hz), alpha 2 (10 – 12.5 Hz), beta (13 – 30 Hz), gamma 1 (30 – 55 Hz), and gamma 2 (65 – 90 Hz) [17]. Power was calculated using the first 80 seconds of data to standardize across participants; if 80 seconds of data was not available the maximum amount of data available was used.

Peak Alpha

Peak alpha frequency (PAF) was included as a biomarker of cognitive function. Frequency bins (0.5 Hz) were formed to create a power density spectrum (PDS) calculated from absolute power. The absolute power spectrogram was converted to a relative power spectrogram and log transformed to find the maximum peak between 6 and 14 Hz for each electrode from a given participant using methods consistent with previous work in FXS [17] (Fig. 3; see supplemental Fig. 2).

Alpha bursts

Neurodynamic metrics in both the alpha and gamma frequency bands were calculated. Methods for calculating alpha burst metrics reflect those proposed by Allen and Cohen (2010) and used by our group previously [20, 30]. Data were bandpass filtered in the alpha range (8 – 13 Hz) and a Hilbert transform was applied. The transformed signal was used to calculate instantaneous power at each timepoint using the natural log of the squared absolute value of the complex result. Burst metrics were calculated by selecting one a priori electrode from each hemisphere and subtracting left hemisphere data (i.e., F3/O1) from the right hemisphere data (i.e., F4/O2) to provide continuous asymmetry power values - [power(Right) – power(Left)]. Alpha bursts were then calculated by defining a threshold as the upper 80th percentile of the continuous asymmetry power value for each participant where burst counts totaled the number of times the threshold value was crossed per second and lengths reflected the average time spent in each burst (i.e., continuous time spent above threshold). Alpha burst metrics were calculated from both frontal (i.e., F3/F4) and occipital (i.e., O1/O2) regions and gamma bursts were calculated only from the frontal region to avoid areas generating increased muscle artifact. Gamma burst metrics were

calculated similarly to alpha burst metrics with the threshold set at the 90th percentile to take a more conservative approach [20, 30].

Biomarker Combinations

Predictors were hypothesis driven combinations of EEG measures previously reported to be affected in FXS. All biomarkers were EEG frequency measures evaluated across multiple metrics using different a priori electrode combinations. Metrics were assessed in the frontal region (i.e., electrodes FP1, FP2, AF3, AF4, F3, F4, F7, F7), occipital region (i.e., electrodes PO3, PO4, O1, O2), and across the whole head using all 32 electrodes. Broadly, predictors were pulled from three categories: 1.) frequency bands: individual frequency bands from combined frontal/occipital regions, specific frequency band combinations (e.g., alpha and theta), and region-specific assessments of all frequency bands across brain regions (frontal, occipital, and whole head), 2.) peak alpha across the same brain regions, and 3.) burst metrics.

Biomarker Combinations: Frequency Bands

Relative and absolute power were computed across each frequency band and assessed from the combination of frontal and occipital regions due to targeted interest in frequency utilization in frontal and occipital regions. Combinations of specific frequency bands of interest were also assessed, including frontal theta-alpha, frontal theta-gamma, occipital theta-alpha, and theta-alpha-gamma combinations. Lastly, all frequency bands calculated from both relative and absolute power (i.e., delta (2-3.5 Hz), theta (3.5-7.5 Hz), alpha 1 (8-10 Hz), alpha 2 (10-12.5 Hz), beta (13-30 Hz), gamma 1 (30-55 Hz), and gamma 2 (65-90 Hz)) were evaluated across frontal and occipital regions to assess region-specific frequency band effects, and then evaluated across the whole head.

Biomarker Combinations: Peak Alpha

PAF predictor variable combinations were assessed by region (i.e., frontal, occipital, whole head, all three regions combined, and frontal/occipital combined). The frontal/occipital combination served as a more targeted assessment of region-specific differences in alpha frequency utilization, as differences in PAF have been reported between these regions in FXS [17], whereas whole head assessments provide a more parsimonious but less dynamic PAF measure.

Biomarker Combinations: Burst Metrics

Burst metrics were included as an exploratory predictor set to explore whether BPN14770 improved dynamic alpha and gamma utilization. Alpha and gamma bursts were assessed in frontal (F3 and F4) and occipital (O1 and O2) regions with predictors reflecting length of time spent or number of times per second (CPS) individuals entered a dynamic alpha/gamma state. Specifically, predictors were: 1.) combination of all variables (F3/F4 CPS and lengths, O1/O2 CPS and lengths), 2.) lengths and CPS at specific electrodes (F3, F4, O1, and O2), or 3.) individual CPS or lengths within respective electrode sets (F3/F4, O1/O2). Gamma bursts were assessed for the frontal region only (F3 and F4) with predictors generated similarly to alpha burst metric predictors due to high muscle artifact from the neck region.

Machine Learning Algorithm

A multiclass naïve Bayes classifier (NBC, see Supplemental Fig. 3 for methodological details), selected for robustness against small sample sizes [31, 32] was used to classify participants into BPN14770, placebo, or baseline conditions based on hypothesis-driven variable combinations using MATLAB R2020b (The Mathworks, Natick, MA, United States). Code is available upon request to corresponding author. The NBC produces simple linear outputs that are easily translatable to clinical threshold values and calculation of composite scores. Importantly, the NBC performed on all three study conditions (baseline, placebo, drug) assumes that placebo is separable from baseline and thus that a placebo effect of some measurable magnitude occurs. We first evaluated all three study conditions on the 17 individuals that had complete usable data to evaluate physiological features robust against placebo effects and potential carryover effects (persistent pharmaceutical effects into the placebo window). Additional classifications were made following the same procedure due to persistent BPN14770 carryover effects in period 2 for those in the placebo condition: the NBC was additionally used to 1.) classify participants into placebo or BPN14770 for period 1 only, and 2.) baseline or

BPN14770 across both period 1 and 2. The naïve Bayes model was trained on predictors using 70% of the data and tested on a holdout sample of 30%. Cross validated classification error (CE) was calculated for each combination using kfoldLoss to assess errors in correctly classifying the whole dataset. Due to the limited sample size, the process was bootstrapped over 2000 iterations to produce an overall average CE, area under the curve (AUC) for the ROC, true positive rate, and false positive rate for assessing model performance. Statistical evaluations were limited to a subset of variable combinations that performed best per AUC and CE values.

