

École polytechnique de Louvain

EEG response to hyperventilation as a biomarker of clinical response and longitudinal changes in brain functional connectivity induced by vagus nerve stimulation

Author: **Marion DE KERCKHOVE**

Supervisors: **Pierre-Antoine ABSIL, Riëm EL TAHRY**

Readers: **Enrique GERMANY MORRISON, John LEE**

Academic year 2025–2026

Master [120] in Biomedical Engineering

Acknowledgements

I would first like to express my gratitude to Professor Pierre-Antoine Absil for supervising this thesis and for his guidance throughout this work. Our regular meetings provided valuable opportunities to discuss progress, refine ideas, and address challenges as they arose. His insightful feedback, availability, and kindness greatly contributed to the development of this project.

I would also like to sincerely thank Professor Riëm El Tahry for her supervision and involvement throughout this project. She consistently followed the progress of this work and provided valuable feedback at its different stages. Her engagement, high standards, and positive appreciation of the work accomplished were deeply valued and represented an important source of motivation throughout this thesis.

I am also grateful to the IoNS team for their involvement throughout this project and for the monthly meetings that allowed me to present and discuss the progress of this work. In particular, I would like to thank Doctor Enrique Germany Morrison for his continued presence during these meetings and for taking the time to answer questions beyond them. I would also like to thank Raphaël Vicini for introducing me to the dataset and providing the necessary explanations for its use, as well as Professor Antoine Nonclercq for his thoughtful questions and feedback during the later stages of this project.

I would also like to thank Céline Bugli from the Support en Méthodologie et Calcul Statistique (SMCS) platform for her methodological guidance and for taking the time to review and discuss aspects of the machine-learning pipeline used in this project.

Finally, I would like to thank my family, friends, and especially my partner for their support and encouragement during the completion of this thesis. Their presence and patience were particularly appreciated during the more stressful periods of this project.

AI disclaimer

Artificial intelligence tools were used to assist in the linguistic refinement of this manuscript (e.g., grammar, clarity, and style). All ideas, analyses, and conclusions are the author's own work.

Abstract

Objectives

Vagus nerve stimulation (VNS) is a neuromodulation therapy used as an adjunctive treatment for patients with drug-resistant epilepsy who are not suitable candidates for resective surgery. Despite its proven efficacy, the mechanisms underlying its therapeutic effects remain incompletely understood, and no reliable biomarkers are currently available to predict or assess patient response. Cortical desynchronization has been proposed as one of the main mechanisms of action of VNS and can be investigated through functional connectivity measures. Hyperventilation (HV) is a well-established activation procedure in routine EEG recordings, but its impact on functional connectivity remains unclear. In this study, we investigated the predictive value of hyperventilation-induced EEG synchronization compared with resting state and evaluated how VNS modulates these synchronization patterns longitudinally.

Materials and methods

Twelve adult patients with drug-resistant epilepsy were followed before and at one, three, and six months after VNS device implantation. High-density EEG recordings were acquired during resting state (eyes open and eyes closed) and hyperventilation, under VNS ON and OFF conditions. The global and regional weighted phase lag index (wPLI) was computed as a measure of phase synchronization and correlated with the VNS response using multiple assessment methods, including binary classification ($> 50\%$ vs $< 50\%$ seizure reduction), percentage of seizure reduction, and the recently developed Clinical-Research Response Scale (CRRS). Linear mixed-effects models were used to assess longitudinal and condition-specific effects. In addition, several regression models were trained using selected feature combinations to evaluate and compare the predictive value of each condition, with

an emphasis on model stability.

Results

During hyperventilation, lower values of global and regional wPLI at the preimplantation visit in the beta band, and more specifically the lower part of the beta band (13–20 Hz), were associated with better response to VNS, particularly when assessed using the CRRS. Although significant resting-state findings were limited, the best-performing predictive model combined global delta-band resting-state wPLI with beta-band hyperventilation wPLI, yielding a mean absolute error of 1.887 on the CRRS scale ranging from 0 to 15.

Longitudinally, a progressive decrease in delta-band wPLI was observed during the eyes-closed VNS OFF condition, which correlated with VNS response over time when assessed using the CRRS. During hyperventilation, longitudinal changes in wPLI were primarily observed in the theta, high-beta, and broadband frequency bands, with better clinical outcomes associated with increasing synchronization over time.

No significant acute differences were observed between VNS ON and OFF conditions.

Conclusion

Hyperventilation reveals distinct EEG synchronization patterns associated with subsequent response to VNS. Preimplantation hyperventilation-induced wPLI in the beta band may serve as a predictive biomarker, providing complementary information to resting-state measures. The chronic effects of VNS on hyperventilation-induced synchronization remain less clear and warrant further investigation. Nonetheless, the observed longitudinal decrease in resting-state delta-band wPLI is consistent with a progressive neural desynchronization.

Contents

1	Introduction	8
1.1	Vagus nerve stimulation (VNS)	8
1.2	Brain functional connectivity	9
1.3	Connectivity metrics for EEG	10
1.4	Phase-synchronization measures	11
1.4.1	Phase-locking value (PLV)	11
1.4.2	Phase coherence (PC)	12
1.4.3	Imaginary part of coherency (IC)	12
1.4.4	Phase lag index (PLI)	13
1.4.5	Weighted phase lag index (wPLI)	14
1.4.6	Phase slope index (PSI)	16
1.5	VNS and brain connectivity	16
1.6	Hyperventilation	19
1.7	Objectives of the study	21
2	Material and methods	22
2.1	Study population and data acquisition	22
2.1.1	Study population	22
2.1.2	Study design and EEG acquisition	23
2.1.3	EEG preprocessing	23
2.1.4	Clinical response evaluation	24
2.1.5	Functional connectivity analysis	25

2.2	Statistical analysis	26
2.2.1	Cross-sectional analyses at baseline	27
2.2.2	Longitudinal analyses	28
2.2.3	Acute stimulation ON/OFF analyses	29
2.3	Machine Learning Pipeline	30
2.3.1	Feature selection	30
2.3.2	Regression models	31
2.3.3	Nested cross-validation	33
2.3.4	Performance metrics	34
2.3.5	Permutation testing	34
2.3.6	Model stability analysis	35
2.3.7	Comparison of resting-state and hyperventilation feature sets	36
2.3.8	Learning curve analysis	37
2.3.9	Power analysis	37
3	Results	40
3.1	Demographic and clinical characteristics	40
3.2	Clinical response to VNS	41
3.3	Predictive value of pre-implantation EEG recordings	42
3.3.1	Binary response scale	42
3.3.2	Seizure reduction percentage	43
3.3.3	CRRS	43
3.3.4	Exploratory trend analysis	44
3.3.5	Machine learning prediction of CRRS	46
3.4	Longitudinal evolution of wPLI following VNS implantation	51
3.4.1	Global longitudinal changes	51
3.4.2	Binary response	51
3.4.3	Seizure reduction percentage	52
3.4.4	CRRS	53

3.5	Acute and chronic effect of VNS comparison	54
4	Discussion	56
4.1	Main findings	56
4.2	Comparison with previous investigations on the same cohort	57
4.3	Interpretation of the machine learning model	59
4.4	Comparison with the broader literature	60
4.5	Limitations	61
4.6	Perspectives and further investigations	61
4.7	Conclusion	62
A	Baseline (V1) statistical results	69
B	Longitudinal evolution of wPLI	76
C	Permutation-test distributions	85

Chapter 1

Introduction

1.1 Vagus nerve stimulation (VNS)

Epilepsy is a common neurological disorder defined by the International League Against Epilepsy (ILAE) as “a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures” [1]. It encompasses a wide range of etiologies and clinical manifestations, affecting individuals of all ages. According to the World Health Organization (WHO), approximately 50 million people worldwide live with epilepsy, making it one of the most prevalent neurological diseases globally. The condition is associated with significant morbidity, including increased risk of injury, cognitive impairment, and psychosocial consequences, as well as elevated mortality rates. Despite the availability of antiseizure medications, up to 40 percent of patients remain drug-resistant, highlighting the need for alternative therapeutic strategies [2].

Vagus nerve stimulation (VNS) is a neuromodulation therapy used as an adjunctive treatment for patients with drug-resistant epilepsy who are not suitable candidates for resective surgery. The technique consists of delivering intermittent electrical stimulation to the left vagus nerve via an implanted pulse generator, typically placed subcutaneously in the chest and connected to a helical electrode wrapped around the cervical portion of the nerve (Figure 1.1). Stimulation is administered according to programmable cycles (e.g., alternating on/off periods), enabling long-term modulation of neural activity [3].

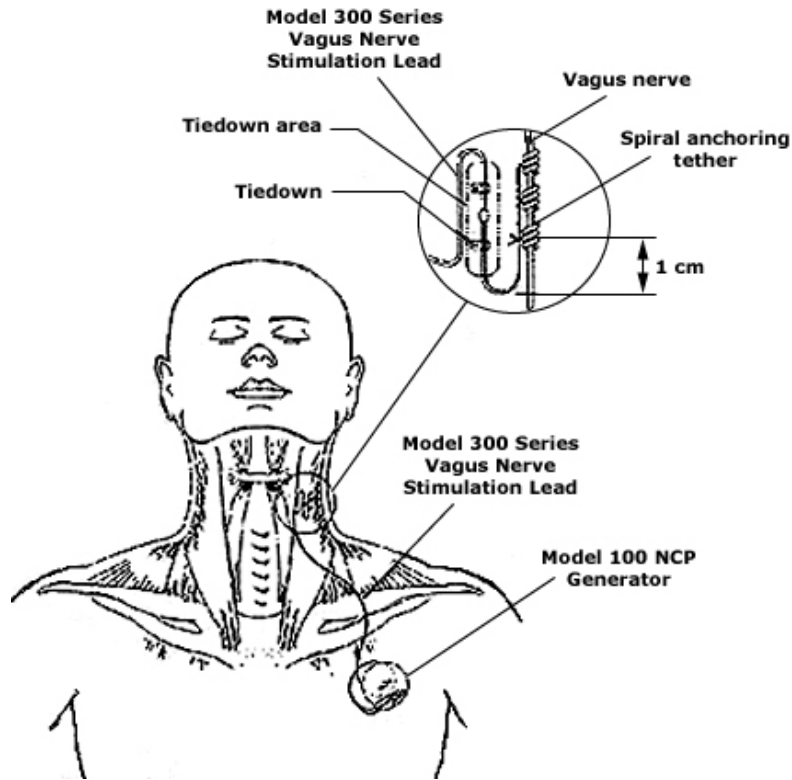


Figure 1.1: Vagal nerve stimulator [3]

VNS has demonstrated efficacy across multiple epilepsy etiologies, with large clinical series reporting responder rates (>50 percent reduction of seizures) of up to 65 percent; the therapy is generally well tolerated, with side effects such as hoarseness, cough, or throat discomfort occurring primarily during stimulation periods. In addition to its antiepileptic effects, VNS has been associated with improvements in mood and quality of life, suggesting broader neuromodulatory influences on brain function. However, the exact mechanisms underlying the therapeutic effects of VNS are still not fully elucidated [3]. One of the main hypotheses is that VNS induces desynchronization of cortical activity, which was first reported in early animal studies in the 1960s [4, 5] and further supported by several functional connectivity studies in humans (see Section 1.5 below).

1.2 Brain functional connectivity

Brain functional connectivity (FC) refers to the statistical dependency between the activity of neural units located in different anatomical regions. It is distinct from

structural connectivity, which describes physical axonal pathways, and from effective connectivity, which captures directed causal influences based on mechanistic models [6].

FC can be studied using different neuroimaging modalities, mainly functional magnetic resonance imaging (fMRI) [7] and electroencephalography (EEG) [8]. The relatively low cost and non-invasive nature of EEG, combined with its high temporal resolution, make it particularly well suited for FC analysis [8, 9, 6]. However, it also presents several challenges, including a high level of acquisition artifacts, although these can be mitigated through appropriate preprocessing, and the issue of volume conduction. Because scalp electrodes integrate neural activity over large and potentially overlapping volumes, the activity of a single neural source can spread across multiple electrodes simultaneously, leading to spurious statistical dependencies between electrode pairs [10, 9, 6].

1.3 Connectivity metrics for EEG

Numerous data-driven methods have been proposed to estimate functional connectivity (FC) from EEG recordings. According to Chiarion et al., these approaches can be classified according to several dimensions, including whether they assess pairwise or multivariate interactions, directed or non-directed relationships, and whether they rely on model-based or model-free formulations [6]. FC estimators can also be categorized according to the domain in which signal interactions are studied: time-domain, frequency-domain, and information-theoretic approaches.

Time-domain approaches evaluate temporal relationships between EEG signals. Non-directed measures mainly rely on correlation-based metrics, including static and dynamic correlation measures and their multivariate or partial extensions, allowing the distinction between direct and indirect interactions. Directed approaches are largely based on the concept of Granger causality and its conditional formulations, which assess time-lagged influences between signals [6].

Frequency-domain approaches are particularly relevant for EEG because they characterize interactions within physiologically meaningful oscillatory bands. Non-directed metrics include coherence-based measures and their partial or multivariate variants, while directed methods comprise measures such as directed coherence, partial directed coherence, and directed transfer function, designed to characterize frequency-specific causal interactions [6].

Information-theoretic approaches provide a model-free framework capable of describing both linear and non-linear dependencies. Their non-directed formulations

include mutual-information-based measures, whereas directed interactions can be assessed through transfer entropy and related conditional or multivariate extensions [6].

Additional approaches have also emerged for characterizing brain interactions, including phase-synchronization measures and graph-theoretical frameworks. Since the present study focuses on phase-based connectivity, these measures are described in greater detail in the following section.

1.4 Phase-synchronization measures

Phase synchronization is a common phenomenon extensively studied in the literature, characterized by the synchronization of instantaneous phases between two signals, which can occur even when their amplitudes remain uncorrelated [6]. The instantaneous phase difference between two signals can be used to estimate FC, resulting in a family of estimators known as phase-based measures [8]. Unlike model-based approaches, they do not require stationarity and are robust to amplitude fluctuations [6, 10]. Because of their advantages, phase-based measures have progressively become the most used FC estimators, representing up to 52% of papers regarding FC studies on EEG recordings in a recent review [8].

The instantaneous phase of a signal can be extracted either via the Hilbert transform or via convolution with a complex continuous wavelet (CCWT, e.g. the Morlet wavelet), after band-pass filtering the signal in the frequency range of interest; both approaches yield comparable results [8]. As the instantaneous phase is well-defined only for narrow band signals, a preliminary band-pass filtering of the signal in the frequency range of interest is required to filter the data, usually within conventional EEG frequency bands [6, 8]. This might seem redundant in the case of CCWT as the wavelet convolution already acts as a filter, but it has actually been shown to lead to a more accurate detection of the phase synchronization [8, 10].

1.4.1 Phase-locking value (PLV)

The phase-locking value (PLV) was introduced by Lachaux et al. [10] as a method for quantifying frequency-specific phase synchronization between two neuroelectric signals. It is applied to responses evoked by repeated stimuli to identify latencies at which the phase difference between signals remains consistent across trials, a phenomenon referred to as phase-locking (or phase synchrony). The instantaneous phase $\phi(t, n)$ is extracted at a target frequency for each time t and trial n via bandpass filtering followed by wavelet convolution. Finally, the PLV is computed

as the consistency of the phase difference $\Delta\varphi$ across trials at a time t :

$$\text{PLV}(t) = \frac{1}{N_{\text{trials}}} \left| \sum_{n=1}^{N_{\text{trials}}} e^{i\Delta\phi(t,n)} \right| \quad (1.1)$$

The PLV ranges from 0 (no phase consistency across trials) to 1 (perfect phase locking), and is by definition independent of signal amplitude and does not require stationarity of the signal[10].

However, the authors highlight two important limitations. First, the PLV is sensitive to volume conduction: apparent synchrony between recordings may not reflect true neural interactions but instead arise from the overlap of the volumes recorded by two electrodes. As a result, high PLV values can be observed even in the absence of true interactions between distinct neural sources [10, 11, 12]. Second, the PLV is sensitive to the choice of reference electrode; its effect on synchrony estimates is difficult to predict, as it may either attenuate or artificially inflate measured synchrony depending on how common signal components are altered [10, 12].

1.4.2 Phase coherence (PC)

Phase coherence, also referred to as mean phase coherence, was introduced by Mormann et al. [13] as a statistical measure of phase synchronization based on the distribution of instantaneous phase differences between two signals. After extracting the instantaneous phases $\phi_a(t)$ and $\phi_b(t)$ (e.g., using the Hilbert transform), the relative phase $\Delta\phi(t)$ is computed and represented as a point on the unit circle in the complex plane. PC is then defined as the length of the mean resultant vector:

$$R = \left| \frac{1}{N} \sum_{k=1}^N e^{i\Delta\phi(t_k)} \right| \quad (1.2)$$

where t_k are discrete time-steps within a recording [13]. Structurally, this formula is identical to that of the PLV. The two measures differ only in the dimension over which averaging is performed: PC averages over time samples within a recording window, making it suited to continuous or resting-state recordings, whereas PLV averages over trials, capturing synchrony that is consistently locked to an experimental event.

1.4.3 Imaginary part of coherency (IC)

To address the problem of spurious synchronization caused by volume conduction and common reference effects, Nolte et al. [14] proposed the use of the imaginary

part of coherency (IC). Their approach relies on the observation that, under the quasi-static approximation generally assumed for EEG (i.e. that an observed scalp potential has no time-lag to the underlying source activity), volume conduction does not introduce time delays between a neural source and the signals recorded at different electrodes. Consequently, correlations arising purely from a common source should have zero phase lag and therefore produce a purely real coherency, while the imaginary part of coherency can only arise from interactions involving a non-zero phase lag between distinct neural processes [14].

Coherency is the normalized complex cross-spectrum between two signals:

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}} \quad (1.3)$$

where $S_{ij}(f)$ is the cross-spectrum between signals i and j at frequency f . If the Fourier transforms of the signals are written as

$$x_i(f) = r_i e^{i\Phi_i}, \quad x_j(f) = r_j e^{i\Phi_j}, \quad (1.4)$$

the cross-spectrum can be expressed as:

$$S_{ij}(f) = \langle r_i r_j e^{i\Delta\Phi} \rangle \quad (1.5)$$

with $\Delta\Phi = \Phi_i - \Phi_j$ the phase difference between the two signals [14].

The IC is simply the imaginary part of this coherency, denoted $IC = \Im\{C_{ij}(f)\}$. Positive and negative values reflect the direction of the phase delay between signals.

The authors emphasize that the robustness of IC to spurious synchronization comes at the cost of reduced sensitivity. Because it only captures interactions associated with non-zero phase delays, it may fail to detect genuine physiological synchronization occurring at or near zero phase lag [14].

1.4.4 Phase lag index (PLI)

To further improve the sensitivity of IC, Stam et al. [11] introduced the phase lag index (PLI). This measure is based on the asymmetry of the distribution of instantaneous phase differences $\Delta\phi(t_k)$ between two signals. Specifically, it quantifies whether phase differences tend to occur preferentially in the interval $(-\pi, 0)$ or $(0, \pi)$, rather than being symmetrically distributed around $0 \bmod \pi$. Such asymmetry indicates the presence of a consistent phase lead or lag. Formally, the PLI is defined as:

$$PLI = \left| \left\langle \text{sign}[\Delta\phi(t_k)] \right\rangle \right| \quad (1.6)$$

where the $\text{sign}[\cdot]$ function extracts only the direction of the phase difference while discarding its magnitude. The PLI ranges from 0 (no coupling or coupling with a phase difference centered around $0 \bmod \pi$) to 1 (perfect phase locking at a consistently non-zero phase lag) [11].

The authors reported that PLI was able to detect true synchronization and showed improved sensitivity in detecting subtle alterations of functional connectivity compared to IC, while being substantially less affected by spurious correlations due to common sources than PC. However, limitations remain. Like IC, the PLI may fail to detect genuine physiological interactions occurring at or near zero phase lag, thus providing only a partial description of functional connectivity. Additionally, the PLI discards the magnitude of the phase difference and retains only the absolute value of its sign, resulting in a loss of quantitative information about both the exact phase delay between signals and the direction of the phase lead or lag [11].

1.4.5 Weighted phase lag index (wPLI)

Vinck et al. [12] introduced the weighted phase lag index (wPLI) as an extension of the PLI: while the PLI only considers the sign of the imaginary component of the coherence, the wPLI weights each observation by the magnitude of $\Im\{X\}$:

$$\text{wPLI} = \frac{|\mathbb{E}[\Im\{X\}]|}{\mathbb{E}[|\Im\{X\}|]} = \frac{|\mathbb{E}[|\Im\{X\}| \text{sign}(\Im\{X\})]|}{\mathbb{E}[|\Im\{X\}|]} \quad (1.7)$$

where $\mathbb{E}[\cdot]$ denotes the expectation across observations. The wPLI ranges from 0 to 1, reaching its maximum when the imaginary components have a consistent sign across observations [12].

The rationale for this weighting is that cross-spectra located close to the real axis have small imaginary components and therefore provide weak evidence for a consistent non-zero phase lag. These observations are also more easily perturbed across the real axis by noise, which can turn a small phase lag into a small phase lead, or vice versa. In the PLI, such a perturbation produces a discontinuous change because only the sign is retained. In the wPLI, the same observation has little influence because its imaginary component is small. Conversely, observations with large imaginary components, corresponding to phase differences farther from 0° or 180° , receive stronger weight [12].

Like PLI and IC, wPLI reduces the contribution of zero-lag interactions, which are typical of common sources caused by volume conduction or active reference electrodes. Vinck et al. showed that, for correlated sources of interest, the value of the wPLI is unaffected by the volume conduction mixture coefficients, because these coefficients cancel out in the expression of the measure [12].

Test-retest reliability

Hardmeier et al. [15] assessed the long-term test-retest reliability of PLI- and wPLI-based functional connectivity and graph measures using 256-channel resting-state EEG in healthy subjects, with follow-up recordings after one and two years. Global wPLI showed good to excellent reliability across frequency bands, comparable to PLI. However, inter-subject variability was higher for wPLI than for PLI. Regional wPLI also showed good reliability for regional degree and fair to good reliability for inter-regional links [15].

Limitations

Despite its advantages, wPLI may fail to detect genuine zero-lag synchronization, because it is based exclusively on the imaginary component of the cross-spectrum. In this situation, measures such as coherence or PLV may be more sensitive, although they are also more susceptible to spurious synchronization caused by volume conduction or common reference effects [12]. This may also affect network analyses, since short-range interactions or local clustering can be underestimated when true interactions occur with very small phase delays [15].

Another practical limitation of several phase-synchronization measures is sample-size bias: when only a limited number of observations is available, random fluctuations can artificially inflate synchronization estimates, even in the absence of true coupling [12]. To address this issue, Vinck et al. analytically derived an unbiased estimator of the squared PLI and a debiased estimator of the squared wPLI, and showed that the debiased squared wPLI had greater statistical power than the unbiased squared PLI estimator [12]. Although the original wPLI formulation remains widely used in applied EEG studies, debiased implementations are also available in analysis toolboxes such as FieldTrip [16].

In addition, wPLI does not capture all types of nonlinear interactions. Imperatori et al. showed that wPLI performed poorly for some purely nonlinear dynamics with non-significant cross-correlation, supporting the use of complementary measures [17].

Finally, although wPLI reduces the influence of volume conduction on connectivity estimates, it does not fully resolve the problem of anatomical interpretation. True interactions between distant neural sources may still be recorded at scalp electrodes located elsewhere due to source mixing, meaning that sensor-level wPLI does not directly localize the underlying interacting brain regions [12].

1.4.6 Phase slope index (PSI)

A final phase-based measure worth mentioning is the phase slope index (PSI), introduced by Nolte et al. [18] as a method for estimating the direction of information flow between two signals. Unlike PLV, PLI, or wPLI, which quantify non-directed phase synchronization, PSI is specifically designed to infer directional interactions from multivariate time series.

The PSI is based on the observation that a consistent temporal delay between two interacting systems produces a frequency-dependent phase shift in the cross-spectrum. For a fixed time delay τ , the phase spectrum varies linearly with frequency:

$$\Phi(f) = 2\pi f\tau \quad (1.8)$$

such that the slope of the phase spectrum reflects the direction of the interaction. To estimate this effect robustly across frequencies, Nolte et al. defined the PSI as:

$$\tilde{\Psi}_{ij} = \Im \left(\sum_{f \in F} C_{ij}^*(f) C_{ij}(f + \delta f) \right) \quad (1.9)$$

where $C_{ij}(f)$ is the complex coherency between signals i and j , δf is the frequency resolution, and $\Im(\cdot)$ denotes the imaginary part [18]. For smooth phase spectra, this expression corresponds to a weighted average of the phase slope across frequencies.

A positive PSI indicates that signal i leads signal j , whereas a negative PSI indicates the opposite direction. Importantly, because the PSI vanishes when the imaginary part of coherency vanishes, the measure is robust against spurious directionality induced by source mixing or volume conduction [18].

Using simulations of coupled autoregressive systems, Nolte et al. showed that the PSI correctly detected the direction of unidirectional information flow while remaining largely insensitive to mixtures of independent noise sources, in contrast to Granger causality, which frequently yielded false-positive directional estimates under source mixing conditions [18].

1.5 VNS and brain connectivity

The hypothesis that VNS exerts its therapeutic effects through modulation of brain synchronization has motivated a growing body of studies examining EEG-derived functional connectivity measures in VNS-treated patients. Overall, the available literature suggests that VNS can alter EEG-derived connectivity, although the direction and spatial distribution of these changes appear to depend on the

frequency band, vigilance state, stimulation condition, clinical response, and the connectivity metric used.

Early evidence in humans was reported by Marrosu et al. [19], who compared EEG recordings obtained before and after VNS implantation in patients with pharmaco-resistant partial epilepsy. Using coherence and a nonlinear interdependence estimator, the authors observed decreased synchronization in the theta band together with increased gamma-band power and synchronization after VNS.

Several subsequent studies focused more specifically on phase-based synchronization measures. Bodin et al. [20] investigated interictal EEG synchronization using the PLI in patients receiving chronic VNS therapy. Responders showed lower global synchronization than non-responders, and VNS ON periods were associated with lower synchronization values than OFF periods, particularly in the delta and alpha bands. Similarly, Sangare et al. [21] reported lower PLI during ON compared with OFF periods in responders, mainly in the delta, theta, and beta bands, whereas non-responders did not show significant ON/OFF differences. In that study, seizure reduction correlated with the PLI OFF/ON ratio, and the theta-band findings were confirmed using PLV.

Lanzone et al. [22] also investigated short-term responses to VNS cycles using graph-theoretical measures derived from high-density EEG. Although spectral activity did not significantly differ between stimulation conditions, VNS cycles were associated with changes in small-world index, betweenness centrality, and global efficiency, particularly in slow frequency bands. Germany et al. [23] further reported dose-dependent EEG effects of VNS, showing that stimulation intensity influenced both EEG synchronization and power topography, with distinct patterns between responders and non-responders. Using stereotactic EEG recordings, Bartolomei et al. [24] observed increased connectivity during VNS ON periods in most patients, although the only responder of the cohort showed decreased connectivity. Together, these studies suggest that acute VNS stimulation can induce measurable modifications of synchronization and network organization, although the direction and magnitude of these effects appear to vary across patients and recording conditions.

Other studies investigated whether these effects depend on vigilance state. Vespa et al. [25] analyzed wakefulness and stage N2 sleep EEG using wPLI and graph-theoretical measures. In responders, stronger theta-band desynchronization and decreased global efficiency were observed during sleep but not during wakefulness, suggesting that sleep EEG may better reveal VNS-related connectivity changes. Similarly, Danthine et al. [26] compared pre-implantation wPLI during sleep and wakefulness in future responders and non-responders. Non-responders showed higher delta-band wPLI during wakefulness than during sleep, whereas lower alpha-band sleep wPLI was associated with greater seizure reduction after VNS. Both

studies therefore support the idea that vigilance state may substantially influence connectivity measures associated with VNS response.

Longitudinal studies have also examined connectivity modifications occurring over months of stimulation. Wang et al. [27] investigated patients with Dravet syndrome (a rare genetic disorder characterized by drug-resistant epilepsy) using mutual information and graph-theoretical measures before and after VNS implantation. The authors reported changes in characteristic path length, global efficiency, and transitivity, particularly in the alpha and beta bands, together with shifts in hub distribution in responders. Coa et al. [28] similarly observed longitudinal modifications of PLI after one year of VNS therapy. In responders, PLI decreased in the delta and gamma bands and increased in the alpha band, whereas non-responders showed no significant PLI modification. The same study also reported opposite longitudinal changes in aperiodic EEG parameters between responders and non-responders.

Connectivity measures have also been explored as predictive biomarkers before implantation. Germany et al. [29] analyzed preoperative EEG using the directed transfer function (DTF) and partial directed coherence (PDC), combined with graph-theoretical metrics including global efficiency, clustering coefficient, modularity, and global reaching centrality. Significant differences between responders and non-responders were found for several graph metrics, suggesting that pre-implantation network organization may help discriminate future responders.

Similar predictive approaches were explored in pediatric populations. Ma et al. [30] combined clinical variables with synchronization measures including PLI, wPLI, and PLV. In their discovery cohort, high-beta PLI was significantly higher in responders. A support vector machine classifier combining clinical and synchronization features showed better performance than models based on one feature type alone, with an accuracy of 75.7% and an AUC of 0.766 during cross-validation, and an accuracy of 61.1% in the validation cohort. Chen et al. [31] used whole-brain functional connectivity estimated by partial correlation. The postcentral gyrus in bilateral parietal regions was identified as the most significant region differentiating responders from non-responders. Their support vector machine classifier reached a mean AUC of 0.88 in cross-validation and 0.91 in the testing cohort, with patient-level prediction accuracy reported as 100%.

Ding et al. [32] specifically investigated ictal EEG recordings obtained before and after VNS implantation using wPLI and graph-theoretical analysis. After VNS treatment, whole-band wPLI, degree, clustering coefficient, and global efficiency increased, whereas characteristic path length decreased. Frequency-specific analyses showed increased connectivity between parieto-occipital and frontal regions in the delta and beta bands, but reduced connectivity within frontal regions and between

frontal and parieto-occipital regions in the beta and gamma bands.

Finally, several animal studies also reported frequency-dependent modifications of neural synchronization after VNS. Kano et al. [33] described increased PLV, particularly in the gamma band, in rat temporal cortex after VNS. Using rat models in both normal and epileptic conditions, Usami et al. [34] observed increased high-gamma synchronization together with decreased delta- and low-beta synchronization in the epileptic state. Nichols et al. [35] further showed in rat auditory cortex that VNS increased and decorrelated spontaneous multi-unit activity while reducing entrainment to repetitive stimulation at 6–8 Hz.

Taken together, these studies suggest that VNS is associated with measurable modifications of brain functional connectivity and network organization. Several studies reported lower synchronization during active stimulation periods or, more broadly, in responders independently of stimulation state [20, 21, 25, 36], whereas others described more heterogeneous patterns involving both increases and decreases in connectivity depending on the analyzed frequency bands, regions, and recording conditions [19, 28, 32]. Overall, the available literature suggests that VNS-related connectivity changes are likely frequency-dependent and state-dependent, supporting the relevance of EEG connectivity analyses as potential biomarkers of VNS effects and response.

1.6 Hyperventilation

Hyperventilation (HV) is a classical EEG activation procedure routinely used in clinical neurophysiology, particularly in the evaluation of epilepsy. It consists of deep and regular breathing, typically at a rate of approximately 20 breaths per minute for several minutes, resulting in hypocapnia and respiratory alkalosis. HV is known to induce diffuse EEG slowing and may facilitate epileptiform abnormalities, especially in generalized epilepsies, although its mechanisms and effects in focal epilepsy remain incompletely understood [37, 38].

Mazzucchi et al. [37] were the first to specifically investigate the effects of HV on functional connectivity and brain network organization in epilepsy. Using scalp EEG in patients with focal epilepsy and healthy controls, the authors estimated FC using lagged coherence and analyzed network topology using graph-theoretical metrics. HV significantly increased lagged coherence in both groups, although the effect was substantially more widespread in epileptic patients: in healthy subjects, increased connectivity was mainly restricted to the theta band, whereas patients with focal epilepsy showed increased connectivity in the delta, theta, alpha, and beta bands. In the theta band, epileptic patients exhibited a particularly diffuse

increase in connectivity between cortical regions. These modifications tended to return toward baseline after the end of HV.

The same study also showed significant effects of HV on graph-theoretical metrics derived from whole-brain connectivity networks. Mean lagged coherence and clustering coefficient increased during HV, whereas path length decreased, suggesting a transient increase in global connectedness and network efficiency during the activation procedure [37]. These effects were observed in both groups but were generally more pronounced in epileptic patients.

More recently, Ferri et al. [38] investigated cortical responses to HV using stereo-EEG (SEEG) recordings in patients with focal epilepsy. Patients underwent HV sessions during SEEG monitoring, and SEEG activity recorded during HV was compared with awake, sleep, and ictal-transition periods. The study showed that HV induced connectivity patterns that partially resembled those observed during ictal transition states, although less prominently, suggesting that HV can transiently shift cortical dynamics toward more seizure-like configurations [38].

Brázdil et al. [39] used pre-VNS implantation scalp EEG recordings acquired during resting conditions, intermittent photic stimulation (IPS), and HV to compare responders and non-responders based on EEG spectral reactivity. Unlike functional connectivity measures, spectral reactivity reflects changes in oscillatory activity power rather than statistical interactions between brain regions. The authors observed significant differences in alpha and gamma dynamics during both HV and IPS. In responders, HV induced marked increases in alpha and gamma activity, whereas non-responders showed persistent alpha desynchronization throughout the recording session. Similarly, responders showed stronger gamma increases during IPS. The authors suggested that these differences might reflect distinct properties of neuronal excitability and thalamocortical or corticocortical processing in responders and non-responders [39].

To date, no study has specifically investigated functional connectivity during HV in the context of VNS. Nevertheless, the available studies suggest that HV induces measurable and frequency-dependent modifications of cortical dynamics and connectivity in epilepsy [37, 38], while EEG reactivity to HV may also differ according to VNS response status [39]. Together, these observations support the relevance of investigating HV-related EEG synchronization in patients treated with VNS.

1.7 Objectives of the study

Building on the evidence reviewed above, two observations motivate the present work. First, VNS-related connectivity changes remain incompletely characterized, and no reliable preimplantation biomarker currently exists to predict patient response. Second, while hyperventilation is known to transiently modify cortical synchronization and network organization, its potential to reveal connectivity patterns relevant to VNS response has never been explored.

The present work therefore aims to investigate EEG functional connectivity during hyperventilation in patients undergoing VNS therapy for drug-resistant epilepsy, and to compare these findings with resting-state recordings. Functional connectivity was assessed using the weighted phase lag index (wPLI), and analyses were performed using both conventional statistical approaches and supervised machine learning methods.

More specifically, this study addresses three main questions: whether preimplantation connectivity measures may predict clinical response to VNS, whether VNS induces longitudinal modifications of brain connectivity over time, and whether acute activation of the stimulator produces measurable short-term connectivity changes.

Chapter 2

Material and methods

2.1 Study population and data acquisition

2.1.1 Study population

The present work is based on the same prospective longitudinal cohort previously described in earlier investigations of this database, which explored resting-state EEG synchronization and intermittent photic stimulation related connectivity changes following VNS implantation [36, 40]. In the current thesis, analyses focused specifically on hyperventilation-induced functional connectivity changes.

Twelve adults with drug-resistant epilepsy (DRE) were prospectively recruited at University Hospital Saint-Luc in Brussels between March 2022 and January 2024. Inclusion criteria comprised: (i) a diagnosis of DRE, (ii) planned cervical implantation of a VNS device, (iii) age between 18 and 70 years, and (iv) an intelligence quotient (IQ) above 55 on the Wechsler scale. Patients with previous neuromodulation therapy, neurodegenerative disease, or dementia were excluded from participation [36].

The study was approved by the local ethics committee of Cliniques Universitaires Saint-Luc/UC Louvain (Approval No. 2022/10JAN/009) and conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to inclusion [36].

2.1.2 Study design and EEG acquisition

Patients underwent longitudinal EEG assessments before and after VNS implantation. Baseline recordings were acquired prior to surgery (V1), followed by post-implantation evaluations at approximately one month (V3), three months (V4), and six months (V5). The implantation procedure itself corresponded to visit V2.

EEG recordings were obtained using a 64-channel Biosemi ActiveTwo acquisition system (Biosemi B.V., Amsterdam, The Netherlands) arranged according to the international 10–20 system, with a sampling frequency of 1024 Hz. Additional electrodes were used to simultaneously record electrocardiographic activity and stimulation-related artifacts, including two electrodes positioned bilaterally at the level of the laryngeal prominence to identify VNS ON periods [36].

Each session included several standard activation procedures routinely used in clinical EEG recordings, namely resting state, hyperventilation, and intermittent photic stimulation (IPS). Resting-state EEG was recorded during alternating eyes-open and eyes-closed periods. Hyperventilation constituted the principal condition of interest in the present study and was performed according to standard clinical procedures and exclusively with eyes open because visual cues were continuously provided to guide the breathing procedure. IPS paradigms were also acquired at multiple stimulation frequencies, as described by Cakiroglu et al. [40]. Because IPS recordings from this cohort are addressed in dedicated work, detailed analyses of this condition are not developed further in the present thesis.

Post-implantation recordings were acquired during alternating VNS ON and VNS OFF periods, allowing assessment of both acute and longitudinal neuromodulatory effects. VNS stimulation parameters were progressively increased during follow-up according to standard clinical titration protocols, with stimulation intensity gradually raised up to a target amplitude of approximately 1.75 mA when tolerated by the patient [36].

2.1.3 EEG preprocessing

EEG preprocessing was performed in MATLAB R2021a (MathWorks, Natick, MA, USA) using custom scripts and the Letswave 6 toolbox (UCLouvain, Brussels) [36, 40].

Signals were re-referenced to the common average reference across all scalp electrodes to reduce regional reference bias. A second-order Butterworth band-pass filter was applied between 0.5 and 40 Hz. Continuous EEG recordings were segmented

into epochs corresponding to each experimental condition. Hyperventilation and intermittent photic stimulation recordings were segmented into 8-second epochs based on trigger onset, whereas resting-state recordings consisted of repeated 10-second eyes-open and eyes-closed periods [40].

Artifact rejection was performed semi-automatically using independent component analysis (ICA) combined with visual inspection [36, 40]. Epochs containing ocular artifacts, movement artifacts, excessive amplitudes, or major interictal epileptiform discharges were excluded from subsequent analyses. In addition, EEG segments exceeding 500 μV were rejected [36]. Final artifact rejection was reviewed visually to ensure adequate data quality.

2.1.4 Clinical response evaluation

Clinical efficacy of VNS was evaluated using three complementary approaches [36]. First, patients were classified using the conventional binary responder status: patients with a seizure frequency reduction of at least 50% at six months post-implantation were considered responders (R), whereas patients with a reduction below 50% were classified as non-responders (NR).

Second, VNS response was also expressed as a continuous variable, corresponding to the percentage of seizure frequency reduction compared with the pre-implantation baseline. This allowed the degree of clinical improvement to be considered without imposing a categorical threshold.

Third, response was assessed using the Clinical-Research Response Scale (CRRS), also referred to as the NOVA scale, previously developed and validated in this cohort [36]. This multidimensional score combines chronic and acute effects of VNS. The chronic component accounts for seizure frequency reduction as well as changes in seizure duration, seizure intensity, and postictal recovery. The acute component evaluates the benefit of magnet activation on seizure suppression, duration, intensity, and postictal recovery. The acute and chronic components are combined into a final score ranging from 0 to 15, with higher values indicating a greater clinical response. The scoring system is detailed in Table 2.1.

Antiseizure medication regimens remained stable throughout the study period whenever possible in order to minimize confounding effects related to medication changes [36].

Table 2.1: Clinical-Research Response Scale, as described by Danthine et al. [36].