Clinical Measures

Clinical trial secondary outcome measures assessed in the current study included the 1) National Institutes of Health-Toolbox Cognition Battery (NIH-TCB) which included 5 subscales (i.e., Cognition Crystallized Composite (CCC), Picture Vocabulary (PV), Oral Reading Recognition (ORR), Picture Sequence Memory (PSM), and Pattern Comparison Processing Speed PCPS) [33], 2) a Visual Analog Scale constructed using patient-specific behavioral anchors selected by the parent/caregiver to assess 3 domains (language, anxiety/irritability, and daily function) [11], 3) Aberrant Behavioral Checklist [34], and 4) Anxiety, Depression, and Mood Scale [35]. The original study found significant improvements with BPN14770 in NIH-TCB CCC, PV, ORR, and both VAS daily functioning and language domains. Clinical variables were selected because they either 1) demonstrated BPN14770 effects or, 2) are related to processes mediated by PAF.

Statistics

Difference scores (i.e., either baseline or placebo effects subtracted from BPN14770 effects) were calculated to match NBC methods for generating probabilities for statistical evaluation of BPN14770 effects. Difference scores were evaluated using one-sided one-sample t-tests based on expected performance in favor of BPN14770. One-sided tests were assessed because statistics were only applied to best-performing variables. Repeated measure ANOVAs assessed differences on best performing EEG variable combinations and clinical variables across conditions with Fisher's LSD to assess for condition differences when significant main effects were present. Effect sizes are reported as partial eta squared. Linear regressions were used to assess causal relationships between best performing variables/variable combinations and clinical variables that previously detected BPN14770 effects [11]. We examined exploratory correlations between best performing EEG variables and clinical variables for both BPN14770, placebo, and baseline using Spearman's rho. Baseline correlations were evaluated to establish relationships between processes captured by clinical measures and EEG to support validity of any clinical correlations with EEG for the BPN14770 condition. Further, baseline correlations were an exploratory effort to better establish relationships between EEG measures and clinical features of FXS. For period 1 only, independent samples t-tests were conducted to assess differences between placebo and BPN14770. Effect sizes for all t-test are reported as Cohen's d.

RESULTS

Naïve Bayes Classifier Performance

150 hypothesis-driven variable combinations per comparison (i.e., all conditions, period 1, and baseline vs. drug) were evaluated by the NBC, for a total of 450 total combinations. PAF was identified as the best performing variable category with the combination of all variables outperforming single regions or other region combinations determined by which variable had the highest AUCs and lowest CE (Table 1). PAF was evaluated statistically for neural effects in favor of BPN14770 across all conditions (N = 19) and in period 1 only (N = 23) due to known carryover effects.

Statistical Evaluation of NBC Best Performers

All Conditions (N = 19): The best performing variable/variable combination was all PAF variables combined (Fig. 4A & B). The average difference score for all variables combined assessing the BPN14770 effect minus pre-dose baseline was significantly different from 0, $t(18) = 3.53$, $p = .001$, $d = .81$. Across all conditions individually, BPN14770 effect minus pre-dose baseline

Table 1. Area Under the Curve (AUC) and Classification Errors for all Variable Combinations.

	AUCs and Classification Error					
	All Conditions (N = 17)		BPN14770 vs. Baseline (N = 23)		BPN14770 vs. Placebo (Period 1; N = 23)	
	AUC	CE	AUC	CE	AUC	CE
<u>ABSOLUTE POWER</u>						
Frequency Band Regions						
Frontal	0.628	0.59	0.616	0.55	0.615	0.57
Occipital	0.593	0.72	0.687	0.58	0.563	0.65
Whole Head	0.659	0.67	0.671	0.49	0.547	0.51
All Variables	0.673	0.64	0.692	0.51	0.673	0.57
Frequency Bands						
Delta	0.589	0.66	0.588	0.51	0.487	0.57
Theta	0.611	0.69	0.662	0.46	0.502	0.57
Alpha 1	0.597	0.67	0.556	0.47	0.614	0.61
Alpha 2	0.624	0.62	0.580	0.63	0.630	0.55
Beta	0.633	0.63	0.581	0.57	0.531	0.72
Gamma 1	0.603	0.63	0.618	0.59	0.567	0.65
Gamma 2	0.591	0.63	0.627	0.57	0.608	0.61
Frequency Band Combinations						
Theta-Alpha-Gamma (Frontal)	0.629	0.62	0.624	0.56	0.561	0.58
Theta-Alpha (Frontal)	0.614	0.61	0.628	0.48	0.548	0.58
Theta-Gamma (Frontal)	0.642	0.60	0.581	0.56	0.575	0.58
Theta-Alpha (Occipital)	0.584	0.67	0.675	0.49	0.681	0.54
<u>RELATIVE POWER</u>						
Frequency Band Regions						
Frontal	0.623	0.68	0.616	0.55	0.643	0.63
Occipital	0.640	0.71	0.687	0.58	0.644	0.75
Whole Head	0.583	0.74	0.669	0.49	0.704	0.66
All Variables	0.633	0.72	0.691	0.50	0.685	0.71
Frequency Bands						
Delta	0.557	0.66	0.590	0.51	0.579	0.51
Theta	0.611	0.63	0.662	0.46	0.554	0.57
Alpha 1	0.599	0.72	0.557	0.47	0.585	0.58
Alpha 2	0.646	0.75	0.580	0.63	0.607	0.61
Beta	0.595	0.66	0.581	0.58	0.665	0.54
Gamma 1	0.569	0.73	0.617	0.59	0.521	0.82
Gamma 2	0.576	0.70	0.627	0.57	0.556	0.72
Frequency Band Combinations						
Theta-Alpha-Gamma (Frontal)	0.639	0.67	0.624	0.56	0.613	0.64
Theta-Alpha (Frontal)	0.630	0.70	0.630	0.48	0.630	0.56
Theta-Gamma (Frontal)	0.585	0.66	0.581	0.56	0.537	0.74
Theta-Alpha (Occipital)	0.659	0.69	0.675	0.49	0.540	0.72
<u>FREQUENCY BURSTS</u>						
Frontal Alpha Bursts						
CPS	0.564	0.71	0.495	0.61	0.541	0.41
Length	0.601	0.65	0.518	0.48	0.707	0.49
F3	0.568	0.61	0.555	0.52	0.642	0.41
F4	0.603	0.66	0.610	0.46	0.689	0.52
All Variables	0.598	0.67	0.607	0.54	0.677	0.44
Occipital Alpha Bursts						
CPS	0.634	0.59	0.617	0.44	0.629	0.51
Length	0.618	0.49	0.629	0.45	0.547	0.62