Response type	VNS effect	Criterion	Score computation	Total score
Acute – magnet response				0–5
	Seizure-suppressing effect of magnet	> 50%	+5	
	Seizures not suppressed	< 50%	+4	
	Seizure duration ↘	Yes / No	If at least one “yes”: +1	
	Seizure intensity ↘	Yes / No		
	Postictal duration ↘	Yes / No		
		0%	+0	
	Seizure duration ↘	Yes / No	If “yes”: +1	
	Seizure intensity ↘	Yes / No	If “yes”: +1	
	Postictal duration ↘	Yes / No	If “yes”: +1	
Chronic				0–10
	Seizure frequency reduction	> 80%	+10	
		> 50%	+8	
	Seizure duration ↘	= 50%	+7	
	Seizure intensity ↘	> 30%	+6	
	Postictal duration ↘	= 30%	+3	
		< 30%	+0	
	Seizure duration ↘	Yes / No	If “yes”: +1	
	Seizure intensity ↘	Yes / No	If “yes”: +1	
	Postictal duration ↘	Yes / No	If “yes”: +1	
Total				0–15

Acute response: Patients rated how often the magnet stopped seizures and described its effects on seizure duration, intensity, and postictal duration, scoring +1 for each observed effect if efficacy was < 50%.

Chronic response: Patients reported weekly seizure frequency changes since the last visit, categorized by reductions of > 80% (+10), > 50% (+8), and > 30% (+6). For 50% or 30% reductions, scores were adjusted to +7 or +3, respectively. Additional effects on seizure duration, intensity, and postictal duration were recorded for reductions < 80%, with +1 for each observed effect if the reduction is < 30%.

2.1.5 Functional connectivity analysis

Functional connectivity was assessed using the weighted Phase Lag Index (wPLI). As detailed in the first chapter of this thesis, wPLI was specifically developed to address several limitations of previously proposed phase-synchronization metrics, particularly their sensitivity to volume conduction, common sources, and noise perturbations close to the real axis [12]. Because of these properties, wPLI has been widely used in EEG functional connectivity studies and was selected as the main connectivity metric in the present work.

The wPLI was computed using custom MATLAB scripts following the methodology applied in previous studies on this cohort [36, 40]. Computations were performed over successive 1-second windows and subsequently averaged within each recording segment [36].

The wPLI was computed between all pairs of EEG electrodes, resulting in 64×64 connectivity matrices describing synchronization strength across the scalp. Analyses

were performed in seven frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), low beta (13–20 Hz), high beta (20–30 Hz), and broadband activity.

Connectivity measures were subsequently averaged across 15 predefined anatomical regions: whole brain, right hemisphere, left hemisphere, frontal lobe, right frontal lobe, left frontal lobe, parietal lobe, right parietal lobe, left parietal lobe, occipital lobe, right occipital lobe, left occipital lobe, temporal lobe, right temporal lobe, and left temporal lobe. For each combination of experimental state (resting state, hyperventilation, or IPS), eye condition (eyes open or eyes closed), and stimulation condition (VNS ON or VNS OFF), wPLI values were computed in the seven frequency bands across the 15 anatomical regions, resulting in a total of 105 regional connectivity features per recording condition and per subject.

In previous investigations performed on this cohort, analyses focused mainly on global synchronization during resting state and occipital beta-band synchronization during IPS [36, 40]. To ensure methodological consistency across studies, wPLI computation for resting-state, IPS, and hyperventilation conditions was performed once using a common processing framework by the same investigator. The present thesis subsequently applied a dedicated statistical and analytical pipeline specifically designed to investigate hyperventilation-related functional connectivity changes.

2.2 Statistical analysis

All statistical analyses and graphical visualizations were performed in R (R Foundation for Statistical Computing, Vienna, Austria) using a custom analysis pipeline. The following packages were primarily used throughout the analyses: `tidyverse` for data wrangling and visualization, `ggplot2` and `patchwork` for graphical outputs, `lme4` and `lmerTest` for mixed-effects modeling, and `rstatix` for non-parametric statistical testing [41].

Analyses were performed for all experimental conditions acquired during the protocol, including resting-state recordings, hyperventilation, and intermittent photic stimulation, and using all investigated response scales (binary response, seizure reduction percentage, and clinical-research response scale [CRRS]). IPS recordings from this cohort are addressed in dedicated work and are therefore not developed further in the present thesis.

2.2.1 Cross-sectional analyses at baseline

Cross-sectional analyses at baseline (V1) were performed to investigate whether pre-implantation wPLI values were associated with subsequent clinical response to VNS.

Prior to inferential analyses, distribution normality was assessed using the Shapiro–Wilk test, following previous analyses on the same cohort [36]. Normality was evaluated independently for each frequency band, brain region, and experimental condition. Because the distributions of many baseline wPLI variables did not satisfy normality assumptions and the cohort size was limited, associations between baseline connectivity values and binary clinical response were assessed using the Wilcoxon rank-sum test (Mann–Whitney U test equivalent in R). This non-parametric approach was selected because it does not require normally distributed data and is appropriate for comparing two independent groups when parametric assumptions are not satisfied [42]. For each brain region and frequency band, wPLI values were compared independently between responder and non-responder groups within each experimental condition and exact p-values were computed.

Associations between baseline connectivity measures and longitudinal or ordinal clinical outcomes were additionally explored using univariate linear regression models implemented with the `lm()` function in R. For each region-band pair, seizure reduction percentage or clinical response score was entered as dependent variable and baseline wPLI value as predictor. Regression coefficients (β), coefficients of determination (R^2), and associated p-values were extracted.

For regression analyses, residual distributions were additionally evaluated because assumptions for parametric inference concern model residuals rather than the original variables. Residual normality was assessed using Shapiro–Wilk tests and representative quantile–quantile plots. No systematic evidence of residual non-normality was observed across fitted models, supporting the use of parametric approaches. Therefore, correlation analyses were performed using Pearson’s correlation coefficient (r) [43, 44].

To account for multiple comparisons, false discovery rate (FDR) correction was applied using the Benjamini-Hochberg procedure [45]. In contrast to more conservative family-wise error rate approaches such as Bonferroni correction, FDR control aims to limit the expected proportion of false positive findings among the rejected hypotheses while preserving greater statistical power in exploratory analyses involving numerous simultaneous tests [45, 46]. This approach was considered particularly appropriate given the large number of region-band comparisons and the exploratory nature of the present analyses. Corrections were performed independently within each frequency band across all tested brain regions and experimental conditions.

Both raw and FDR-corrected p -values were reported. Statistical significance was defined as $p < 0.05$ after correction.

An additional exploratory trend analysis was performed to investigate whether the direction of associations between baseline connectivity and clinical outcomes showed consistent patterns across frequency bands and brain regions. For each region-band pair, the slope of the fitted linear relationship between wPLI values and clinical outcomes was extracted. The overall directionality of these slopes was subsequently summarized across features using a sign-based analysis to evaluate whether positive and negative associations occurred with similar frequency.

2.2.2 Longitudinal analyses

Longitudinal changes in functional connectivity following VNS implantation were investigated using linear mixed-effects models (LMMs). Mixed-effects modeling was selected to account for the repeated-measures structure of the data and the intra-subject correlation induced by multiple observations originating from the same patient [36, 40]. Separate models were fitted independently for each experimental condition, frequency band, and brain region, using wPLI as the dependent variable.

First, global longitudinal changes following VNS implantation were assessed independently of clinical response. For this analysis, visit was included as a fixed effect and patient identity as a random intercept:

$$\text{wPLI} \sim \text{Visit} + (1 \mid \text{Patient_ID}) \quad (2.1)$$

In this model, V1 served as the reference visit, and the V5 fixed effect estimated the average change in wPLI between baseline and six-month follow-up across all patients.

Second, response-dependent longitudinal changes were evaluated by including the response variable of interest, visit, and their interaction as fixed effects, with patient identity again included as a random intercept:

$$\text{wPLI} \sim \text{Response} * \text{Visit} + (1 \mid \text{Patient_ID}) \quad (2.2)$$

In these models, the $\text{response} \times \text{V5}$ interaction term assessed whether the change in wPLI between baseline and six-month follow-up differed according to clinical response. For analyses using the binary response scale, non-responders served as the reference category. For analyses using seizure reduction percentage and CRRS as continuous predictors, a value of 0 represented the reference response value.

Models were implemented using the `lmer()` function from the `lme4` package in R. Fixed-effects significance testing was performed using the `lmerTest` package, which computes p-values using Satterthwaite approximations for denominator degrees of freedom [47]. To improve model convergence given the relatively small sample size and repeated-measures structure, models were optimized using the `bobyqa` optimizer with an increased maximum number of function evaluations (`maxfun = 2e5`). Potential singular model fits were assessed using the `isSingular()` function from the `lme4` package with a tolerance threshold of 10^{-4} .

Because multiple tests were performed across brain regions, FDR correction was applied using the Benjamini-Hochberg procedure. Correction was performed separately for each frequency band and each fixed-effect term, across the 15 predefined brain regions. Corrected p-values below $\alpha = 0.05$ were considered statistically significant. Fixed-effect estimates and associated FDR-corrected p-values were extracted for each frequency band and brain region.

Covariate analyses

Additional exploratory mixed-effects analyses were performed to assess the potential influence of demographic and clinical variables on longitudinal connectivity findings.

The following covariates were individually introduced as additional fixed effects: age, sex, epilepsy type, epilepsy duration, IQ (categorized as < 70 and ≥ 70), number of antiseizure medications, and individual antiseizure medication treatments. Each covariate was independently added to the base model according to:

$$\text{wPLI} \sim \text{Response} * \text{Visit} + \text{Covariate} + (1 \mid \text{Patient_ID}) \quad (2.3)$$

This approach enabled exploratory evaluation of whether longitudinal findings remained consistent after adjustment for individual demographic and clinical variables.

2.2.3 Acute stimulation ON/OFF analyses

Additional analyses were performed to investigate the acute effect of VNS stimulation on functional connectivity at the six-month follow-up visit (V5). For these analyses, EEG recordings acquired with stimulation switched OFF and ON were directly compared within subjects. Two populations were evaluated independently: the complete cohort and the subgroup of clinical responders.

For each frequency band and brain region, wPLI values were reshaped into paired

observations at the patient level and compared using paired Wilcoxon signed-rank tests. This non-parametric approach was selected because observations were paired within subjects and because distributional assumptions for parametric paired tests could not be guaranteed [42]. FDR correction was applied independently within each frequency band across all investigated regions to account for multiple testing.

2.3 Machine Learning Pipeline

Machine learning analyses were performed in Python using a custom analysis pipeline based on the `scikit-learn` ecosystem [48, 49]. All analyses were restricted to baseline recordings obtained before VNS implantation (V1).

2.3.1 Feature selection

Because the cohort included only 11 patients whereas the complete connectivity dataset contained up to 315 wPLI variables per subject (7 frequency bands $\times$ 15 brain regions $\times$ 3 experimental conditions, corresponding to resting-state eyes closed, resting-state eyes open, and hyperventilation eyes open), several reduced feature sets were constructed prior to model training in order to limit dimensionality and reduce the risk of overfitting. These feature sets were specifically designed to: (i) compare resting-state and hyperventilation-derived connectivity, (ii) compare the predictive contribution of different frequency bands, and (iii) evaluate the added value of regional connectivity measures relative to whole-brain connectivity alone.

Because connectivity was computed separately for eyes-closed and eyes-open resting-state recordings, but only during eyes-open recordings for hyperventilation, resting-state datasets systematically contained twice as many condition-specific features as hyperventilation datasets.

Several categories of datasets were generated. First, datasets including only global wPLI values across all frequency bands were constructed. Second, frequency-specific datasets were generated separately for each frequency band (delta, theta, alpha, beta, low-beta, high-beta, and broadband). For each frequency band, two versions were created: one including only the global wPLI value(s) and another including regional connectivity values from the 15 anatomical regions. Finally, an exploratory feature set combining delta-band resting-state global connectivity with beta-band hyperventilation global connectivity was constructed. This dataset was motivated by the results of the previous statistical analyses, which identified delta-band resting-state connectivity and beta-band hyperventilation connectivity as the most discriminative features associated with clinical outcome. Table 2.2 summarizes the

forty-six resulting datasets and the corresponding number of features included in each.

Table 2.2: Summary of the 46 machine learning datasets constructed from baseline wPLI connectivity measures.

Dataset category (and number)	Dataset type	Bands	Regions	Features
Global, all bands (3)	RS	7	1	14
	HV	7	1	7
	RS + HV	7	1	21
Regional, frequency-specific (21)	RS (1 per band)	1	15	30
	HV (1 per band)	1	15	15
	RS + HV (1 per band)	1	15	45
Global, frequency-specific (21)	RS (1 per band)	1	1	2
	HV (1 per band)	1	1	1
	RS + HV (1 per band)	1	1	3
Special dataset (1)	RS + HV	2	1	3

2.3.2 Regression models

Several regression algorithms were evaluated to predict the CRRS from baseline wPLI connectivity features. CRRS was selected as the primary prediction target instead of percentage seizure reduction or binary response because the preceding statistical analyses demonstrated stronger and more consistent associations with CRRS. Given the limited sample size and the relatively high dimensionality of several feature sets, the analyses focused exclusively on linear and regularized linear models in order to reduce model complexity and limit overfitting.

For all models, predictions were expressed as a linear combination of the input features:

$$\hat{y} = Xw + b, \quad (2.4)$$

where X is the feature matrix, w is the vector of regression coefficients, and b is the intercept. The following regression models were evaluated [48]:

- **LinearRegression()**. Ordinary least-squares (OLS) linear regression was used as an unregularized baseline model; it estimates regression coefficients by minimizing the residual sum of squares between observed and predicted values:

$$\min_w ||Xw - y||_2^2. \quad (2.5)$$

- **Ridge()**. Ridge regression extends OLS by introducing an L2 regularization penalty on regression coefficients, reducing coefficient magnitude and limiting

overfitting in the presence of correlated predictors:

$$\min_w ||Xw - y||_2^2 + \alpha ||w||_2^2, \quad (2.6)$$

where α controls the strength of coefficient shrinkage.

- **Lasso()**. Lasso regression applies an L1 regularization penalty that can shrink some coefficients exactly to zero, thereby performing implicit feature selection:

$$\min_w \frac{1}{2n_{\text{samples}}} ||Xw - y||_2^2 + \alpha ||w||_1. \quad (2.7)$$

- **ElasticNet()**. Elastic net regression combines L1 and L2 regularization penalties, allowing simultaneous sparse feature selection and coefficient stabilization in the presence of correlated variables:

$$\min_w \frac{1}{2n_{\text{samples}}} ||Xw - y||_2^2 + \alpha \rho ||w||_1 + \frac{\alpha(1 - \rho)}{2} ||w||_2^2, \quad (2.8)$$

where ρ corresponds to the `l1_ratio` parameter and controls the balance between L1 and L2 penalties.

- **QuantileRegressor()**. Quantile regression was implemented with `quantile = 0.5`, corresponding to least absolute deviations (LAD) or median regression. Compared with squared-error regression, this approach is less sensitive to large residuals because it estimates the conditional median rather than the conditional mean. The model minimizes the pinball loss:

$$\min_w \frac{1}{n_{\text{samples}}} \sum_i PB_q(y_i - X_i w) + \alpha ||w||_1, \quad (2.9)$$

with

$$PB_q(t) = \begin{cases} 0.5t, & t > 0, \\ 0, & t = 0, \\ -0.5t, & t < 0. \end{cases} \quad (2.10)$$

- **HuberRegressor()**. Huber regression was also included as a robust regression model. It combines squared and absolute error minimization through the Huber loss function:

$$\min_{w, \sigma} \sum_{i=1}^n \left(\sigma + H_\epsilon \left(\frac{X_i w - y_i}{\sigma} \right) \sigma \right) + \alpha ||w||_2^2, \quad (2.11)$$

where the Huber loss is defined as:

$$H_{\epsilon}(z) = \begin{cases} z^2, & \text{if } |z| < \epsilon, \\ 2\epsilon|z| - \epsilon^2, & \text{otherwise.} \end{cases} \quad (2.12)$$

This loss behaves quadratically for small residuals and linearly for large residuals, thereby reducing the influence of outlying observations.

Hyperparameter optimization was performed independently within each training fold using grid-search procedures integrated within the nested cross-validation framework. Hyperparameter grids were intentionally restricted in order to limit model complexity given the very small sample size. The grid of hyperparameters evaluated for each regression model is summarized in Table 2.3.

Table 2.3: Hyperparameter grids evaluated for regression models.

Model	Hyperparameter(s)	Tested values
Linear regression	None	-
Ridge regression	alpha	0.01, 0.1, 1, 10, 100
Lasso regression	alpha	0.01, 0.1, 1, 10
Elastic Net regression	alpha; l1_ratio	0.01, 0.1, 1, 10 0.2, 0.5, 0.8
LAD regression	alpha	0.01, 0.1, 1, 10
Huber regression	alpha; epsilon	0.01, 0.1, 1, 10 1.35, 1.5, 2.0

2.3.3 Nested cross-validation

For all models, preprocessing steps were incorporated within `scikit-learn` pipelines to prevent data leakage during cross-validation. Continuous predictors were standardized using z-score normalization with `StandardScaler`. In accordance with the pipeline framework, scaling parameters were estimated separately within each training set and subsequently applied to the corresponding test data, ensuring that no information from test observations contributed to preprocessing.

Given the very small sample size, model evaluation was performed using nested cross-validation in order to separate hyperparameter optimization from final performance estimation. This approach reduces the risk of overly optimistic performance estimates that may occur when hyperparameter tuning and model evaluation are performed on the same data.

The outer cross-validation loop used leave-one-out cross-validation (LOOCV), such that one subject was left out for testing at each iteration while the remaining

subjects were used for training. Within each outer training set, a second inner LOOCV loop was used for hyperparameter selection through exhaustive grid search. Hyperparameter optimization was performed using mean absolute error (MAE) as the scoring metric.

For each outer fold, the model associated with the best-performing hyperparameter configuration identified in the inner loop was then applied to the held-out test subject. Because CRRS values are bounded between 0 and 15, predicted scores were clipped to the valid score range after prediction:

$$\hat{y}_{\text{clip}} = \min(15, \max(0, \hat{y})) . \quad (2.13)$$

For each outer fold, the optimal hyperparameter configuration selected during the inner loop was stored. A final consensus hyperparameter set was subsequently defined for each model as the configuration most frequently selected across outer folds.

2.3.4 Performance metrics

Model performance was primarily evaluated using mean absolute error (MAE) [48]:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (2.14)$$

MAE was selected as the primary metric because it is directly interpretable on the CRRS scale and is less sensitive to extreme prediction errors than squared-error metrics. Root mean squared error (RMSE) was additionally computed as a complementary metric [48]:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (2.15)$$

To provide a baseline reference for regression performance, a `DummyRegressor` predicting the mean CRRS score of the training set was evaluated using the same outer LOOCV procedure. Regression models were therefore considered informative only if they achieved lower prediction errors than this baseline model.

2.3.5 Permutation testing

Permutation testing was used to assess whether the observed cross-validated model performance exceeded what would be expected under the null hypothesis of no

association between the input features and the target variable. Permutation tests constitute a nonparametric inference framework that relies on fewer distributional assumptions than classical parametric tests and are particularly useful in small-sample and high-dimensional settings [48, 50]. Under the null hypothesis, labels are assumed to be exchangeable, allowing the empirical null distribution of the test statistic to be estimated by repeatedly permuting the target labels [50].

This procedure was implemented using `permutation_test_score` from `scikit-learn`, which evaluates a cross-validated score on the original data and compares it with scores obtained after randomly permuting the target labels [48]. For each selected model, the target labels were randomly permuted while preserving the same leave-one-out cross-validation structure used in the outer evaluation loop, with the same scoring functions.

For computational feasibility, permutation testing was performed using the consensus hyperparameters identified during nested cross-validation rather than re-running the full nested hyperparameter search for each permutation. Thus, the permutation test assessed whether the selected model configuration performed better than expected by chance under the same outer cross-validation scheme. Predictions were clipped to the valid CRRS range $[0, 15]$ during permutation testing, ensuring consistency with the nested cross-validation procedure.

A total of 1000 permutations was used for each selected model-dataset combination. The permutation-derived p-value corresponds to the proportion of permuted datasets producing a model performance at least as extreme as the performance obtained with the true labels [50]. Because multiple models and feature sets were tested, p-values were corrected using the Benjamini–Hochberg false discovery rate (FDR) procedure [45]. The correction was applied across the full family of model-dataset permutation tests.

2.3.6 Model stability analysis

In addition to predictive performance, model stability was explicitly assessed. This step was included because the number of subjects was small relative to the number of candidate predictors, increasing the risk that apparently predictive models would depend strongly on the inclusion or exclusion of individual subjects.

Stability was evaluated using an additional leave-one-out procedure. In each fold, the model was refit on the training subjects using the consensus hyperparameters identified during nested cross-validation. Coefficient stability was assessed using three complementary quantities: selection frequency, sign consistency, and coefficient variability. Selection frequency was defined as the proportion of leave-one-out

folds in which a coefficient was non-zero. Sign consistency was defined as the proportion of folds in which the sign of a coefficient matched the median sign of that coefficient across folds. Coefficient variability was quantified as the ratio between the standard deviation and the absolute mean coefficient value across folds:

$$\text{Stability} = \frac{\sigma_{\text{coef}}}{|\mu_{\text{coef}}|}. \quad (2.16)$$

To summarize stability at the model level, only predictors selected in more than 50% of leave-one-out folds were retained for the computation of average coefficient stability and average sign consistency. This filtering step was introduced to avoid instability estimates being dominated by coefficients that appeared only sporadically across folds. The average coefficient stability and average sign consistency were then computed across these selected predictors. Regression models were retained as stable only if they satisfied three predefined criteria:

- Mean coefficient stability below 0.3;
- Mean sign consistency above 0.8;
- Standard deviation of leave-one-out MAE below 2.5.

These thresholds were chosen a priori as pragmatic filtering criteria for this exploratory small-sample analysis. They should therefore be interpreted as arbitrary decision thresholds rather than universally validated cut-offs.

2.3.7 Comparison of resting-state and hyperventilation feature sets

Additional analyses were performed to evaluate whether the inclusion of hyperventilation derived connectivity measures improved predictive performance compared with resting-state features alone. These comparisons aimed to determine whether hyperventilation provided complementary information beyond baseline resting-state connectivity.

Prediction performance was compared using subject-wise absolute prediction errors:

$$e_i = |y_i - \hat{y}_i|, \quad (2.17)$$

where y_i denotes the observed CRRS value for subject i , and $\hat{y}_i$ the corresponding model prediction. Differences between resting-state and combined resting-state

plus hyperventilation feature sets were evaluated using paired permutation tests. Statistical significance indicated whether adding hyperventilation features produced systematic improvements in prediction performance across subjects.

2.3.8 Learning curve analysis

A learning curve analysis was performed to evaluate how predictive performance evolved as the training sample size increased, and to determine whether model performance showed evidence of stabilization or continued improvement with additional data. In the context of machine learning, learning curves are commonly used to assess model generalization, estimate potential gains associated with larger datasets, and contextualize limitations related to small sample sizes [51, 52].

For each possible training set size, all combinations of training subjects were exhaustively generated. For every combination, the regression pipeline was trained on the selected training subset and evaluated on the remaining held-out subjects. Prediction performance was quantified using the mean absolute error (MAE), consistently with the primary evaluation metric used throughout the nested cross-validation analyses, and MAE values obtained were averaged to estimate the expected predictive performance at that sample size.

2.3.9 Power analysis

A power analysis was performed to estimate the sample size requirements of the regression models. The procedure followed Cohen’s framework for multiple regression and correlation analysis [53]. According to Cohen, statistical power analysis exploits the relationships among the significance criterion (α), effect size, statistical power ($1 - \beta$), sample size (N), and model parameters. For a given statistical model, these quantities are interrelated such that one parameter can be estimated from the others when the remaining values are specified [54].

In the multiple regression framework, significance testing is based on an F statistic evaluating whether the proportion of variance explained by a source differs from zero in the population [53]. Cohen expresses this statistic as:

$$F = \frac{PV_S}{PV_E} \times \frac{v}{u}, \quad (2.18)$$

where PV_S denotes the proportion of variance explained by the source of interest (denoted R^2 for a standard regression model), PV_E the residual variance proportion (corresponding to $1 - R^2$), u the numerator degrees of freedom corresponding to the

number of predictors, and v the denominator degrees of freedom. Cohen describes the first term as the effect-size index:

$$f^2 = \frac{R^2}{1 - R^2}. \quad (2.19)$$

and the second as reflecting experiment size and model complexity. For a model with u predictors and total sample size N , denominator degrees of freedom are:

$$v = N - u - 1. \quad (2.20)$$

Power calculations further depend on the noncentrality parameter:

$$\lambda = f^2(u + v + 1), \quad (2.21)$$

which simplifies to:

$$\lambda = f^2 N. \quad (2.22)$$

Cohen states that the noncentrality parameter absorbs information from effect size and model degrees of freedom, and that statistical power increases as a function of λ , v , and α , while decreasing with increasing numerator degrees of freedom u [53].

In the present work, Cohen’s original table-based procedure was replaced by direct numerical evaluation using the `scipy.stats` implementation of the central F distribution and the noncentral F distribution. The critical F value corresponding to the significance threshold α was obtained from the percent-point function of the central F distribution:

$$F_{\text{crit}} = F^{-1}(1 - \alpha; u, v), \quad (2.23)$$

implemented using `scipy.stats.f.ppf()`.

The noncentrality parameter was then computed directly from Cohen’s formulation (Eq. 2.22). Statistical power was evaluated from the cumulative distribution function of the noncentral F distribution parameterized by numerator degrees of freedom (u), denominator degrees of freedom (v), and noncentrality parameter (λ), implemented using `scipy.stats.ncf.cdf()` [55, 56]. For each candidate sample size, power was iteratively evaluated and the smallest sample size reaching the predefined power threshold was retained.

To account for uncertainty in the estimated predictive effect size, the observed coefficient of determination was complemented by a bootstrap confidence interval. At each bootstrap iteration, patients were resampled with replacement and the coefficient of determination was recomputed. The power analysis was therefore not based on a single observed value of R^2 , but on a range of plausible effect-size values derived from the bootstrap distribution. For each value of R^2 within this interval, the required sample size was estimated as the smallest N achieving the target power, given the chosen significance threshold α and the number of predictors included in the model. This produced a range of required sample sizes rather than a single point estimate. This approach was used to reflect the uncertainty surrounding the effect-size estimate in this small exploratory cohort.

Chapter 3

Results

3.1 Demographic and clinical characteristics

The present thesis is based on the same longitudinal cohort previously investigated in earlier studies focusing on resting-state EEG synchronization and IPS-induced connectivity changes after VNS implantation [36, 40].

Twelve patients with drug-resistant epilepsy were included in the study, comprising five men and seven women. Mean age at inclusion was approximately 38 years, and mean epilepsy duration exceeded 20 years. Both focal and generalized epilepsy syndromes were represented in the cohort, with structural, idiopathic, and unknown etiologies.

No participant withdrew during follow-up, and all twelve patients completed the longitudinal study. However, several EEG recordings were unavailable across visits and experimental conditions because of missed visits or incomplete acquisitions. The availability of recordings across patients, visits, and experimental conditions is summarized in Figure 3.1.

Antiseizure medication regimens remained generally stable throughout follow-up. VNS stimulation parameters were progressively titrated following standard clinical practice, and most patients reached target stimulation intensity within the first months after implantation.

Detailed demographic and clinical characteristics are summarized in Table 3.1.

Figure 3.1: Availability of EEG recordings across patients, longitudinal visits, and experimental conditions. Green cells indicate available recordings and gray cells indicate missing data. Columns represent recordings obtained at baseline before VNS implantation (V1), and at one-month (V3), three-month (V4), and six-month (V5) follow-up visits for resting-state (RS), intermittent photic stimulation (IPS), and hyperventilation (HV) conditions.

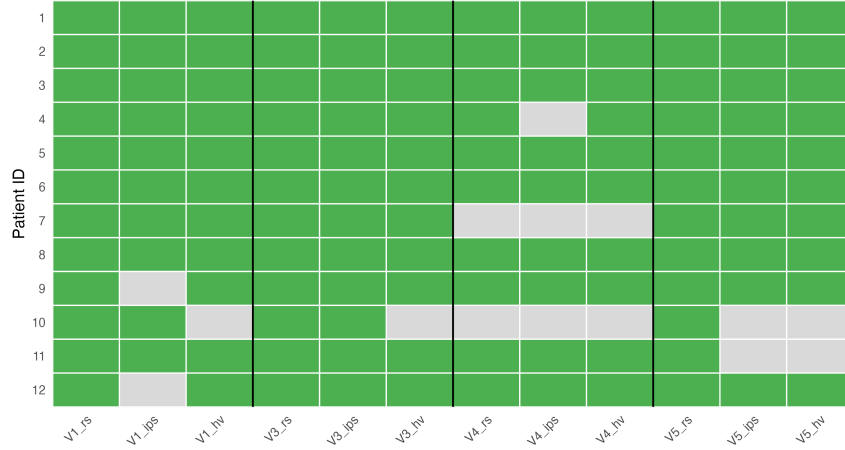


Table 3.1: Demographic and clinical characteristics of the study population, adapted from Danthine et al. [36]. ASM, antiseizure medication; BVR, brivaracetam; CBD, cannabidiol; CBM, cenobamate; CBZ, carbamazepine; F, female; LCM, lacosamide; LEV, levetiracetam; LTG, lamotrigine; M, male; OXC, oxcarbazepine; TPM, topiramate; VPA, sodium valproate.

ID	Age	Sex	Epilepsy type and etiology	Duration (years)	ASM	IQ
1	24	M	Focal epilepsy, structural etiology	16	LCM, BVR, CBD	>70
2	40	M	Focal epilepsy, structural etiology	27	LCM, BVR	>70
3	37	F	Focal epilepsy, unknown etiology	4	TPM, CBM, LTG	84
4	34	M	Adult Lennox–Gastaut syndrome	21	LEV, LTG, OXC	56
5	53	F	Generalized epilepsy, idiopathic	32	LCM, TPM	>70
6	66	F	Focal epilepsy, unknown etiology	58	LEV, LCM	100
7	41	F	Generalized epilepsy, idiopathic	34	VPA, LCM	84
8	25	F	Generalized epilepsy, idiopathic	14	LTG	72
9	28	M	Focal epilepsy, structural etiology	6	VPA, BRV	>70
10	47	F	Generalized epilepsy, idiopathic	33	LTG, VPA	84
11	33	F	Focal epilepsy, structural etiology	10	BVR, LCM	72
12	35	M	Focal epilepsy, unknown etiology	22	VPA, LTG, LCM	88

3.2 Clinical response to VNS

Clinical response at six months post-implantation was evaluated using the three predefined outcome measures described in Section 2.1.4: binary responder status, seizure frequency reduction percentage, and the Clinical-Research Response Scale

(CRRS).

Binary classification identified both responders and non-responders within the cohort, reflecting interindividual variability in clinical outcomes following VNS implantation. Seizure reduction percentages and CRRS scores likewise showed substantial variability across patients.

Complete clinical response data, including responder status, seizure reduction percentages, and CRRS scores, are summarized in Table 3.2.

Table 3.2: Clinical response at six months post-implantation, adapted from Danthine et al. [36].

ID	Binary response	Seizure reduction (%)	Chronic CRRS (0–10)	Acute CRRS (0–5)	Final CRRS (0–15)
1	R	80	10	5	15
2	R	55	8	5	13
3	NR	17	1	4	5
4	NR	14	0	1	1
5	R	94	10	5	15
6	NR	50	7	5	12
7	R	67	8	4	12
8	NR	13	2	5	7
9	R	60	8	1	9
10	NR	0	0	0	0
11	NR	46	6	5	11
12	NR	0	0	0	0

3.3 Predictive value of pre-implantation EEG recordings

Complete statistical results for the hyperventilation condition are reported in Appendix A.

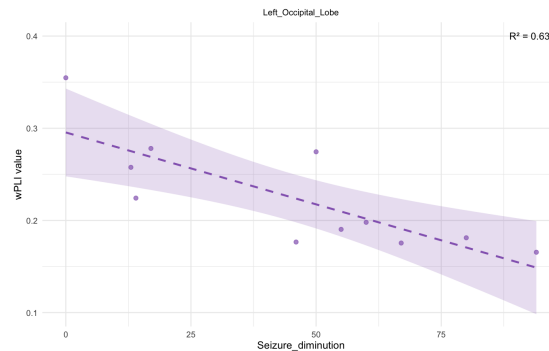
3.3.1 Binary response scale

No statistically significant association was identified between baseline wPLI values and binary VNS responder status after FDR correction. This absence of significant findings was consistent across all frequency bands, brain regions, and experimental conditions. Visual inspection of boxplot distributions did not reveal any clear separation between responders and non-responders.

3.3.2 Seizure reduction percentage

Among all evaluated frequency bands, brain regions, and experimental conditions, only one association remained statistically significant after FDR correction. During hyperventilation, theta-band connectivity within the left occipital lobe was significantly associated with seizure reduction percentage ($p_{FDR} = 0.048$). The corresponding linear model yielded an R^2 value of 0.637 (Figure 3.2).

Figure 3.2: Relationship between baseline theta-band left occipital wPLI values during hyperventilation at V1 and seizure reduction percentage following VNS implantation.



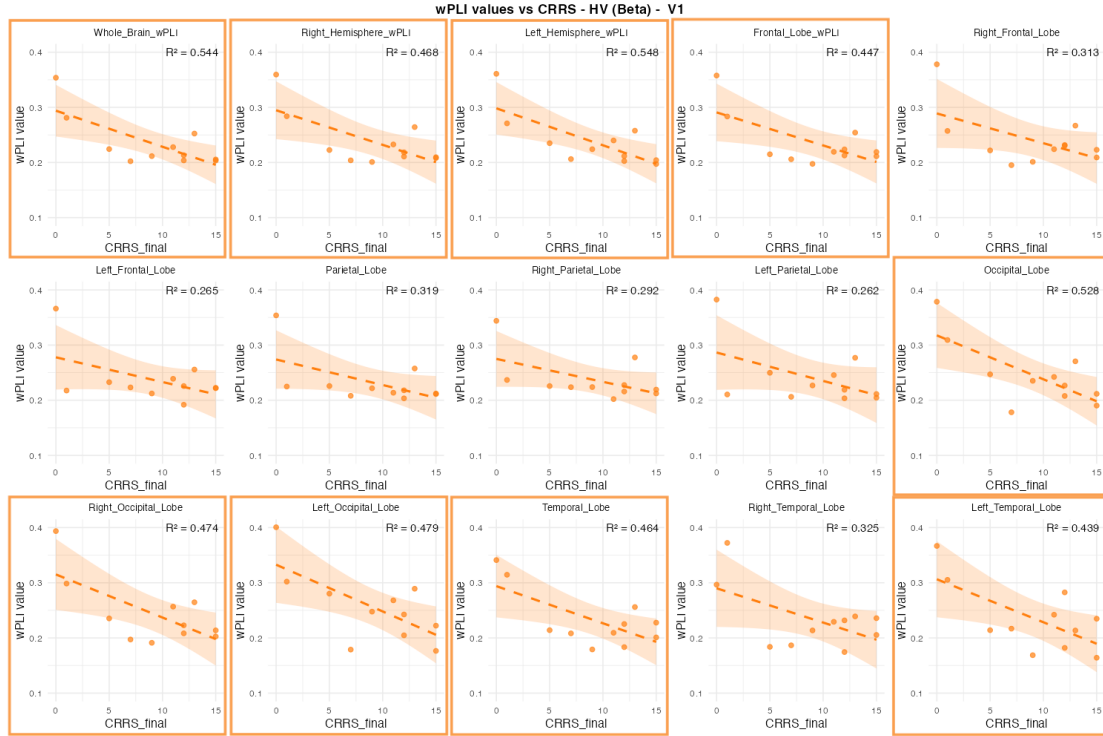
No additional associations remained significant after FDR correction.

3.3.3 CRRS

Several associations between baseline connectivity measures and CRRS remained significant after FDR correction. Within hyperventilation recordings, significant findings were primarily observed in the beta and low-beta frequency bands, with FDR-corrected p -values ranging from 0.037 to 0.049. These findings involved multiple connectivity measures, including whole-brain, hemispheric, occipital, temporal, and frontal regions. Figures 3.3 and 3.4 illustrate regional connectivity patterns within the beta and low-beta bands during hyperventilation.

In contrast, resting-state recordings yielded only two significant associations in the eyes-closed condition, both involving broadband connectivity within parietal regions, illustrated in Figure 3.5 ($p_{FDR} = 0.006$).

Figure 3.3: Relationship between baseline beta-band wPLI values during hyperventilation at V1 and CRRS following VNS implantation. Significant regions after FDR correction are framed.



3.3.4 Exploratory trend analysis

An exploratory analysis was performed to investigate whether the direction of associations between baseline connectivity and clinical outcomes showed consistent patterns across frequency bands and brain regions. For each feature, the slope of the linear relationship between baseline wPLI values and clinical outcomes was extracted, and the proportion of negative slopes was summarized across all evaluated features (Table 3.3).

Table 3.3: Number of negative regression slopes among all evaluated frequency band–region combinations ($n = 105$) for each experimental condition.

Outcome	HV	RS (EC)	RS (EO)
Seizure reduction (%)	99/105	34/105	79/105
CRRS	101/105	38/105	84/105

Hyperventilation showed a marked predominance of negative slopes across frequency bands and brain regions for both seizure reduction and CRRS outcomes.

Figure 3.4: Relationship between baseline low-beta-band wPLI values during hyperventilation at V1 and CRRS following VNS implantation. Significant regions after FDR correction are framed.

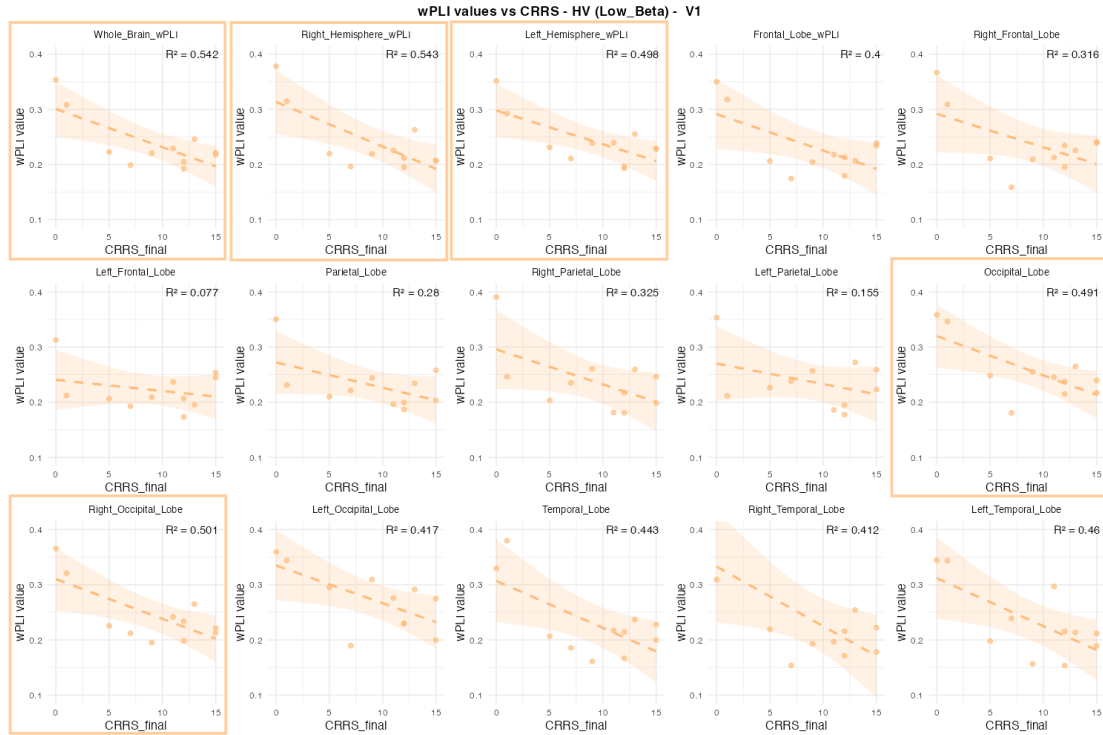
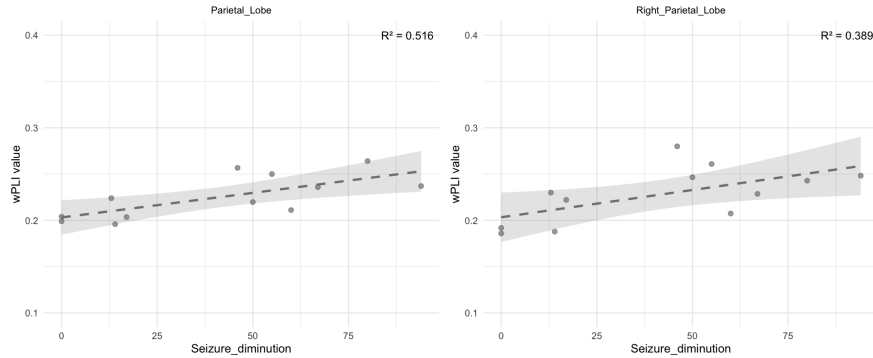


Figure 3.5: Relationship between baseline broadband parietal and right parietal wPLI values during resting-state recordings at V1 and CRRS following VNS implantation.