Table 1. continued

	AUCs and Classification Error					
	All Conditions (N = 17)		BPN14770 vs. Baseline (N = 23)		BPN14770 vs. Placebo (Period 1; N = 23)	
	AUC	CE	AUC	CE	AUC	CE
O1	0.620	0.51	0.564	0.49	0.566	0.59
O2	0.640	0.62	0.546	0.50	0.630	0.52
All Variables	0.640	0.48	0.551	0.48	0.627	0.56
Frontal Gamma Bursts						
CPS	0.606	0.73	0.536	0.49	0.514	0.61
Lengths	0.646	0.72	0.504	0.45	0.526	0.57
F3	0.616	0.73	0.625	0.56	0.549	0.57
F4	0.637	0.66	0.699	0.42	0.504	0.59
All Variables	0.629	0.74	0.684	0.46	0.523	0.53
PEAK ALPHA						
Peak Alpha Frequency						
Frontal	0.581	0.58	0.691	0.36	0.615	0.49
Occipital	0.569	0.49	0.701	0.40	0.563	0.61
Whole Head	0.617	0.58	0.738	0.35	0.547	0.58
Frontal-Occipital	0.613	0.69	0.743	0.37	0.667	0.49
All Variables	0.638	0.61	0.756	0.37	0.673	0.47

Note. Table columns are paired by the area under the curve (AUC) value for trial condition comparison and the classification error (CE). ML performance depicted (in order) include AUCs for all conditions (i.e., BPN14770, baseline, and placebo), AUCs for the BPN14770 – baseline prediction, finally followed by AUCs for the BPN14770 – placebo prediction.

PAF from frontal, occipital, and whole head regions were all significantly different from 0 (frontal: $t(18) = 2.34$, $p = .031$, $d = .54$; occipital: $t(18) = 2.62$, $p = .017$, $d = .60$; whole head: $t(18) = 4.01$, $p = .001$, $d = .92$). The average difference score (i.e., all head regions) for the BPN14770 effect minus placebo was not significantly different from 0, $t(18) = 1.29$, $p = .211$, $d = .30$. BPN14770 effect minus placebo was not different from 0 across all regions for PAF (frontal: $t(18) = 1.57$, $p = .134$, $d = .36$; occipital: $t(18) = .38$, $p = .708$, $d = .09$; whole head: $t(18) = 1.52$, $p = .145$, $d = .34$).

A repeated measure ANOVA was conducted on PAF from all variables combined created by averaging PAF across all regions (i.e., frontal, occipital, and whole head) to assess for differences across conditions (Fig. 4). There was a main effect of condition on PAF where BPN14770 increased PAF ($M = 8.02$, $SE = 0.20$) compared to baseline ($M = 7.36$, $SE = 0.13$) but not placebo ($M = 7.71$, $SE = 0.17$), $F(2, 17) = 7.59$, $p = .004$, $ES = .47$. Placebo represented an intermediate value, likely due to carryover effects from period 1, and was also significantly increased from baseline based on a post hoc comparison ($p = .044$).

The clinical variables were re-evaluated to determine whether significant clinical effects were detectable with reduced measurement time-points [i.e., 3 instead of the 5 reported in Berry-Kravis et al. (2021)] and reported in Supplemental Tables 1 and 2. Correlations between PAF region difference scores (BPN14770-baseline) and difference scores for clinical variables of interest were assessed and found no significant relationships (Supplemental Table 3). Additionally, exploratory correlations were assessed for the BPN14770 condition only for PAF raw values (Supplemental Table 4) and found a significant correlation between whole head PAF and ABC lethargy/withdrawal, $r = .53$, $p = .019$. Another significant correlation was noted between frontal PAF and ADAMS obsessive/compulsive behavior, $r = .46$, $p = .049$. Other exploratory assessments between PAF and clinical variables were assessed and reported in supplement (Supplemental Table 5).

Period 1 (N = 23): BPN14770 differences from baseline were confirmed in period 1 with paired samples t-tests across all regions and found a similar pattern to that observed across both periods apart from occipital PAF. All between subject comparisons met criteria for equal variances across groups via Levene's test. Frontal PAF in the BPN14770 condition ($M = 7.69$, $SD = .90$) was significantly different from baseline ($M = 6.88$, $SD = .47$), occipital PAF in the BPN14770 condition ($M = 7.93$, $SD = 1.28$) was not significantly different from baseline ($M = 7.4$, $SD = .79$), and whole head PAF in the BPN14770 condition ($M = 7.79$, $SD = .75$) trended towards a significance from baseline ($M = 7.26$, $SD = .48$), (frontal: $t(11) = 2.53$, $p = .028$, $d = .73$; occipital: $t(11) = 1.18$, $p = .265$, $d = .34$; whole head: $t(11) = 2.08$, $p = .062$, $d = .60$). Given the presence of carry-over effects, BPN14770 was tested against placebo in period 1. Independent sample t-tests were conducted on period 1 PAF data across frontal, occipital, and whole head regions and found no significant differences between BPN14770 and placebo (Frontal: $t(21) = .302$, $p = .383$, $d = .126$; Occipital: $t(21) = .657$, $p = .259$, $d = .274$; Whole head: $t(21) = .833$, $p = .209$, $d = .348$). Paired samples t-tests were then conducted on period 1 PAF placebo and baseline data across all regions and found no significant differences between placebo and baseline except a marginal differences between baseline and placebo for occipital likely driven by a single individual (Frontal: $t(8) = -.069$, $p = .947$, $d = -.023$; Occipital: $t(8) = 2.28$, $p = .052$, $d = .761$; Whole head: $t(8) = 1.75$, $p = .119$, $d = .526$).

Baseline Correlations (N = 24): Exploratory correlations assessed relationships between pre-dose baseline PAF and pre-dose baseline clinical measures to determine the baseline relationships between PAF and clinical measures of interest (Table 2). A significant positive correlation was observed between frontal PAF and ABC Inappropriate Speech, $r = .43$, $p = .036$. A significant positive correlation was also observed between whole head PAF and ABC inappropriate speech ($r = .46$, $p = .025$), ABC stereotypy

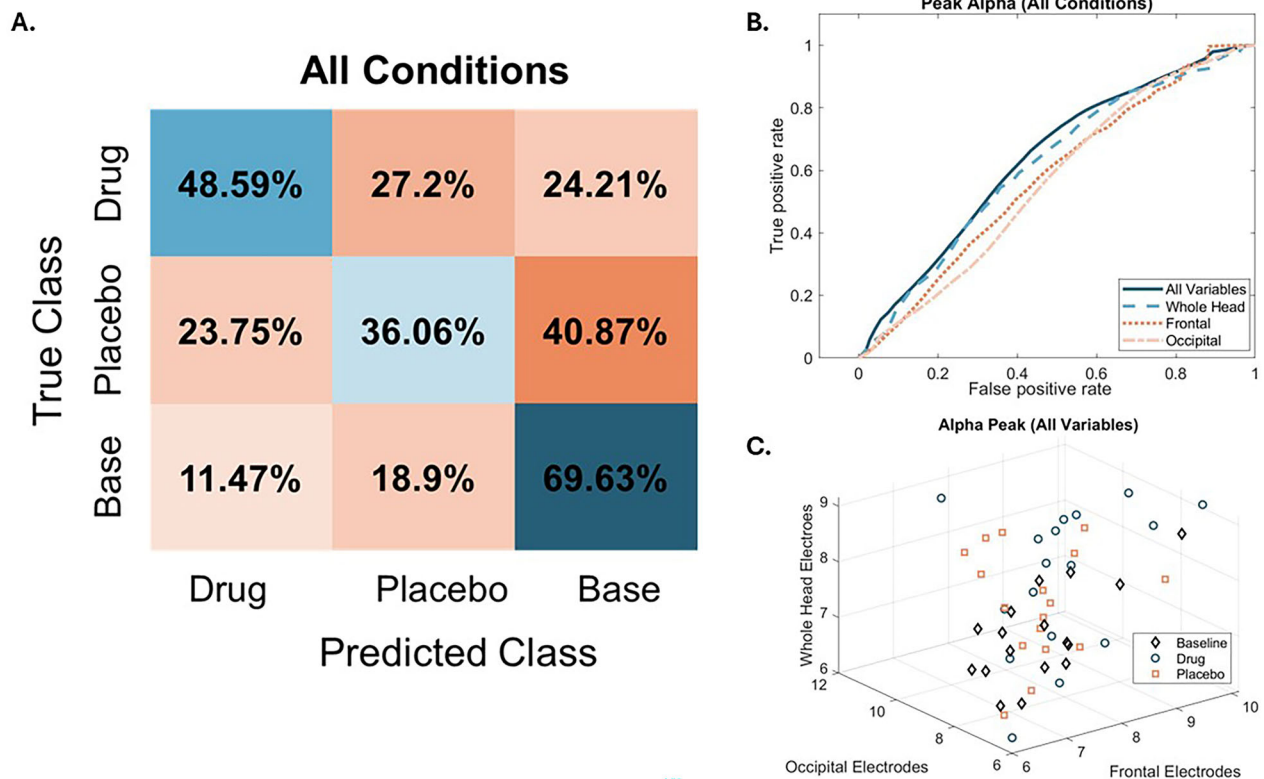


Fig. 4 Average Classification Confusion Matrix Across All PAF Conditions, ROC curve, and Individual Data Points across All Conditions for all peak alpha. **A)** Average classification error confusion matrix showing classification performance in percentages of whether the true class was identified by the NBC. Percentage values reflect the average NBC performance across 2000 iterations of model building and across the 17 participants with data for all conditions. Misclassification is largely due to the overlap of placebo and baseline, but misclassification of drug and placebo can be attributed to carryover effects. **B)** ROC plot for PAF across head regions. Electrode selections included all electrodes (whole head), all frontal electrodes (frontal), all occipital electrodes (occipital), and all variables together (whole head, frontal, and occipital). Peaking toward the left upper corner indicates better performance (i.e., maximizing true positive rate and minimizing false positive rate). **C)** Individual data points for PAF across frontal, occipital, and all (i.e., whole head) electrodes.

($r = .46$, $p = .025$), and VAS daily functioning ($r = .49$, $p = .015$). Finally, a significant positive correlation was found between occipital PAF and ADAMs general anxiety ($r = .41$, $p = .044$), social anxiety ($r = .42$, $p = .044$), and manic and hyperactive behavior ($r = .41$, $p = .040$). Group differences for BPN14770 and placebo for clinical variables of interest are reported in Supplemental Table 6.