Resting-state recordings with eyes closed showed substantially fewer negative slopes, whereas the eyes-open resting-state condition demonstrated an intermediate pattern. Overall, these descriptive observations indicate that negative associations were more frequently observed during hyperventilation than during resting-state conditions. Because frequency bands and brain regions are not statistically independent, these

findings should be considered exploratory and interpreted cautiously.

3.3.5 Machine learning prediction of CRRS

The machine learning pipeline was applied to evaluate whether combinations of baseline wPLI features could predict CRRS following VNS implantation, beyond the predictive value of individual connectivity measures evaluated in the preceding statistical analyses. Among all tested model–dataset combinations, only three models satisfied all predefined selection criteria: prediction error lower than the dummy regressor baseline, FDR-corrected permutation-test significance ($p_{FDR} < 0.05$), and fulfilment of the predefined stability criteria. Table 3.4 summarizes the characteristics and performance metrics of the retained models.

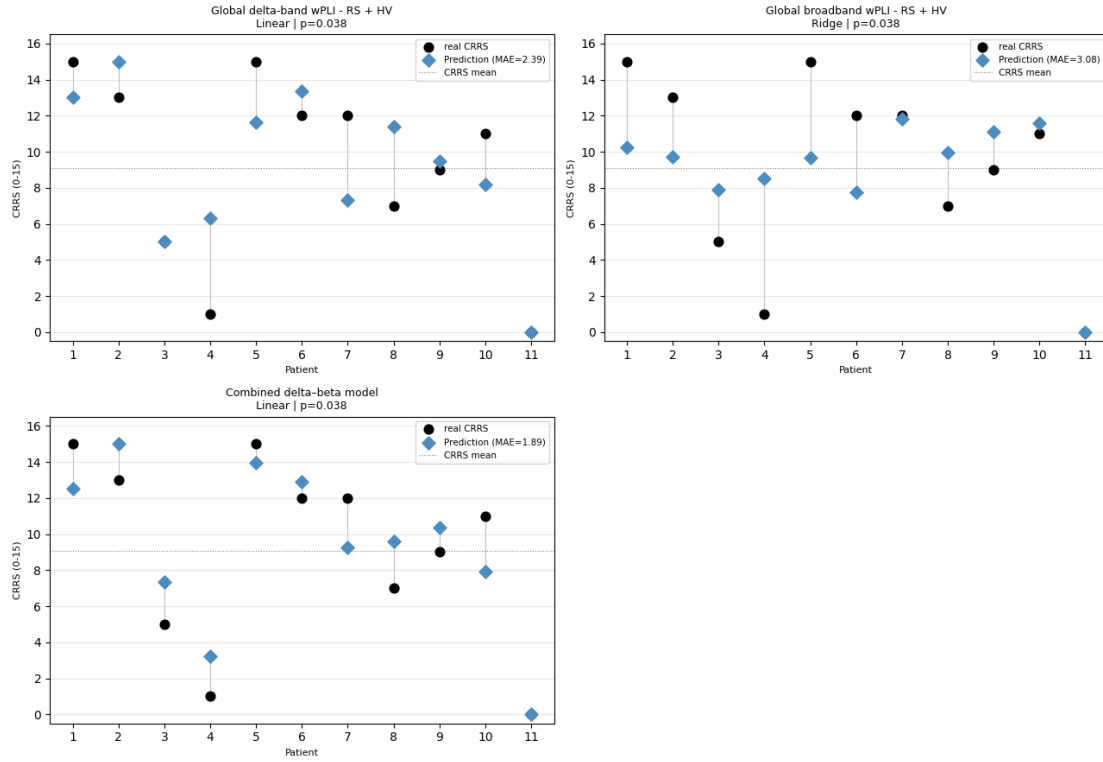
Table 3.4: Machine learning models retained after applying the predefined selection criteria. MAE SD corresponds to the standard deviation of leave-one-out prediction errors across folds. Stability quantifies coefficient variability and is defined as the ratio between the standard deviation and absolute mean of model coefficients across leave-one-out refits ($\sigma_{\text{coef}}/|\mu_{\text{coef}}|$), with lower values indicating greater coefficient stability. Sign consistency represents the proportion of folds in which coefficient signs remained unchanged. Further details regarding the stability analysis are provided in Section 2.3.6.

Dataset	Model	p_{FDR}	MAE	RMSE	MAE SD	Stability	Sign consistency
Global - delta - RS + HV	Linear	0.038	2.393	2.988	1.790	0.222	0.970
Global - broadband - RS + HV	Ridge	0.038	3.077	3.794	2.113	0.149	1.000
Global - delta RS + beta HV	Linear	0.038	1.887	2.089	0.895	0.289	1.000

All retained models were based exclusively on global wPLI values and incorporated features derived from both resting-state and hyperventilation recordings. The best-performing model was the exploratory dataset combining delta-band resting-state connectivity with beta-band hyperventilation connectivity (hereafter referred to as the combined delta–beta model), which achieved the lowest prediction error (MAE = 1.887; RMSE = 2.089). The two remaining retained models were based on global delta-band and broadband connectivity features, respectively.

Figure 3.6 presents the out-of-sample predictions generated during the outer leave-one-out cross-validation procedure for the three retained models. For each patient, the displayed prediction corresponds to the prediction obtained when that patient was excluded from model training and hyperparameter optimization. Consequently, these predictions represent cross-validated out-of-sample estimates rather than fitted values derived from models trained on the complete dataset, providing an estimate of model performance on unseen data. Visual inspection indicated that the combined delta–beta model more closely reproduced the observed CRRS values than the other retained models.

Figure 3.6: Observed and predicted CRRS values for the three retained machine learning models. Black circles represent observed CRRS values, blue diamonds represent out-of-sample predictions obtained during the outer leave-one-out cross-validation procedure, and the dashed horizontal line indicates the mean CRRS value.



Added value of combined resting-state and hyperventilation features

For each retained model, prediction errors were compared with the best corresponding model trained exclusively on resting-state or hyperventilation-derived features. Comparisons were performed using subject-wise absolute errors from leave-one-out predictions, represented in Figure 3.7. Paired permutation tests were then used to assess whether the observed reduction in prediction error for combined feature sets was statistically significant relative to the corresponding single-condition alternatives. Results are summarized in Table 3.5, while permutation distributions are provided in Appendix C.

Figure 3.7: Subject-wise absolute prediction errors for the three retained combined models and their corresponding single-condition models. Bars represent patient-wise prediction errors. Dashed horizontal lines indicate mean absolute error for each model, and the dotted red line indicates dummy regressor performance.

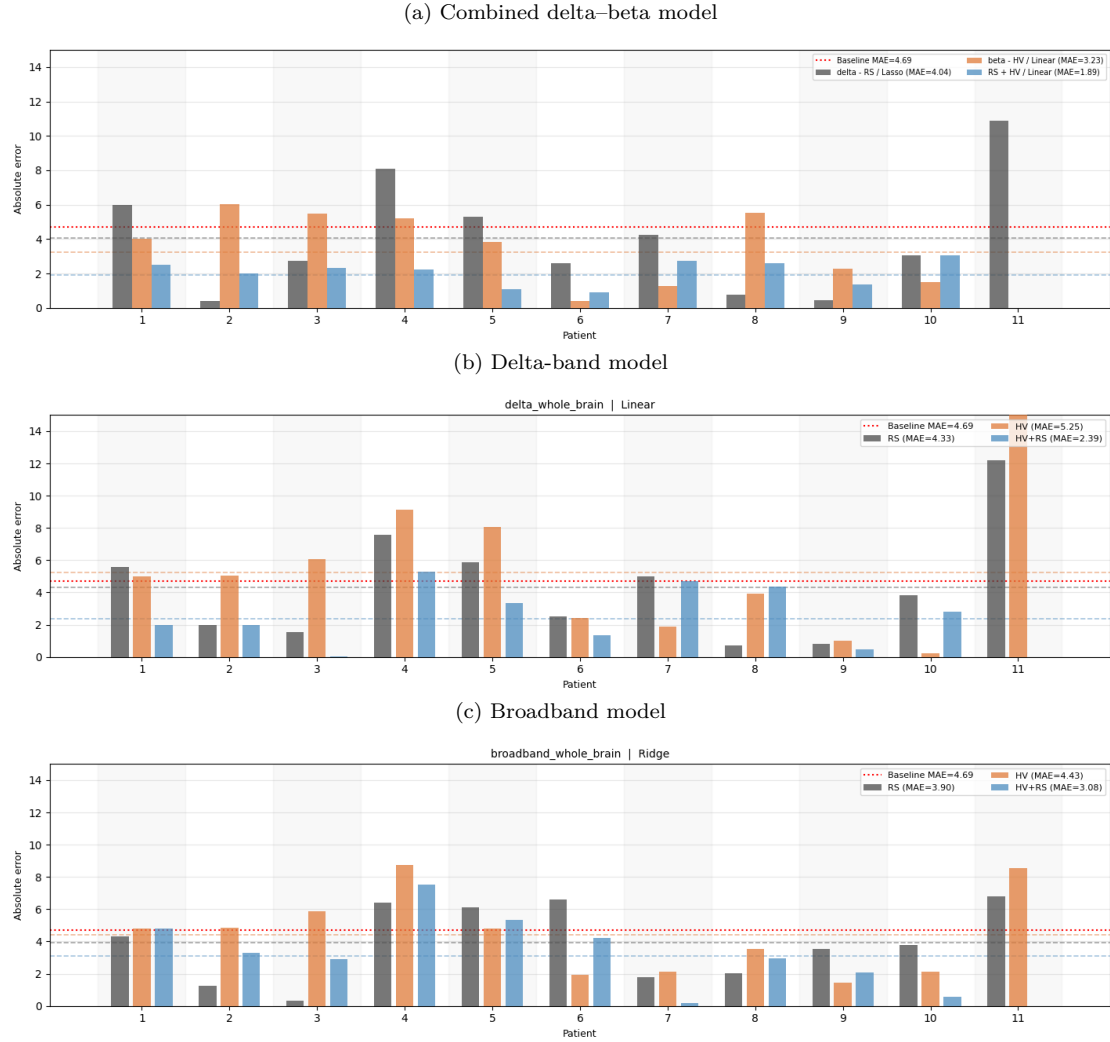


Table 3.5: Comparison using paired permutation tests between retained combined models and corresponding resting-state-only and hyperventilation-only models. Negative ΔMAE values indicate lower prediction error for the combined model.

Combined model	vs HV only		vs RS only	
	ΔMAE	p -value	ΔMAE	p -value
Combined delta-beta	-1.348	0.044	-2.154	0.024
Delta RS + HV	-2.493	0.021	-1.649	0.087
Broadband RS + HV	-0.371	0.355	-0.445	0.314

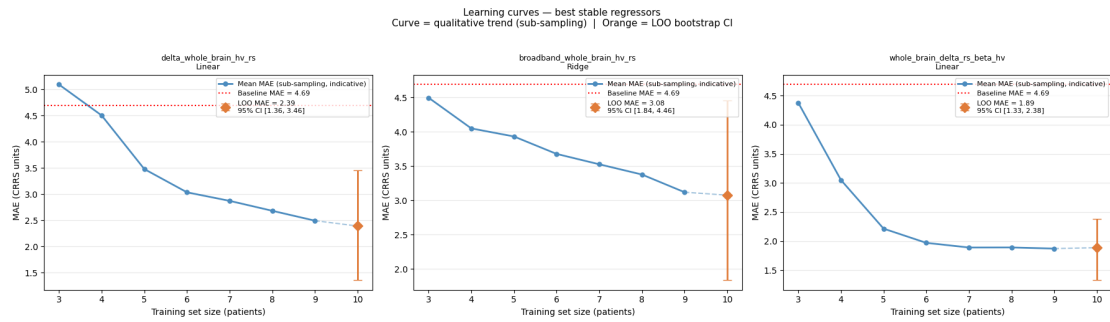
The combined delta–beta model showed significantly lower prediction error than both corresponding single-condition models. In contrast, the delta RS + HV model showed significantly lower prediction error only relative to the hyperventilation-only model, whereas no significant improvements were observed for the broadband RS + HV model.

Overall, the strongest benefit of combining resting-state and hyperventilation-derived connectivity was observed for the combined delta–beta model.

Learning curve analysis

Learning curves were evaluated for the retained models to investigate the evolution of predictive performance as training sample size increased. Figure 3.8 presents the average prediction error obtained for progressively larger training subsets.

Figure 3.8: Learning curves for retained machine learning models. Curves represent mean MAE estimated from exhaustive subsampling across training-set sizes. Orange markers indicate leave-one-out performance with corresponding bootstrap confidence intervals. The red dotted line represents dummy regressor performance.



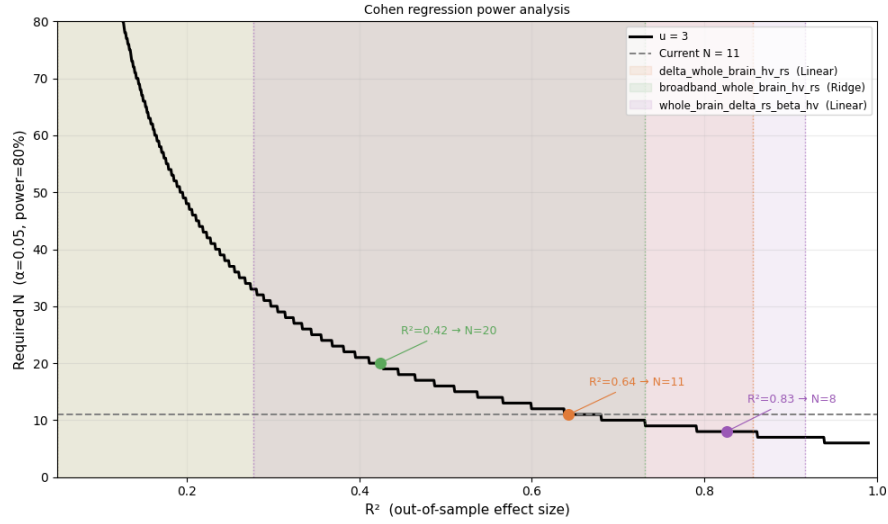
All models showed decreasing prediction error as training sample size increased. However, the combined delta–beta model demonstrated a more rapid reduction in MAE and appeared to approach a plateau at larger sample sizes. In contrast, the broadband and delta whole-brain RS + HV models showed more gradual decreases, suggesting that additional training samples could potentially further improve performance.

Power analysis

Power analyses were performed to estimate the sample sizes required to achieve 80% statistical power according to Cohen’s regression framework. Analyses used out-of-sample R^2 values derived from leave-one-out predictions and incorporated

bootstrap confidence intervals to account for uncertainty in predictive effect-size estimation. Figure 3.9 illustrates the relationship between predictive effect size and required sample size.

Figure 3.9: Regression power analysis according to Cohen’s framework ($u = 3$, $\alpha = 0.05$, power=80%). Colored points indicate observed out-of-sample R^2 values for retained models and corresponding required sample sizes. Shaded regions represent bootstrap confidence intervals.



The combined delta–beta model showed the largest out-of-sample effect size ($R^2 = 0.825$, 95% CI [0.277, 0.917]), corresponding to an estimated sample size requirement of eight subjects. However, uncertainty remained substantial, with the lower confidence bound corresponding to an estimated requirement of 33 subjects. The delta-band and broadband models showed smaller observed effect sizes ($R^2 = 0.643$, 95% CI [-0.280, 0.856] and $R^2 = 0.424$, 95% CI [-0.616, 0.731], respectively), corresponding to estimated sample sizes of 11 and 20 subjects.

Notably, the lower confidence bounds of the delta-band and broadband models included negative R^2 values. Under bootstrap resampling, this indicates that some resampled models performed worse than prediction using the sample mean. Such findings illustrate the instability of effect-size estimation in very small datasets and suggest that larger cohorts may be required to establish robust predictive performance.

To further contextualize these findings, Table 3.6 summarizes the relationship between predictive effect size and required sample size for regression models with three predictors.

Table 3.6: Required sample sizes according to Cohen’s regression framework ($u = 3$, $\alpha = 0.05$, power=80%) for different hypothetical out-of-sample R^2 values.

R^2	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
Required N	103	48	30	21	16	12	10	8	7

3.4 Longitudinal evolution of wPLI following VNS implantation

Longitudinal analyses first assessed global wPLI changes following VNS implantation independently of clinical outcome. Response-dependent longitudinal changes were subsequently evaluated by testing the interaction between clinical response and the six-month follow-up visit (V5). Additional exploratory analyses incorporating demographic and clinical covariates did not substantially alter the longitudinal mixed-effects model results, indicating that the observed findings remained stable after adjustment for these variables. Graphical representations of all significant region $\times$ band combinations together with associated statistical summaries (estimate (β), standard error (SE), raw p -value, and p_{FDR}) are provided in Appendix B.

3.4.1 Global longitudinal changes

No significant V5 fixed effects remained after FDR correction in hyperventilation, resting-state eyes-open, or resting-state eyes-closed recordings. These results indicate that no global longitudinal changes in wPLI between baseline and six-month follow-up were detected when considering the cohort as a whole, independently of clinical response.

3.4.2 Binary response

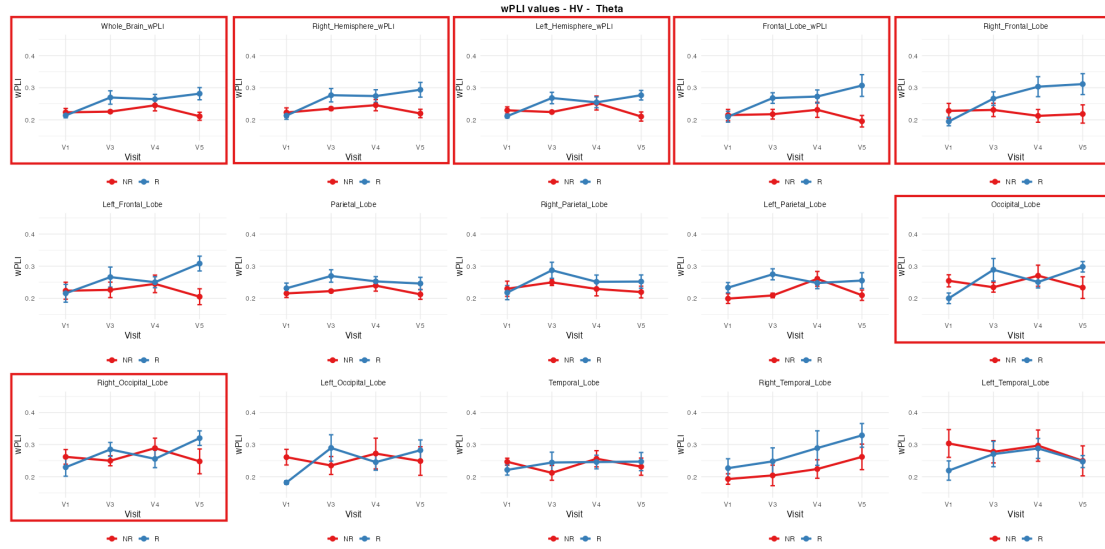
In the binary response models, no significant response $\times$ V5 interaction effects were observed during resting-state recordings, either with eyes open or eyes closed.

In contrast, hyperventilation recordings revealed several significant interaction effects, indicating differential longitudinal wPLI changes between responders and non-responders. Significant effects were predominantly observed in the theta band (7 regions), with fewer effects in the alpha (2 regions) and broadband (2 regions) bands.

Within the theta band, significant interaction effects involved whole-brain, bilateral

hemispheric, occipital, and frontal measures. Interaction coefficients ranged from $\beta = 0.081$ to $\beta = 0.126$, consistent with greater longitudinal increases in wPLI among responders relative to non-responders. Longitudinal trajectories for theta-band hyperventilation recordings are illustrated in Figure 3.10. Significant alpha-band interaction effects were observed in the temporal lobe ($\beta = 0.116$) and right temporal lobe ($\beta = 0.196$), whereas broadband effects involved the left frontal lobe ($\beta = 0.079$) and right occipital lobe ($\beta = 0.072$). Corresponding figures are provided in Appendix B.