DISCUSSION

The current study reflects an exploratory evaluation of secondary outcome measures from a recent, successful phase 2a clinical trial demonstrating BPN14770 efficacy in a small sample of males with FXS. The NBC identified PAF as a variable that adequately separated BPN14770 from baseline and placebo with the strongest predictor of condition being the combination of all PAF variables (frontal, occipital, and whole head PAF) which then demonstrated statistical significance for BPN14770 vs baseline but not placebo vs baseline. Our novel identification of BPN14770 efficacy in PAF adds to previous findings of cognitive improvements with BPN214770, as peak alpha is notably reduced in individuals with FXS and related to cognitive performance in typical development [17, 19, 36, 37].

Improvements in PAF

Phase 2 clinical trials are frequently underpowered where multiple statistical comparisons across proposed biomarkers render secondary outcomes susceptible to false outcomes (i.e., type I and type II error). Utilization of classification algorithms represents a simplistic approach to secondary biomarker assessments in clinical

trials. Not only are outcomes clinically interpretable and mechanistically insightful, where physiological outcomes help bridge known therapeutic mechanisms (i.e., pharmaceutical mechanisms of action) and externally measurable behavior/physiology, but the statistical evaluation is more robust against false discovery/type 1 error [21, 33, 38]. Further, the current methods are easily employed under blinded conditions suggesting clinical trials in NDDs could move toward more data-driven outcome measures to avoid both type 1 and type 2 errors with more definitive efficacy determination.

Using the NBC to explore BPN14770-related physiological shifts identified adequate condition separation (i.e., baseline, drug, placebo) for PAF. Difficulty separating drug from placebo arose from known carryover effects which negatively affected NBC performance due to a lack of washout period in the cross-over design and persistent pharmacological effects in the placebo [11, 16]. The combination of all PAF variables separated drug from baseline with an area under the curve in the fair performance range for clinical use which survived statistical comparison [39]. Importantly, NBC performance on the PAF variable combination was achieved with a very small sample size ($N = 17$) of participants with clean EEG data across all timepoints. Statistical evaluations showed significant improvements in PAF with BPN14770 where PAF shifted into the alpha frequency range for all participants but were limited to evaluations comparing differences in BPN14770 effect from baseline and not placebo due to underpowered comparisons and carryover effects. Evaluating BPN14770 effects from placebo was only possible in period 1 rather than within participant comparisons. Despite the statistical limitations, the

Table 2. Exploratory Correlations Assessing Pre-Dose Relationships between Peak Alpha Frequency and Clinical Measures.

Variable (N)	Baseline Peak Alpha Frequency		
	Frontal PAF	Occipital PAF	Whole Head PAF
NIH-TCB (N = 24)^a			
CCC	-.120	-.081	.156
PV	.064	-.113	.166
ORR	-.286	-.012	.063
PSM (N = 21)	-.103	-.200	.031
PCPS	.159	-.044	-.155
VAS (N = 24)			
Language	.065	-.177	-.040
Anxiety/Irritability	-.093	.088	.103
Daily Functioning	.122	.236	.490*
ABC (N = 24)			
Hyperactivity	.318	.199	.422*
Inappropriate Speech	.429*	.232	.456*
Irritability	.176	.094	.266
Lethargy/Withdrawal	.085	.278	.362
Social Avoidance	-.017	.092	.206
Stereotypy	.233	.327	.457*
ADAMS (N = 24)			
Depressed Mood	-.127	.219	.089
General Anxiety	.048	.414*	.232
Manic/Hyperactive Behavior	.170	.422*	.278
Obsessive/Compulsive Behavior	.034	.311	.149
Social Avoidance	-.086	.415*	.266
Deviation IQ (Z-score)	-.070	-.417*	

Note. All correlations are Spearman's rho. Sample sizes are listed for each task and are the same across all sub-variables unless otherwise noted as N^a. No asterisk = N.S., *p < 0.05, **p < 0.01.

effect size for the linear composite variable assessing BPN14770 vs. Baseline created from all conditions was large and the effect size for BPN14770 vs placebo during period 1 was medium sized suggesting a moderate effect of BPN14770 on PAF. Further, the moderate effect of BPN14770 was driven by frontal measures of PAF indicating the effect may be specific to the thalamocortical generator of alpha [17].

Shifts in PAF are relevant given the nature of the clinical improvements initially observed in the cognitive domain despite not sharing relationships with clinical variables that initially demonstrated improvements [11]. PAF is associated with cognitive performance in typically developed individuals and correlates highly with measures of cognitive performance in both idiopathic ASD and samples of FXS that include females [17, 40, 41]. Further, PAF is typically considered highly stable in typically developed younger and older adults with evidence showing PAF was not easily modified by cognitive interventions alone on a larger time scale [42]. While PAF can fluctuate on smaller timescales and with varying task demands, males with FXS demonstrate difficulties initiating alpha frequency oscillations on a dynamic scale suggesting any intraindividual shifts into the alpha frequency range across time may be meaningful [43]. The current neocortical

hyperexcitability model of FXS includes increased power in gamma and theta with decreased power in alpha [18]. Increased N1 amplitude in the ERP to a novel/initial auditory stimulus are thought to reflect neural hyperexcitability in FXS and related to blood serum levels of BPN14770 [15, 16, 44]. Combining increases in PAF with marginal improvements in the N1 amplitude measured from frontocentral electrodes reported in the original article detailing the clinical trial findings, BPN14770 may enhance an individual with FXS's ability to initiate alpha frequency and organize neural networks necessary for supporting the process of temporal integration of information underlying both sensory processing and cognitive performance in frontocentral brain regions [11, 20, 43]. Ultimately, given the lack of improvement observed from placebo, the adequate but not excellent AUCs for separating conditions, and the sample size limitations on statistical power, more work and a replication are required to add support to the current conclusions and establish PAF changes as a biomarker of efficacy for BPN14770.

Limitations and Conclusions

First, the original manuscript reported clinical findings before and after each treatment arm (i.e., 5 total measurements) but EEG was recorded only at the end of each treatment arm (i.e., 3 total measurements) [11]. Thus, the current study was limited to one measurement per treatment arm for both clinical and EEG evaluations (see Fig. 2). Clinical measure outcome differences between the original and current study (see supplement) may reflect the use of only 3 measurement points in the current study where 5 were used in the original or the reduction in sample size from 30 with clinical outcomes to 17 with complete EEG outcomes. Second, carryover effects were a major limitation for evaluating differences between BPN14770 and placebo. Third, we were unable to include the same covariates from the original study (i.e., IQ and baseline measures) due to 1.) the use of baseline measures in the analyses, and 2.) the relationship between IQ and PAF. Fourth, the use of machine learning to separate conditions also assumes that baseline values for the sample do not overlap significantly with the drug sample values, and so is only appropriate for variables where a mean shift at the group level in values is expected. In circumstances where there is a large range of variability in baseline values, with some individuals showing minimal impairment that overlaps with improved scores with treatment for those with lower baseline values, the classifier technique may underperform. We selected input variables with known deficits at the group level for FXS, therefore despite some overlap in conditions (Fig. 4C) we do not expect this limitation to be significantly problematic for our study. However, this is an important consideration for use of this technique in trials with larger ranges of baseline ability. Finally, the NBC has an assumption of independence of features and certain EEG measures are likely correlated (e.g., power across bands) in certain instances. Although there is debate on how impactful this assumption is, dependence of some features likely reduces the effectiveness of the ML approach [45]. However, one benefit of the NBC independence assumption is that it allows it to learn high dimensional features and achieve model fit with much smaller training sets than many other classifiers [32, 46]. Despite these limitations, this study demonstrates the possibility of screening large numbers of exploratory variable combinations using machine learning, ROC evaluation, and composite variable generation in small sample sizes to detect novel effects of treatment on brain physiology. The composite PAF variable identified is relevant to the significant cognitive outcomes for BPN14770, has been demonstrated to be impaired in FXS [17], and is more robust to data loss than ERP measures, thus more scalable to larger trials. Future work will be necessary to validate PAF in larger sample sizes to assess the extent to which BPN14770 modulates cognition-related biomarkers in FXS.

DATA AVAILABILITY

Data are property of Shionogi & Company. Requests for data can be made to Shionogi & Company. Analytic scripts are available upon request to the corresponding author.