Figure 3.10: Longitudinal evolution of theta-band wPLI values during hyperventilation recordings following VNS implantation according to binary clinical response. Mean values across visits are shown separately for responders (blue) and non-responders (red). Error bars indicate standard errors of the mean. Brain regions exhibiting significant response $\times$ V5 interaction effects after FDR correction are outlined.



3.4.3 Seizure reduction percentage

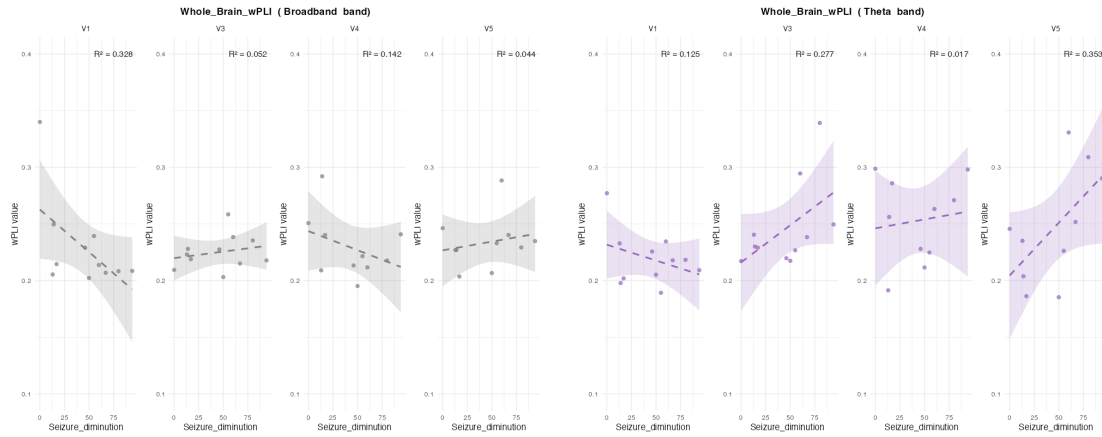
No significant response $\times$ V5 interaction effects were identified during resting-state recordings with either eyes-open or eyes-closed conditions.

In contrast, hyperventilation recordings revealed several significant interaction effects between seizure reduction percentage and V5, indicating that longitudinal changes in wPLI varied according to clinical outcome. Significant findings were observed in the theta (4 regions), alpha (1 region), high-beta (4 regions), and broadband (9 regions) frequency bands.

Theta-band findings involved whole-brain, right-hemisphere, left-hemisphere, and right frontal measures. Interaction estimates ranged from $\beta = 0.0012$ to $\beta = 0.0021$, indicating larger longitudinal increases in wPLI values with increasing seizure reduction percentage.

Significant high-beta findings involved the left hemisphere, left frontal lobe, left parietal lobe, and right occipital lobe. Broadband findings showed the widest regional distribution, involving whole-brain, hemispheric, frontal, parietal, and occipital measures. A single alpha-band finding remained significant after FDR correction in the right temporal lobe and represented the strongest effect observed across all frequency bands ($\beta = 0.0036$, FDR-corrected $p = 0.0029$). Longitudinal trajectories of all significant region $\times$ frequency-band combinations are provided in Appendix B; two illustrative examples are presented in Figure 3.11.

Figure 3.11: Representative examples of longitudinal associations between hyperventilation wPLI values and seizure reduction percentage across follow-up visits. Each panel corresponds to a visit (V1, V3, V4, and V5) and displays individual patient data with fitted linear regression lines (dashed) and corresponding 95% confidence intervals (shaded areas). The coefficient of determination (R^2) is displayed for each visit.



3.4.4 CRRS

No significant response $\times$ V5 interaction effects were observed during resting-state recordings with eyes open. However, in contrast to the previous outcome measures, significant interaction effects were identified during resting-state recordings with eyes closed.

During hyperventilation, significant findings were observed in the alpha (1 region), high-beta (3 regions), and broadband (7 regions) frequency bands. Importantly, all significant hyperventilation region $\times$ band combinations identified for CRRS had

also been identified in the seizure reduction analyses, indicating substantial overlap between both outcome measures. The alpha band yielded a single significant finding in the right temporal lobe, representing the strongest effect observed in this analysis ($\beta = 0.018$, FDR-corrected $p = 0.032$). High-beta findings involved the left hemisphere, left frontal lobe, and right occipital lobe, with interaction estimates ranging from $\beta = 0.006$ to $\beta = 0.009$. Broadband findings involved whole-brain, hemispheric, frontal, parietal, and occipital measures, with interaction estimates ranging from $\beta = 0.005$ to $\beta = 0.007$. All significant hyperventilation interaction coefficients were positive, indicating greater longitudinal increases in wPLI values with increasing CRRS scores.

In contrast, resting-state eyes-closed recordings revealed four significant delta-band findings involving whole-brain, right hemispheric, parietal, and right parietal measures. Interaction estimates ranged from $\beta = -0.006$ to $\beta = -0.008$. Unlike hyperventilation findings, all significant coefficients were negative, indicating lower longitudinal wPLI changes with increasing CRRS scores. The significant longitudinal trajectories are shown in Figure 3.12.

3.5 Acute and chronic effect of VNS comparison

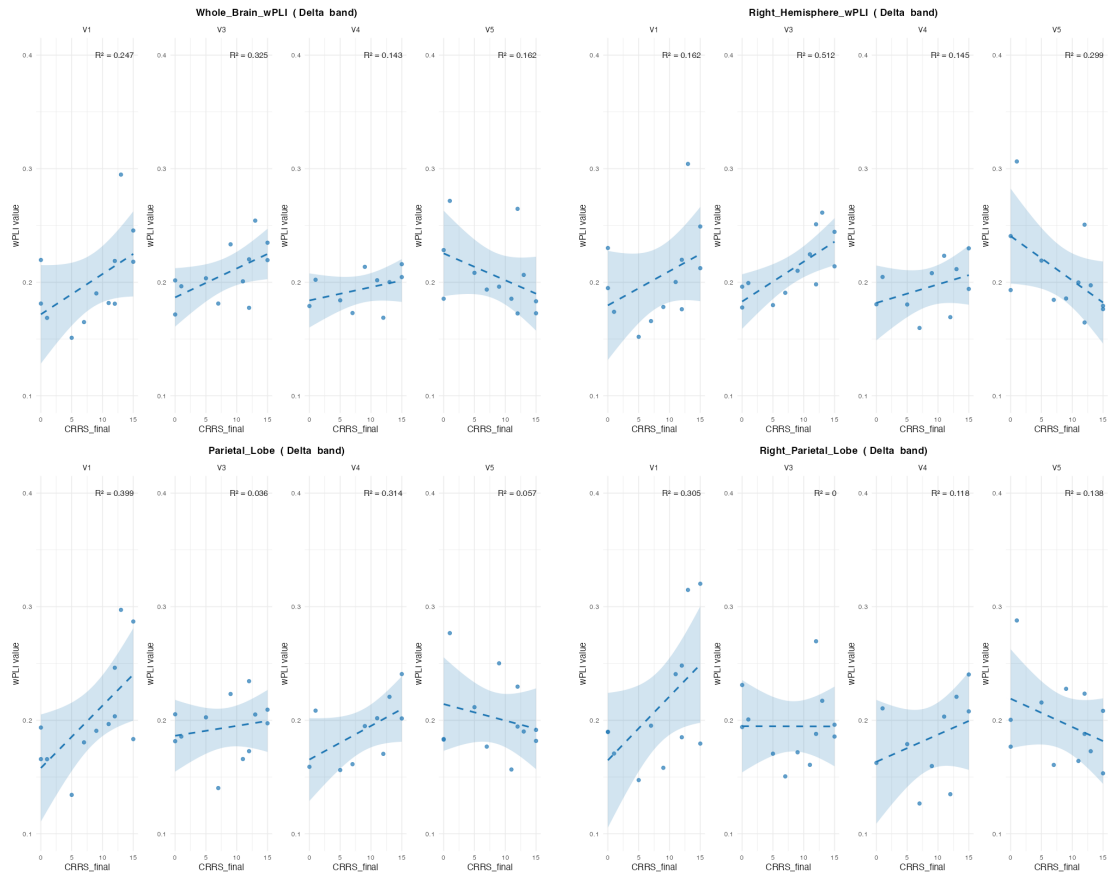
Acute effects of VNS stimulation were evaluated by comparing EEG recordings acquired during stimulation OFF and ON conditions at V5.

In the analysis including all patients, only a limited number of significant findings remained after FDR correction. Significant differences were observed exclusively during resting-state recordings under eyes-closed conditions. Four region–band combinations reached statistical significance after correction: the occipital lobe and right occipital lobe in the beta band, and the occipital lobe and right occipital lobe in the broadband range (FDR-corrected $p = 0.0402$ for all comparisons). Effect sizes were large ($r = 0.804$ – 0.827).

No significant ON/OFF differences were observed during hyperventilation or intermittent photic stimulation after multiple-comparison correction.

When analyses were restricted to responders only, no region–band combination remained significant following FDR correction. Therefore, restricting analyses to responders did not reveal additional significant acute stimulation effects.

Figure 3.12: Longitudinal associations between resting-state eyes-closed wPLI values and CRRS across follow-up visits for all significant region $\times$ band combinations. Each panel corresponds to a visit (V1, V3, V4, and V5) and displays individual patient data with fitted linear regression lines (dashed) and corresponding 95% confidence intervals (shaded areas). The coefficient of determination (R^2) is displayed for each visit.



Chapter 4

Discussion

4.1 Main findings

The present study investigated hyperventilation-induced EEG connectivity as a potential biomarker of VNS response and compared it with resting-state connectivity measures in three contexts: prediction of subsequent clinical response from pre-implantation recordings, longitudinal evolution of brain synchronization following VNS implantation, and acute connectivity changes during VNS ON versus OFF periods.

Regarding the first research question, several baseline connectivity measures were associated with subsequent clinical outcome, although the observed patterns differed according to experimental condition and response metric. No significant associations were identified for binary responder classification after correction for multiple comparisons. In contrast, analyses using continuous outcome measures, particularly CRRS, revealed several significant findings.

No clear spatial pattern consistently emerged across analyses. Significant findings were distributed across multiple regions rather than concentrating within a specific anatomical area. In contrast, frequency-specific patterns appeared more reproducible. Hyperventilation showed the largest number of significant associations, predominantly involving the beta and low-beta frequency bands, whereas resting-state recordings yielded only a limited number of findings in broadband connectivity.

The second research question addressed longitudinal modifications in connectivity following VNS implantation. Again, hyperventilation and resting-state recordings showed distinct patterns. During hyperventilation, longitudinal changes associated

with clinical response were mainly observed in theta, high-beta, and broadband activity. Across all significant region $\times$ frequency-band combinations, greater clinical improvement was associated with increasing hyperventilation wPLI over time.

In contrast, resting-state eyes-closed recordings demonstrated a different profile. Significant findings were concentrated in the delta band and showed the opposite directionality, with greater clinical response associated with lower longitudinal wPLI changes over time.

As for baseline analyses, longitudinal findings did not reveal a clearly reproducible anatomical organization. Significant regions varied across frequency bands and outcome measures, while frequency-related effects appeared more stable.

Finally, the third research question examined whether acute VNS stimulation induced measurable connectivity changes. Only a small number of differences between VNS ON and OFF conditions remained significant after correction for multiple comparisons, and these findings were restricted to resting-state eyes-closed recordings. No significant acute effects were observed during hyperventilation, and no significant findings remained when analyses were restricted to responders.

Taken together, the present results suggest that hyperventilation may reveal synchronization patterns distinct from resting-state recordings, supporting its potential value as an additional activation paradigm in EEG-based biomarker research. Acute stimulation appeared to produce only limited connectivity differences, whereas several longitudinal changes were observed in association with clinical response.

4.2 Comparison with previous investigations on the same cohort

The present work extends two previous investigations performed on the same prospective cohort, which respectively focused on resting-state EEG synchronization and intermittent photic stimulation (IPS) [36, 40]. Together, these studies provide an important framework for interpreting the present hyperventilation findings.

Danthine et al. investigated resting-state EEG synchronization before and after VNS implantation using global wPLI values during eyes-open and eyes-closed recordings [36]. Their main longitudinal finding was a progressive decrease in global delta-band wPLI during the eyes-closed VNS OFF condition, which correlated with clinical response over time, particularly when response was assessed using the CRRS. They also investigated whether baseline resting-state connectivity predicted VNS outcome and reported that higher pre-implantation global delta-band wPLI

was associated with better clinical response.

Importantly, this predictive relationship was sensitive to influential observations and reached significance only after applying robust regression (`lmrob()`) [36]. Unlike conventional linear regression (`lm()`), robust regression reduces the influence of observations with large residuals and limits the impact of outliers on parameter estimation. In the present study, baseline analyses relied on standard linear models and additionally incorporated FDR correction because regional wPLI measures substantially increased the number of statistical comparisons. In contrast, Danthine et al. investigated only global wPLI measures and therefore did not apply multiple-comparison correction.

For this reason, the absence of significant baseline delta-band findings in the present work should not be interpreted as contradictory evidence but as a reflect of methodological differences. This distinction becomes particularly relevant given that resting-state delta-band connectivity later contributed to the best-performing machine learning model of the present study.

Cakiroglu et al. investigated IPS-induced EEG synchronization in the same cohort, with a particular focus on occipital beta-band wPLI [40]. They reported that non-responders exhibited higher beta-band synchronization during IPS compared with resting state, whereas responders did not show the same increase. Higher CRRS scores were also associated with lower IPS-induced beta-band synchronization at baseline, with similar tendencies during follow-up [40].

These IPS findings are particularly relevant because both IPS and hyperventilation identified beta-band synchronization during activation procedures as potentially associated with VNS outcome. Unlike hyperventilation, IPS primarily stimulates visual cortical systems; therefore, the predominance of occipital findings in Cakiroglu et al. is physiologically expected. In contrast, hyperventilation induces broader physiological changes and significant findings in the present work were distributed across multiple regions.

Taken together, the three studies suggest that EEG synchronization findings within this cohort depend strongly on recording condition: resting-state recordings highlighted delta-band synchronization as potentially relevant, whereas IPS and hyperventilation both revealed beta-related activation findings associated with clinical outcome. These observations support the idea that activation procedures may reveal information not fully captured by resting-state recordings alone.

4.3 Interpretation of the machine learning model

In addition to the univariate statistical analyses, machine learning models were developed to assess whether multiple baseline connectivity features jointly contributed to VNS outcome prediction.

The best-performing predictive model combined resting-state delta-band wPLI with hyperventilation beta-band wPLI and achieved the lowest prediction error for CRRS, with a MAE of 1.887. This dataset was intentionally constructed based on prior observations: hyperventilation beta-band features were selected because they emerged repeatedly in the baseline statistical analyses of the present study, whereas resting-state delta-band features were included based on findings reported by Danthine et al. [36].

Model significance was assessed using permutation testing. In permutation testing, the outcome labels are repeatedly shuffled and model performance recalculated under the null hypothesis of no association between predictors and outcome. A significant result therefore suggests that the observed prediction performance is unlikely to arise by chance alone under random label assignment.

Model stability was also explicitly evaluated because prediction models developed on small datasets may be highly dependent on the specific sample used during development [57]. Stability analyses indicated that the selected models remained relatively consistent across leave-one-out iterations, suggesting that its predictive performance was not entirely driven by a small subset of patients.

Learning-curve analyses provided additional insight into the effect of sample size on model performance. For the best-performing model, the MAE decreased markedly as the training set size increased and then appeared to approach a plateau with the available sample size. This pattern suggests that predictive performance became relatively stable within the current cohort and that the model captured reproducible information rather than continuously improving with each additional patient. Furthermore, the leave-one-out estimate showed a relatively narrow confidence interval, suggesting reasonably consistent predictive performance despite the limited sample size.

Power analyses further contextualized the sample size requirements of the predictive models. According to Cohen’s regression framework [53], estimated sample sizes required to achieve 80% statistical power should not be interpreted as invalidating findings obtained with smaller samples. Rather, they provide an indication of the sample size required to reliably detect effects of a given magnitude.

For the best-performing combined delta–beta model, the estimated sample size

requirement was close to the size of the current cohort. However, confidence intervals around predictive effect-size estimates remained broad enough to produce substantially different sample size estimates. This suggests that the main limitation is not necessarily the observed sample size itself, but uncertainty regarding the true magnitude of the predictive effect. Such uncertainty is expected in small exploratory cohorts, where effect-size estimates may remain unstable.

Taken together, these observations suggest that the present machine learning findings provide encouraging indications regarding potentially relevant feature combinations. The apparent stabilization of the learning curve, relatively consistent prediction performance, and significant permutation results support the presence of predictive information within the current dataset. Nevertheless, larger cohorts remain necessary to obtain more precise effect-size estimates and determine whether these relationships generalize beyond the present sample.

4.4 Comparison with the broader literature

As discussed in Section 1.5, previous studies investigating EEG-derived functional connectivity in VNS reported heterogeneous findings that depended on the recording condition, frequency band, synchronization metric, and analytical approach. Moreover, most investigations focused on resting-state recordings, sleep, ictal activity, or VNS ON/OFF effects, whereas hyperventilation-induced functional connectivity has not previously been specifically investigated in the context of VNS response. Consequently, direct comparison with the present results remains limited.

The current findings suggest that hyperventilation may reveal connectivity patterns distinct from those observed during resting-state recordings and may provide complementary information regarding clinical outcome. However, because previous literature specifically addressing hyperventilation-induced connectivity in VNS is currently lacking, these observations should be interpreted cautiously and primarily as an indication that hyperventilation may represent a relevant avenue for further investigation.

Interestingly, only limited acute VNS ON/OFF effects were identified in the present study, whereas several previous investigations reported stimulation-related connectivity changes [20, 21, 22]. However, given the considerable methodological differences across studies and the small number of significant findings in the current cohort, no strong conclusion can be drawn from this discrepancy.

4.5 Limitations

Several limitations should be acknowledged when interpreting the present findings.

The principal limitation of this work is the limited sample size, reflecting the difficulty of recruiting a prospective longitudinal VNS cohort with repeated EEG recordings. Small cohorts reduce statistical power and increase uncertainty regarding effect estimates. Although several conservative approaches, including nested cross-validation, permutation testing, FDR correction, and stability analyses, were implemented to reduce overfitting and optimistic bias, replication remains necessary.

Another important consideration concerns the large number of evaluated region $\times$ frequency-band combinations. Although multiple-comparison correction was applied, the exploratory nature of these analyses increases the possibility that some findings may reflect sample-specific effects rather than robust underlying relationships.

Finally, machine learning analyses relied exclusively on internal validation. Consequently, although the results provide indications regarding potentially relevant feature combinations, their generalizability remains to be confirmed in independent cohorts.

4.6 Perspectives and further investigations

Several further investigations could help clarify the role of hyperventilation-induced connectivity as a potential biomarker of VNS response.

A natural continuation of the present work would involve replication in a larger cohort. Increasing the number of included patients would improve statistical power, reduce uncertainty around estimated effects, and provide more reliable evaluation of predictive model stability.

The present analyses also focused on absolute hyperventilation wPLI values. Future work could investigate relative hyperventilation responses by quantifying connectivity changes with respect to each patient’s own resting-state baseline. A HV/RS ratio could therefore characterize the connectivity modification induced by hyperventilation itself rather than the absolute synchronization level. Conceptually, this approach resembles OFF/ON ratios previously used in studies investigating VNS-induced synchronization changes [21, 25].

Comparison between hyperventilation and IPS may also represent an interesting direction. Both procedures are standard EEG activation paradigms and both

appeared associated with beta-related findings in this cohort. Evaluating these paradigms within a common analytical framework could help determine whether they provide complementary or partially redundant information regarding VNS response.

Finally, future analyses could extend beyond regional averages of wPLI values and investigate graph-theoretical approaches. Several previous studies suggested that network-level metrics may reveal information not fully captured by averaged synchronization measures, potentially providing additional insight into VNS-related network organization.

4.7 Conclusion

The present work suggests that hyperventilation may provide complementary information to resting-state EEG for investigating VNS-related functional connectivity changes. Hyperventilation revealed several frequency-specific associations with clinical outcome, particularly involving beta-band activity, and showed patterns distinct from those observed during resting-state recordings. In addition, machine learning analyses suggested that combining hyperventilation and resting-state connectivity features may improve outcome prediction, supporting the idea that both recording conditions provide complementary information.