REFERENCES

- Santoro MR, Bray SM, Warren ST. Molecular mechanisms of fragile X syndrome: a twenty-year perspective. *Annu Rev Pathol.* 2021;7:219–45. <https://doi.org/10.1146/annurev-pathol-011811-132457>.
- Straub D, Schmitt LM, Boggs AE, Horn PS, Dominick KC, Gross C, et al. A sensitive and reproducible qRT-PCR assay detects physiological relevant trace levels of FMR1 mRNA in individuals with Fragile X syndrome. *Sci Rep.* 2023;13:3808. <https://doi.org/10.1038/s41598-023-29786-4>.
- Salcedo-Arellano MJ, Dufour B, McLennan Y, Martinez-Cerdeno V, Hagerman R. Fragile X syndrome and associated disorders: Clinical aspects and pathology. *Neurobiol Dis.* 2020;136:104740. <https://doi.org/10.1016/j.nbd.2020.104740>.
- Schmitt LM, Wang J, Pedapati EV, Thurman AJ, Abbeduto L, Erickson CA, et al. A neurophysiological model of speech production deficits in fragile X syndrome. *Brain Commun.* 2020;2:fcz042. <https://doi.org/10.1093/braincomms/fcz042>.
- Hagerman PJ, Hagerman R. Fragile X syndrome. *Curr Biol.* 2021;31:R273–R275. <https://doi.org/10.1016/j.cub.2021.01.043>.
- Abbeduto L, Thurman AJ, McDuffie A, Klusek J, Feigles RT, Brown TW, et al. ASD Comorbidity in Fragile X Syndrome: Symptom Profile and Predictors of Symptom Severity in Adolescent and Young Adult Males. *J Autism Dev Disord.* 2019;49:960–77. <https://doi.org/10.1007/s10803-018-3796-2>.
- Thurman, Swinehart AJ, Klusek SS, Roberts J, Bullard JE, Marzan L, et al. Daily Living Skills in Adolescent and Young Adult Males With Fragile X Syndrome. *Am J Intellect Dev Disabil.* 2022;127:64–83. <https://doi.org/10.1352/1944-7558-127.1.64>.
- Usher LV, DaWalt LS, Hong J, Greenberg JS, Mailick MR. Trajectories of Change in the Behavioral and Health Phenotype of Adolescents and Adults with Fragile X Syndrome and Intellectual Disability: Longitudinal Trends Over a Decade. *J Autism Dev Disord.* 2020;50:2779–92. <https://doi.org/10.1007/s10803-020-04367-w>.
- Weber JD, Smith E, Berry-Kravis E, Cadavid D, Hessl D, Erickson C. Voice of People with Fragile X Syndrome and Their Families: Reports from a Survey on Treatment Priorities. *Brain Sci.* 2019;9. <https://doi.org/10.3390/brainsci9020018>.
- Berry-Kravis EM, Lindemann L, Jonch AE, Apostol G, Bear MF, Carpenter RL, et al. Drug development for neurodevelopmental disorders: lessons learned from fragile X syndrome. *Nat Rev Drug Discov.* 2018;17:280–99. <https://doi.org/10.1038/nrd.2017.221>.
- Berry-Kravis EM, Harnett MD, Reines SA, Reese MA, Ethridge LE, Outtersen AH, et al. Inhibition of phosphodiesterase-4D in adults with fragile X syndrome: a randomized, placebo-controlled, phase 2 clinical trial. *Nat Med.* 2021;27:862–70. <https://doi.org/10.1038/s41591-021-01321-w>.
- Berry-Kravis E, Hicar M, Ciurlionis R. Reduced cyclic AMP production in Fragile X Syndrome: cytogenetic and molecular correlations. *Pediatric Res.* 1995;38:638–43. 0031-3998/95/3805-0638\$03.00/0.
- Berry-Kravis E. Overexpression of fragile X gene (FMR-1) transcripts increases cAMP production in neural cells. *J Neurosci Res.* 1998;51:41–48. [https://doi.org/10.1002/\(sici\)1097-4547](https://doi.org/10.1002/(sici)1097-4547).
- Kanellopoulos AK, Semelidou O, Kotini AG, Anezaki M, Skoulakis EM. Learning and memory deficits consequent to reduction of the fragile X mental retardation protein result from metabotropic glutamate receptor-mediated inhibition of cAMP signaling in *Drosophila*. *J Neurosci.* 2012;32:13111–24. <https://doi.org/10.1523/JNEUROSCI.1347-12.2012>.
- Ethridge LE, White SP, Mosconi MW, Wang J, Byerly MJ, Sweeney JA. Reduced habituation of auditory evoked potentials indicate cortical hyper-excitability in Fragile X Syndrome. *Transl Psychiatry.* 2016;6:e787. <https://doi.org/10.1038/tp.2016.48>.
- Norris JE, Berry-Kravis EM, Harnett MD, et al. Auditory N1 event-related potential amplitude is predictive of serum concentration of BPN14770 in fragile X syndrome. *Mol Autism.* 2024;15:47. <https://doi.org/10.1186/s13229-024-00626-0>.
- Pedapati EV, Schmitt LM, Ethridge LE, Miyakoshi M, Sweeney JA, et al. Neocortical localization and thalamocortical modulation of neuronal hyperexcitability contribute to Fragile X Syndrome. *Commun Biol.* 2022;5:442 <https://doi.org/10.1038/s42003-022-03395-9>.
- Wang J, Ethridge LE, Mosconi MW, White SP, Binder DK, Pedapati EV, et al. A resting EEG study of neocortical hyperexcitability and altered functional connectivity in fragile X syndrome. *J Neurodev Disord.* 2017;9. <https://doi.org/10.1186/s11689-017-9191-z>.
- Angelakis E, Stathopoulou S, Frymiare JL, Green DL, Lubar JF, Kounios J. EEG neurofeedback: a brief overview and an example of peak alpha frequency training for cognitive enhancement in the elderly. *Clin Neuropsychol.* 2007;21:110–29. <https://doi.org/10.1080/13854040600744839>.
- Norris JE, DeStefano LA, Schmitt LM, Pedapati EV, Erickson CA, Sweeney JA. Hemispheric Utilization of Alpha Oscillatory Dynamics as a Unique Biomarker of Neural Compensation in Females with Fragile X Syndrome. *ACS Chem Neurosci.* 2022;13:3389–402. <https://doi.org/10.1021/acscchemneuro.2c00404>.
- Rathees S, Bhatia D, Punia V, Singh R. Peak Alpha Frequency in Relation to Cognitive Performance. *J Neurosci Rural Pr.* 2020;11:416–9. <https://doi.org/10.1055/s-0040-1712585>.
- Richard Clark C, Veltmeyer MD, Hamilton RJ, et al. Spontaneous alpha peak frequency predicts working memory performance across the age span. *Int J Psychophysiol.* 2004;53:1–9. <https://doi.org/10.1016/j.jpsycho.2003.12.011>.
- Busch N, Geyer T, Zinchenko A. Individual peak alpha frequency does not index individual differences in inhibitory cognitive control. *Psychophysiology.* 2024;61:e14586. <https://doi.org/10.1111/psyp.14586>.
- Fogel DB. Factors associated with clinical trials that fail and opportunities for improving the likelihood of success: A review. *Contemp Clin Trials Commun.* 2018;11:156–64. <https://doi.org/10.1016/j.conctc.2018.08.001>.
- Hsu M-J, Chang Y-CI, Hsueh H-M. Biomarker selection for medical diagnosis using the partial area under the ROC curve. *BMC Research Notes.* 2014;7. <https://doi.org/10.1186/1756-0500-7-25>.
- Liu D, Zhou XH. ROC analysis in biomarker combination with covariate adjustment. *Acad Radio.* 2013;20:874–82. <https://doi.org/10.1016/j.acra.2013.03.009>.
- Ethridge LE, Pedapati EV, Schmitt LM, et al. Validating brain activity measures as reliable indicators of individual diagnostic group and genetically mediated subgroup membership Fragile X Syndrome. *Sci Rep.* 2024;14:22982. <https://doi.org/10.1038/s41598-024-72935-6>. Published 2024 Oct 3.
- Ahsan MM, Luna SA, Siddique Z. Machine-Learning-Based Disease Diagnosis: A Comprehensive Review. *Healthcare (Basel).* 2022;10:541 <https://doi.org/10.3390/healthcare10030541>.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods.* 2004;134:9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- Allen JJB, Cohen MX. Deconstructing the “resting” state: exploring the temporal dynamics of frontal alpha asymmetry as an endophenotype for depression. *Front Hum Neurosci.* 2010;4:232 <https://doi.org/10.3389/fnhum.2010.00232>.
- Sordo M, Zeng Q. On Sample Size and Classification Accuracy: A Performance Comparison. In: Oliveira, JL, Maojo, V, Martín-Sánchez, F, Pereira, AS (eds) *Biological and Medical Data Analysis*. ISBMDA 2005. *Lecture Notes in Computer Science*, 2005;3745:193–201. https://doi.org/10.1007/11573067_20.
- Guo Y, Graber A, McBurney RN, Balasubramanian R. Sample size and statistical power considerations in high-dimensionality data settings: a comparative study of classification algorithms. *BMC Bioinforma.* 2010;11:447. <https://doi.org/10.1186/1471-2105-11-447>.
- Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, et al. Cognition assessment using the NIH Toolbox. *Neurology.* 2013;80:S54–64. <https://doi.org/10.1212/WNL.0b013e3182872ded>.
- Aman MG, Singh NN, Stewart AW, Field CJ. The Aberrant Behavior Checklist: a behavior rating scale for the assessment of treatment effects. *Am J Ment Defic.* 1985;89:485–91. <https://doi.org/10.1037/t10453-000>.
- Esbensen AJ, Rojahn J, Aman MG, Ruedrich S. Reliability and validity of an assessment instrument for anxiety, depression, and mood among individuals with mental retardation. *J Autism Dev Disord.* 2003;33:617–29. <https://doi.org/10.1023/b:jadd.000005999.27178.55>.
- Cellier D, Riddle J, Petersen I, Hwang K. The development of theta and alpha neural oscillations from ages 3 to 24 years. *Dev Cogn Neurosci.* 2021;50:100969. <https://doi.org/10.1016/j.dcn.2021.100969>.
- Proteau-Lemieux M, Knoth IS, Agbogba K, Cote V, Barlahan Biag HM, Thurman AJ, et al. EEG Signal Complexity Is Reduced During Resting-State in Fragile X Syndrome. *Front Psychiatry.* 2021;12:716707. <https://doi.org/10.3389/fpsyt.2021.716707>.
- Gupta C, Chandrashekar P, Jin T, He C, Khullar S, Chang Q, et al. Bringing machine learning to research on intellectual and developmental disabilities: taking inspiration from neurological diseases. *J Neurodev Disord.* 2022;14:28 <https://doi.org/10.1186/s11689-022-09438-w>.
- Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol.* 2022;75:25–36. <https://doi.org/10.4097/kja.21209>.
- Angelakis E, Lubar JF, Stathopoulou S, Kounios J. Peak alpha frequency: an electroencephalographic measure of cognitive preparedness. *Clin Neurophysiol.* 2004;115:887–97. <https://doi.org/10.1016/j.clinph.2003.11.034>.
- Dickinson A, DiStefano C, Senturk D, Jeste SS. Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *Eur J Neurosci.* 2018;47:643–51. <https://doi.org/10.1111/ejn.13645>.
- Grandy TH, Werkle-Bergner M, Chicherio C, Schmiedek F, Lövdén M, Lindenberger U. Peak individual alpha frequency qualifies as a stable neurophysiological trait marker in healthy younger and older adults. *Psychophysiology.* 2013;50:570–82. <https://doi.org/10.1111/psyp.12043>.