Although the present findings remain exploratory and require replication in larger cohorts, they support the potential relevance of hyperventilation as an additional EEG paradigm for investigating biomarkers of VNS response.

Bibliography

- [1] R. S. Fisher, C. Acevedo, A. Arzimanoglou, et al. “ILAE official report: a practical clinical definition of epilepsy”. In: *Epilepsia* 55.4 (2014), pp. 475–482. DOI: 10.1111/epi.12550.
- [2] World Health Organization. *Epilepsy*. <https://www.who.int/news-room/fact-sheets/detail/epilepsy>. Accessed: 2026-04-13. 2023.
- [3] S. C. Schachter and J. I. Sirven. “Vagus nerve stimulation therapy for seizures”. In: *UpToDate*. Ed. by R. F. Connor. Accessed: 2026-04-13. Wolters Kluwer, 2026.
- [4] M. Chase, M. Sterman, and C. Clemente. “Cortical and subcortical patterns of response to afferent vagal stimulation”. In: *Experimental Neurology* 16.1 (1966), pp. 36–49. DOI: [https://doi.org/10.1016/0014-4886\(66\)90084-7](https://doi.org/10.1016/0014-4886(66)90084-7).
- [5] M. H. Chase, Y. Nakamura, C. D. Clemente, and M. B. Sterman. “Afferent vagal stimulation: Neurographic correlates of induced eeg synchronization and desynchronization”. In: *Brain Research* 5.2 (1967), pp. 236–249. DOI: [https://doi.org/10.1016/0006-8993\(67\)90089-3](https://doi.org/10.1016/0006-8993(67)90089-3).
- [6] G. Chiarion, L. Sparacino, Y. Antonacci, et al. “Connectivity Analysis in EEG Data: A Tutorial Review of the State of the Art and Emerging Trends”. In: *Bioengineering* 10.3 (2023). DOI: 10.3390/bioengineering10030372.
- [7] M. G. Preti, T. A. Bolton, and D. Van De Ville. “The dynamic functional connectome: State-of-the-art and perspectives”. In: *NeuroImage* 160 (2017), pp. 41–54. DOI: <https://doi.org/10.1016/j.neuroimage.2016.12.061>.
- [8] S. Coelli, M. Corda, and A. M. Bianchi. “The time-varying brain: a comprehensive review of dynamic functional connectivity analysis in EEG and MEG”. In: *Journal of neural engineering* 22.5 (2025). DOI: 10.1088/1741-2552/ae1258.
- [9] B. He, L. Astolfi, P. A. Valdes-Sosa, et al. “Electrophysiological Brain Connectivity: Theory and Implementation”. In: *IEEE Transactions on Biomedical Engineering* 66.7 (2019), pp. 2115–2137. DOI: 10.1109/TBME.2019.2913928.

- [10] J.-P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela. “Measuring phase synchrony in brain signals”. In: *Human Brain Mapping* 8.4 (1999), pp. 194–208. DOI: 10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C.
- [11] C. J. Stam, G. Nolte, and A. Daffertshofer. “Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources”. In: *Human Brain Mapping* 28.11 (2007), pp. 1178–1193. DOI: 10.1002/hbm.20346.
- [12] M. Vinck, R. Oostenveld, M. Van Wingerden, et al. “An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias”. In: *NeuroImage* 55.4 (2011), pp. 1548–1565. DOI: 10.1016/j.neuroimage.2011.01.055.
- [13] F. Mormann, K. Lehnertz, P. David, and C. E. Elger. “Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients”. In: *Physica D: Nonlinear Phenomena* 144.3 (2000), pp. 358–369. DOI: 10.1016/S0167-2789(00)00087-7.
- [14] G. Nolte, O. Bai, L. Wheaton, et al. “Identifying true brain interaction from EEG data using the imaginary part of coherency”. In: *Clinical Neurophysiology* 115.10 (2004), pp. 2292–2307. DOI: <https://doi.org/10.1016/j.clinph.2004.04.029>.
- [15] M. Hardmeier, F. Hatz, H. Bousleiman, et al. “Reproducibility of functional connectivity and graph measures based on the phase lag index (PLI) and weighted phase lag index (wPLI) derived from high resolution EEG”. In: *PLoS ONE* 9.10 (2014). DOI: 10.1371/journal.pone.0108648.
- [16] FieldTrip Development Team. *FieldTrip Connectivity Analysis Documentation*. https://github.com/fieldtrip/fieldtrip/blob/master/ft_connectivityanalysis.m. Accessed: 2026-05-07. 2025.
- [17] L. S. Imperatori, M. Betta, L. Cecchetti, et al. “EEG functional connectivity metrics wPLI and wSMI account for distinct types of brain functional interactions”. In: *Scientific Reports* 9.1 (2019). DOI: 10.1038/s41598-019-45289-7.
- [18] G. Nolte, A. Ziehe, V. V. Nikulin, et al. “Robustly estimating the flow direction of information in complex physical systems”. In: *Physical Review Letters* 100.23 (2008). DOI: 10.1103/PhysRevLett.100.234101.
- [19] F. Marrosu, F. Santoni, M. Puligheddu, et al. “Increase in 20-50 Hz (gamma frequencies) power spectrum and synchronization after chronic vagal nerve stimulation”. In: *Clinical Neurophysiology* 116.9 (2005), pp. 2026–2036. DOI: 10.1016/j.clinph.2005.06.015.

- [20] C. Bodin, S. Aubert, G. Daquin, et al. “Responders to vagus nerve stimulation (VNS) in refractory epilepsy have reduced interictal cortical synchronicity on scalp EEG”. In: *Epilepsy Research* 113 (2015), pp. 98–103. DOI: 10.1016/j.epilepsyres.2015.03.018.
- [21] A. Sangare, A. Marchi, E. Pruvost-Robieux, et al. “The effectiveness of vagus nerve stimulation in drug-resistant epilepsy correlates with vagus nerve stimulation-induced electroencephalography desynchronization”. In: *Brain Connectivity* 10.10 (2020), pp. 566–577. DOI: 10.1089/brain.2020.0798.
- [22] J. Lanzone, M. Boscarino, T. Tufo, et al. “Vagal nerve stimulation cycles alter EEG connectivity in drug-resistant epileptic patients: A study with graph theory metrics”. In: *Clinical Neurophysiology* 142 (2022), pp. 59–67. DOI: 10.1016/j.clinph.2022.07.503.
- [23] E. Germany Morrison, V. Danthine, R. Santalucia, et al. “Characterization of Vagus Nerve Stimulation (VNS) Dose-Dependent Effects on EEG Power Spectrum and Synchronization”. In: *Biomedicines* 12.3 (2024). DOI: 10.3390/biomedicines12030557.
- [24] F. Bartolomei, F. Bonini, E. Vidal, et al. “How does vagal nerve stimulation (VNS) change EEG brain functional connectivity?” In: *Epilepsy Research* 126 (2016), pp. 141–146. DOI: 10.1016/j.epilepsyres.2016.06.008.
- [25] S. Vespa, J. Heyse, L. Stumpp, et al. “Vagus Nerve Stimulation Elicits Sleep EEG Desynchronization and Network Changes in Responder Patients in Epilepsy”. In: *Neurotherapeutics* 18.4 (2021), pp. 2623–2638. DOI: 10.1007/s13311-021-01124-4.
- [26] V. Danthine, L. Cottin, A. Berger, et al. “Electroencephalogram synchronization measure as a predictive biomarker of Vagus nerve stimulation response in refractory epilepsy: A retrospective study”. In: *PLoS ONE* 19.6 June (2024). DOI: 10.1371/journal.pone.0304115.
- [27] Z.-J. Wang, E.-S. Kim, B. H. Noh, et al. “Alteration in brain connectivity in patients with Dravet syndrome after vagus nerve stimulation (VNS): Exploration of its effectiveness using graph theory analysis with electroencephalography”. In: *Journal of Neural Engineering* 17.3 (2020). DOI: 10.1088/1741-2552/ab914f.
- [28] R. Coa, S. M. La Cava, G. Baldazzi, et al. “Estimated EEG functional connectivity and aperiodic component induced by vagal nerve stimulation in patients with drug-resistant epilepsy”. In: *Frontiers in Neurology* 13 (2022). DOI: 10.3389/fneur.2022.1030118.

- [29] E. Germany, I. Teixeira, V. Danthine, et al. “Functional brain connectivity indexes derived from low-density EEG of pre-implanted patients as VNS outcome predictors”. In: *Journal of Neural Engineering* 20.4 (2023). DOI: 10.1088/1741-2552/acf1cd.
- [30] J. Ma, Z. Wang, T. Cheng, et al. “A prediction model integrating synchronization biomarkers and clinical features to identify responders to vagus nerve stimulation among pediatric patients with drug-resistant epilepsy”. In: *CNS Neuroscience and Therapeutics* 28.11 (2022), pp. 1838–1848. DOI: 10.1111/cns.13923.
- [31] H. Chen, Y. Wang, T. Ji, et al. “Brain functional connectivity-based prediction of vagus nerve stimulation efficacy in pediatric pharmacoresistant epilepsy”. In: *CNS Neuroscience and Therapeutics* 29.11 (2023), pp. 3259–3268. DOI: 10.1111/cns.14257.
- [32] Y. Ding, K. Guo, X. Wang, et al. “Brain functional connectivity and network characteristics changes after vagus nerve stimulation in patients with refractory epilepsy”. In: *Translational Neuroscience* 14.1 (2023). DOI: 10.1515/tnsci-2022-0308.
- [33] R. Kano, K. Usami, T. Noda, et al. “Vagus nerve stimulation-induced synchrony modulation of local field potential in the rat cerebral cortex”. In: *Electronics and Communications in Japan* 98.3 (2015), pp. 14–23. DOI: 10.1002/ecj.11674.
- [34] K. Usami, R. Kano, K. Kawai, et al. “State-dependent modulation of cortical synchrony by vagus nerve stimulation in adult rats”. In: *Electronics and Communications in Japan* 98.6 (2015), pp. 44–50. DOI: 10.1002/ecj.11735.
- [35] J. Nichols, A. Nichols, S. Smirnakis, et al. “Vagus nerve stimulation modulates cortical synchrony and excitability through the activation of muscarinic receptors”. In: *Neuroscience* 189 (2011), pp. 207–214. DOI: 10.1016/j.neuroscience.2011.05.024.
- [36] V. Danthine, E. I. Germany Morrison, L. Cottin, et al. “Effect of Vagus Nerve Stimulation on Electroencephalogram Synchronization: A Longitudinal Study Using a Clinical-Research Response Scale”. In: *Neuromodulation* (2025). DOI: 10.1016/j.neurom.2025.03.068.
- [37] E. Mazzucchi, C. Vollono, A. Losurdo, et al. “Hyperventilation in patients with focal epilepsy: Electromagnetic tomography, functional connectivity and graph theory: A possible tool in epilepsy diagnosis?” In: *Journal of Clinical Neurophysiology* 34.1 (2017), pp. 92–99. DOI: 10.1097/WNP.0000000000000329.

- [38] L. Ferri, F. Mason, L. Di Vito, et al. “Cortical Connectivity Response to Hyperventilation in Focal Epilepsy: A Stereo-EEG Study”. In: *Applied Sciences (Switzerland)* 14.18 (2024). DOI: 10.3390/app14188494.
- [39] M. Brázdil, I. Doležalová, E. Koritáková, et al. “EEG reactivity predicts individual efficacy of vagal nerve stimulation in intractable epileptics”. In: *Frontiers in Neurology* 10.MAY (2019). DOI: 10.3389/fneur.2019.00392.
- [40] I. Cakiroglu, V. Danthine, E. I. Vicine Raphaeland Germany Morrison, et al. “Vagus nerve stimulation modulates Intermittent Photic stimulation induced EEG synchronization in drug resistant epilepsy”. Submitted for publication. 2025.
- [41] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2025.
- [42] M. P. Fay and M. A. Proschan. “Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules”. In: *Statistical Surveys* 4 (2010), pp. 1–39. DOI: 10.1214/09-SS051.
- [43] M. M. Mukaka. “Statistics corner: A guide to appropriate use of correlation coefficient in medical research”. In: *Malawi Medical Journal* 24.3 (Sept. 2012), pp. 69–71.
- [44] P. Schober, C. Boer, and L. A. Schwarte. “Correlation Coefficients: Appropriate Use and Interpretation”. In: *Anesthesia & Analgesia* 126.5 (May 2018), pp. 1763–1768. DOI: 10.1213/ANE.0000000000002864.
- [45] Y. Benjamini and Y. Hochberg. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 57.1 (1995), pp. 289–300. DOI: 10.1111/j.2517-6161.1995.tb02031.x.
- [46] M. E. Glickman, S. R. Rao, and M. R. Schultz. “False Discovery Rate Control Is a Recommended Alternative to Bonferroni-Type Adjustments in Health Studies”. In: *Journal of Clinical Epidemiology* 67.8 (2014), pp. 850–857. DOI: 10.1016/j.jclinepi.2014.03.012.
- [47] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen. “lmerTest Package: Tests in Linear Mixed Effects Models”. In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26. DOI: 10.18637/jss.v082.i13.
- [48] L. Buitinck, G. Louppe, M. Blondel, et al. “API design for machine learning software: experiences from the scikit-learn project”. In: *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*. 2013, pp. 108–122.

- [49] F. Pedregosa, G. Varoquaux, A. Gramfort, et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [50] “A review of multivariate permutation tests: Findings and trends”. In: *Journal of Multivariate Analysis* 207 (2025), p. 105421. DOI: <https://doi.org/10.1016/j.jmva.2025.105421>.
- [51] T. Viering and M. Loog. “The Shape of Learning Curves: A Review”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.6 (2023), pp. 7799–7819. DOI: [10.1109/TPAMI.2022.3220744](https://doi.org/10.1109/TPAMI.2022.3220744).
- [52] F. Mohr and J. N. van Rijn. “Learning curves for decision making in supervised machine learning: a survey”. In: *Machine Learning* 113.11 (2024), pp. 8371–8425. DOI: [10.1007/s10994-024-06619-7](https://doi.org/10.1007/s10994-024-06619-7).
- [53] J. Cohen. “Multiple Regression and Correlation Analysis”. In: *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988. Chap. 9, pp. 407–465. ISBN: 9780805802832.
- [54] C. L. Chuan and J. Penyelidikan. “Sample size estimation using Krejcie and Morgan and Cohen statistical power analysis: A comparison”. In: *Jurnal Penyelidikan IPBL* 7 (2006), pp. 78–86.
- [55] SciPy Community. *scipy.stats.f* — *F distribution*. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.f.html>. SciPy v1.16.0 documentation. Accessed: 2026-05-14. 2026.
- [56] SciPy Community. *scipy.stats.ncf* — *Non-central F distribution*. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ncf.html>. SciPy v1.16.0 documentation. Accessed: 2026-05-14. 2026.
- [57] R. D. Riley and G. S. Collins. “Stability of clinical prediction models developed using statistical or machine learning methods”. In: *Biometrical Journal* 65.8 (2023). DOI: [10.1002/bimj.202200302](https://doi.org/10.1002/bimj.202200302).

Appendix A

Baseline (V1) statistical results

This appendix summarizes the relationships between baseline (V1) hyperventilation wPLI values and clinical outcomes following VNS implantation.

Figure A.1 presents boxplots comparing wPLI values between responders and non-responders across the seven frequency bands and the 15 predefined brain regions. Figures A.2–A.6 present scatterplots illustrating the relationship between baseline (V1) hyperventilation wPLI values and CRRS at six months following VNS implantation. CRRS was selected as the representative continuous outcome because it yielded the largest number of significant associations. Similar graphical patterns were observed for seizure reduction percentage; therefore, only CRRS results are shown for clarity and conciseness. Dashed lines represent linear regression fits with corresponding 95% confidence intervals (shaded areas). The coefficient of determination (R^2) is reported for each region.

Tables A.1–A.7 provide the corresponding Wilcoxon statistics, raw p-values and FDR-corrected p-values for the three response scales.

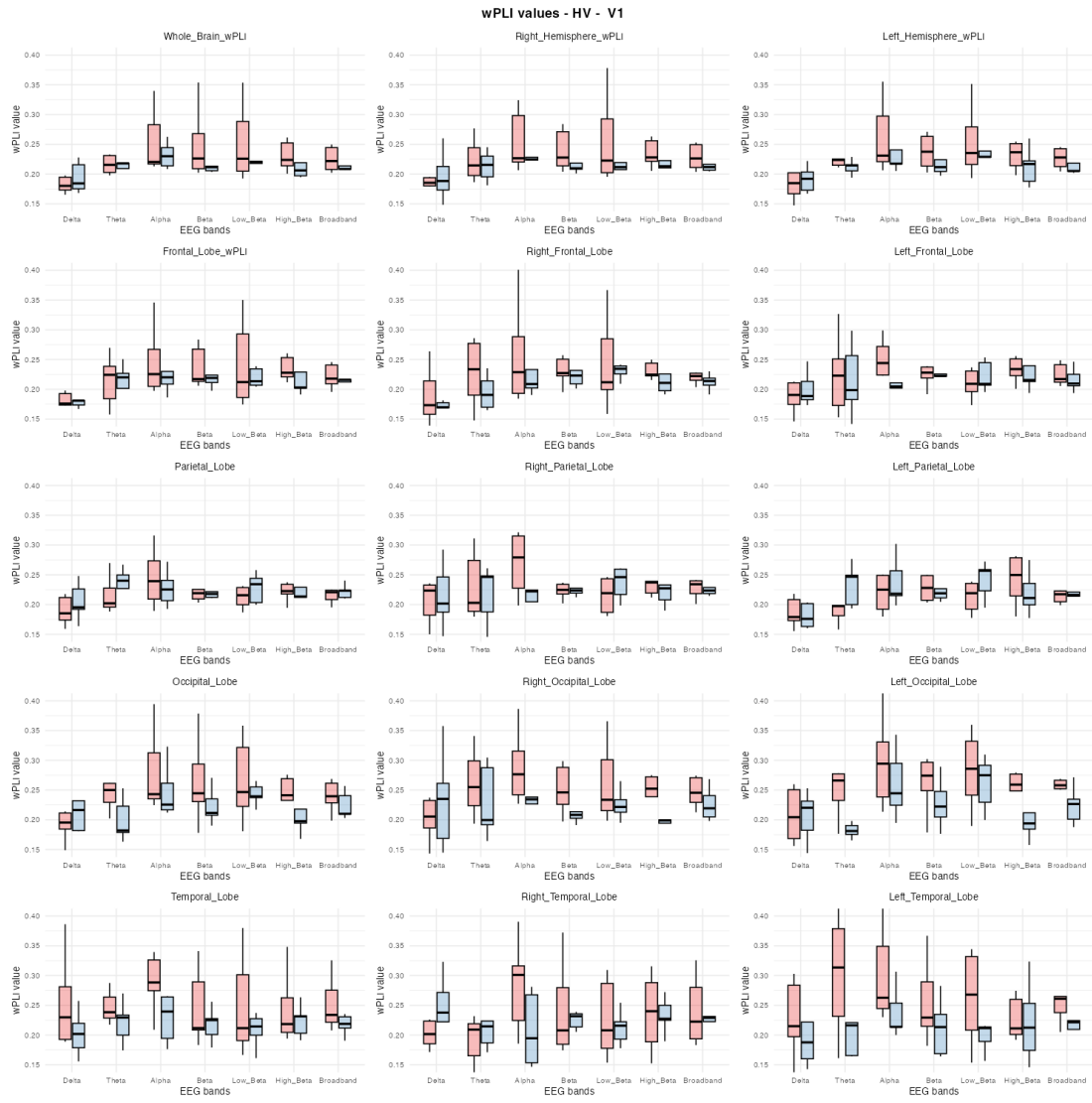


Figure A.1: Baseline (V1) hyperventilation wPLI values in responders (blue) and non-responders (red).