43. Mierau A, Klimesch W, Lefebvre J. State-dependent alpha peak frequency shifts: Experimental evidence, potential mechanisms and functional implications. *Neuroscience*. 2017;360:146–54. <https://doi.org/10.1016/j.neuroscience.2017.07.037>.
44. Ethridge LE, De Stefano LA, Schmitt LM, Woodruff NE, Brown KL, Tran M, et al. Auditory EEG Biomarkers in Fragile X Syndrome: Clinical Relevance. *Front Integr Neurosci*. 2019;13:60 <https://doi.org/10.3389/fnint.2019.00060>.
45. Lewis, DD Naive (Bayes) at forty: The independence assumption in information retrieval. In: Nédellec, C, Rouveilol, C (eds) *Machine Learning: ECML-98. Lecture Notes in Computer Science*, vol 1398. 1998; Springer, Berlin, Heidelberg. <https://doi.org/10.1007/BFb0026666>.
46. Sordo M & Zeng Q in *Biological and Medical Data Analysis*. (eds Oliveira José Luís, Maojo Víctor, Martín-Sánchez Fernando, & Pereira António Sousa) 193–201 (Springer; Berlin Heidelberg).

ACKNOWLEDGEMENTS

We thank the patients and their families for participating in the clinical trial. Direct clinical costs were funded by the FRAXA Research Foundation. Access to and training on the NIH-TCB was obtained in association with work on HD076189 (David Hessl). Tetra Therapeutics provided drug product and funded trial administration, independent data analysis, and provided publication support. NFXF Randi Hagerman Summer Scholar Award funded J.E.N and the secondary analysis for the current study. Additional financial support for publication was provided by the University of Oklahoma Libraries' Open Access Fund.

AUTHOR CONTRIBUTIONS

J.E.N. conducted all analyses and drafted the manuscript, E.M.B.-K. designed and conducted the clinical trial, M.D.H. led the biostatistical analysis for the original manuscript, S.A.R. designed the clinical trial and served as medical monitor, M.R.S. and E.K.A. preprocessed and organized the EEG dataset, A.H.O., and J.F. conducted the clinical trial and obtained EEG recordings, M.E.G. contributed to the clinical protocol, L.E.E. supervised all aspects of EEG experimental design, data analysis, and manuscript preparation, All authors contributed significantly to manuscript preparation.

COMPETING INTERESTS

E.M.B.-K., M.R.S., E.K.A., A.O., C.M. and J.F. declare no competing interests. M.D.H. and S.A.R. are paid consultants to Tetra Therapeutics. M.E.G. is an employee of Tetra Therapeutics, which is a wholly owned subsidiary of Shionogi & Company that has a financial interest in BPN14770. J.E.N. and L.E.E. received research funding from Shionogi & Company for independent data analysis; all funds are contracted to and managed by the University of Oklahoma.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-025-03558-2>.

Correspondence and requests for materials should be addressed to Lauren E. Ethridge.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ethics approval and consent to participate: The clinical trial is registered with ClinicalTrials.gov identifier: NCT03569631. All participants or their legal guardians signed informed consent which included consent for EEG data collection. The clinical site was Rush University Medical Center (RUMC) where all study documents, including study protocol, consent documents, recruitment materials, safety information for participants, and information about study compensation were approved by the RUMC institutional review board (Approval #00000482) [11]. Participant recruitment was managed by RUMC and supplemented by FXS patient advocacy groups. All EEG procedures were approved by the University of Oklahoma institutional review board (Approval #10129). All study methods were conducted according to the ethical guidelines of the Declaration of Helsinki, seventh revision, 64th World Medical Association General Assembly Meeting and are consistent with the International Conference on Harmonization/good clinical practice, applicable regulatory requirements and the sponsor or its delegate's policy on bioethics.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025