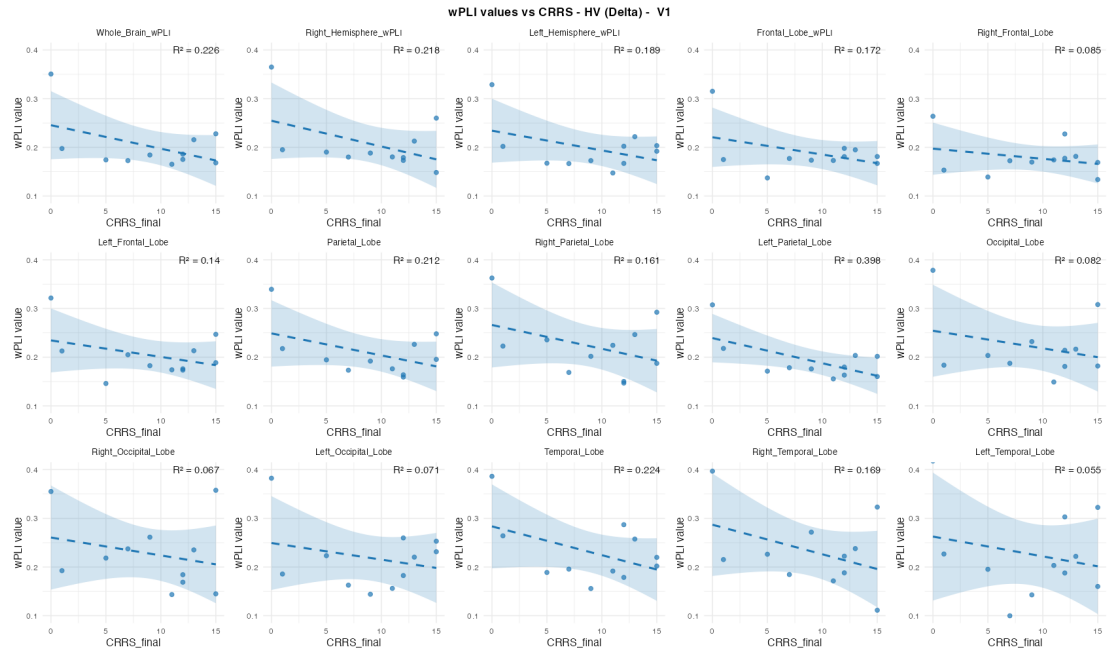


Figure A.2: Baseline (V1) delta-band hyperventilation wPLI values versus CRRS.

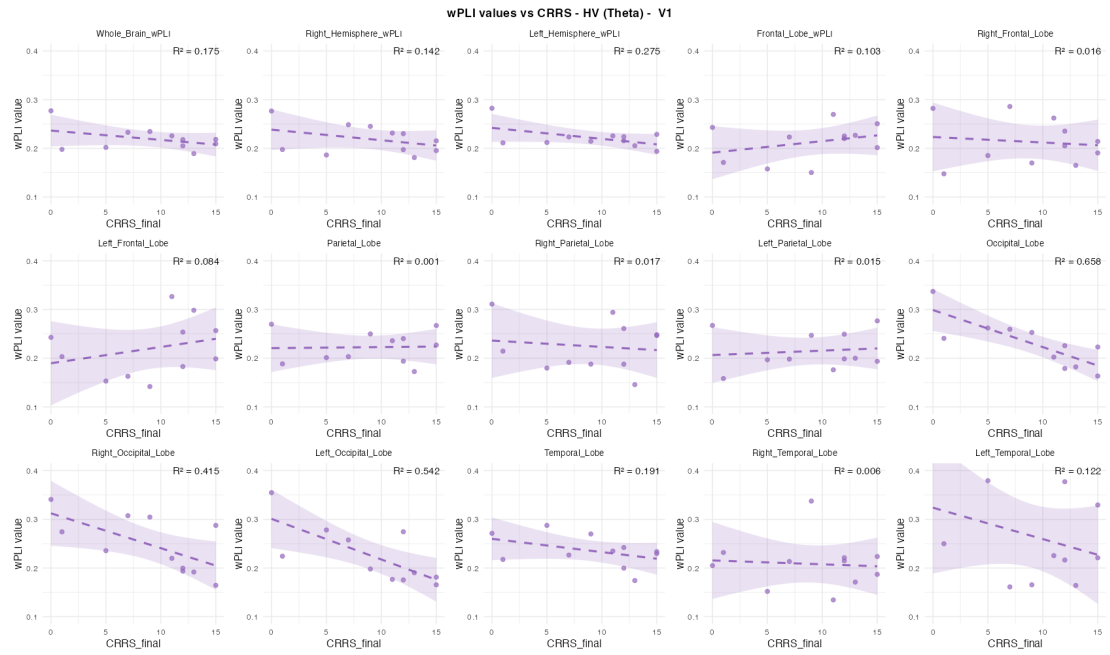


Figure A.3: Baseline (V1) theta-band hyperventilation wPLI values versus CRRS.

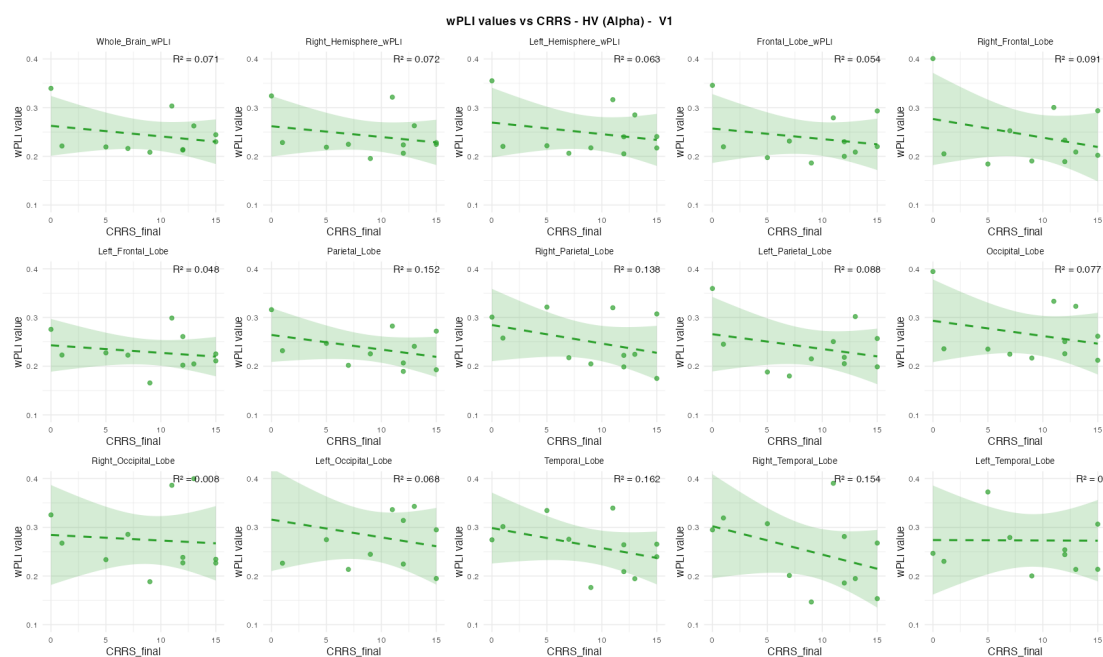


Figure A.4: Baseline (V1) alpha-band hyperventilation wPLI values versus CRRS.

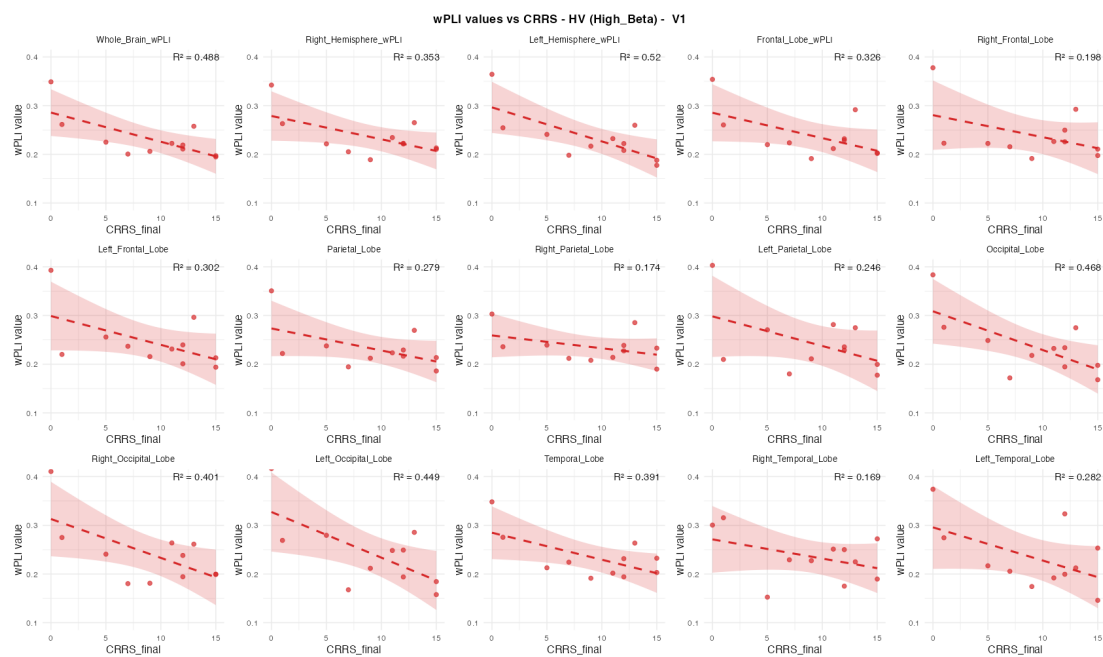


Figure A.5: Baseline (V1) high beta-band hyperventilation wPLI values versus CRRS.

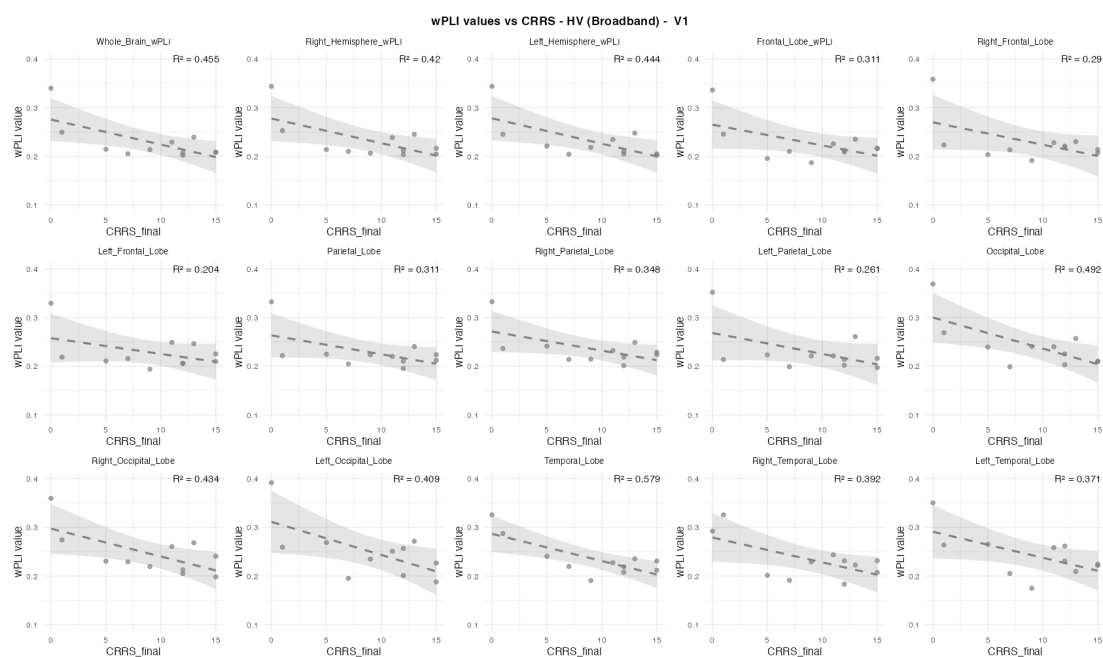


Figure A.6: Baseline (V1) broadband hyperventilation wPLI values versus CRRS.

Table A.1: Statistical results - alpha band - HV - V1. p-val, p-value; bin, binary response; % SR, percentage of seizure reduction; CRRS, Clinical-Response Research Scale.

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,792	0,849	0,383	0,541	0,429	0,598
Right_Hemisphere	0,537	0,805	0,316	0,541	0,426	0,598
Left_Hemisphere	0,429	0,804	0,354	0,541	0,457	0,598
Frontal_Lobe	0,792	0,849	0,657	0,657	0,490	0,598
Right_Frontal_Lobe	0,931	0,931	0,456	0,541	0,367	0,598
Left_Frontal_Lobe	0,017	0,227	0,281	0,541	0,518	0,598
Parietal_Lobe	0,662	0,828	0,267	0,541	0,236	0,598
Right_Parietal_Lobe	0,329	0,804	0,196	0,541	0,260	0,598
Left_Parietal_Lobe	0,662	0,828	0,469	0,541	0,376	0,598
Occipital_Lobe	0,329	0,804	0,274	0,541	0,409	0,598
Right_Occipital_Lobe	0,429	0,804	0,403	0,541	0,799	0,856
Left_Occipital_Lobe	0,537	0,805	0,284	0,541	0,438	0,598
Temporal_Lobe	0,030	0,227	0,163	0,541	0,219	0,598
Right_Temporal_Lobe	0,082	0,411	0,290	0,541	0,233	0,598
Left_Temporal_Lobe	0,178	0,666	0,581	0,623	0,985	0,985

Table A.2: Statistical results - beta band - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,537	0,990	0,040	0,136	0,010	0,044
Right_Hemisphere	0,247	0,823	0,054	0,136	0,020	0,044
Left_Hemisphere	0,329	0,823	0,033	0,136	0,009	0,044
Frontal_Lobe	0,662	0,990	0,074	0,138	0,025	0,044
Right_Frontal_Lobe	0,792	0,990	0,130	0,152	0,073	0,092
Left_Frontal_Lobe	0,792	0,990	0,143	0,153	0,106	0,108
Parietal_Lobe	1,000	1,000	0,126	0,152	0,070	0,092
Right_Parietal_Lobe	1,000	1,000	0,131	0,152	0,086	0,099
Left_Parietal_Lobe	1,000	1,000	0,130	0,152	0,108	0,108
Occipital_Lobe	0,329	0,823	0,049	0,136	0,011	0,044
Right_Occipital_Lobe	0,178	0,823	0,045	0,136	0,019	0,044
Left_Occipital_Lobe	0,247	0,823	0,040	0,136	0,018	0,044
Temporal_Lobe	0,792	0,990	0,073	0,138	0,021	0,044
Right_Temporal_Lobe	0,662	0,990	0,288	0,288	0,067	0,092
Left_Temporal_Lobe	0,247	0,823	0,097	0,152	0,026	0,044

Table A.3: Statistical results - high beta band - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,178	0,617	0,040	0,137	0,017	0,090
Right_Hemisphere	0,429	0,643	0,079	0,137	0,054	0,115
Left_Hemisphere	0,329	0,617	0,028	0,137	0,012	0,090
Frontal_Lobe	0,247	0,617	0,065	0,137	0,067	0,125
Right_Frontal_Lobe	0,247	0,617	0,124	0,168	0,170	0,197
Left_Frontal_Lobe	0,537	0,671	0,076	0,137	0,080	0,130
Parietal_Lobe	0,537	0,671	0,135	0,168	0,095	0,130
Right_Parietal_Lobe	0,329	0,617	0,200	0,231	0,201	0,209
Left_Parietal_Lobe	0,429	0,643	0,121	0,168	0,121	0,151
Occipital_Lobe	0,178	0,617	0,039	0,137	0,020	0,090
Right_Occipital_Lobe	0,178	0,617	0,054	0,137	0,037	0,099
Left_Occipital_Lobe	0,247	0,617	0,030	0,137	0,024	0,090
Temporal_Lobe	0,931	0,931	0,082	0,137	0,039	0,099
Right_Temporal_Lobe	0,792	0,849	0,433	0,433	0,209	0,209
Left_Temporal_Lobe	0,662	0,764	0,258	0,276	0,093	0,130

Table A.4: Statistical results - low beta band - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,662	0,903	0,073	0,256	0,010	0,049
Right_Hemisphere	0,662	0,903	0,058	0,256	0,010	0,049
Left_Hemisphere	0,662	0,903	0,085	0,256	0,015	0,049
Frontal_Lobe	0,931	0,997	0,243	0,304	0,037	0,055
Right_Frontal_Lobe	0,662	0,903	0,349	0,374	0,072	0,090
Left_Frontal_Lobe	0,662	0,903	0,764	0,764	0,408	0,408
Parietal_Lobe	0,537	0,903	0,215	0,294	0,094	0,109
Right_Parietal_Lobe	0,537	0,903	0,167	0,278	0,067	0,090
Left_Parietal_Lobe	0,429	0,903	0,345	0,374	0,230	0,247
Occipital_Lobe	0,931	0,997	0,133	0,262	0,016	0,049
Right_Occipital_Lobe	0,537	0,903	0,059	0,256	0,015	0,049
Left_Occipital_Lobe	0,537	0,903	0,140	0,262	0,032	0,055
Temporal_Lobe	0,792	0,990	0,116	0,262	0,025	0,054
Right_Temporal_Lobe	1,000	1,000	0,195	0,293	0,033	0,055
Left_Temporal_Lobe	0,247	0,903	0,051	0,256	0,022	0,054

Table A.5: Statistical results - delta band - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,792	0,914	0,299	0,650	0,140	0,369
Right_Hemisphere	0,792	0,914	0,319	0,650	0,148	0,369
Left_Hemisphere	0,662	0,914	0,318	0,650	0,182	0,369
Frontal_Lobe	1,000	1,000	0,280	0,650	0,205	0,369
Right_Frontal_Lobe	0,662	0,914	0,335	0,650	0,384	0,472
Left_Frontal_Lobe	0,792	0,914	0,375	0,650	0,257	0,386
Parietal_Lobe	0,662	0,914	0,390	0,650	0,154	0,369
Right_Parietal_Lobe	0,931	0,997	0,458	0,652	0,221	0,369
Left_Parietal_Lobe	0,662	0,914	0,086	0,650	0,037	0,369
Occipital_Lobe	0,792	0,914	0,735	0,772	0,392	0,472
Right_Occipital_Lobe	0,792	0,914	0,772	0,772	0,441	0,472
Left_Occipital_Lobe	0,792	0,914	0,533	0,652	0,428	0,472
Temporal_Lobe	0,329	0,914	0,161	0,650	0,142	0,369
Right_Temporal_Lobe	0,537	0,914	0,545	0,652	0,209	0,369
Left_Temporal_Lobe	0,537	0,914	0,565	0,652	0,486	0,486

Table A.6: Statistical results - theta band - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,931	0,931	0,287	0,623	0,200	0,501
Right_Hemisphere	0,537	0,732	0,363	0,623	0,254	0,544
Left_Hemisphere	0,329	0,705	0,087	0,437	0,098	0,366
Frontal_Lobe	0,792	0,931	0,728	0,780	0,336	0,559
Right_Frontal_Lobe	0,429	0,732	0,374	0,623	0,715	0,830
Left_Frontal_Lobe	0,931	0,931	0,689	0,780	0,386	0,579
Parietal_Lobe	0,537	0,732	0,535	0,730	0,919	0,919
Right_Parietal_Lobe	0,931	0,931	1,000	1,000	0,700	0,830
Left_Parietal_Lobe	0,178	0,666	0,525	0,730	0,719	0,830
Occipital_Lobe	0,052	0,390	0,010	0,076	0,002	0,037
Right_Occipital_Lobe	0,329	0,705	0,130	0,487	0,033	0,163
Left_Occipital_Lobe	0,030	0,390	0,003	0,048	0,010	0,073
Temporal_Lobe	0,247	0,705	0,307	0,623	0,179	0,501
Right_Temporal_Lobe	0,537	0,732	0,645	0,780	0,818	0,877
Left_Temporal_Lobe	0,178	0,666	0,321	0,623	0,292	0,548

Table A.7: Statistical results - broadband - HV - V1

Region	p-val (bin)	FDR p-val (bin)	p-val (% SR)	FDR p-val (% SR)	p-val (CRRS)	FDR p-val (CRRS)
Whole_Brain	0,662	0,914	0,066	0,153	0,023	0,073
Right_Hemisphere	0,537	0,895	0,071	0,153	0,031	0,073
Left_Hemisphere	0,329	0,895	0,054	0,153	0,025	0,073
Frontal_Lobe	0,792	0,914	0,129	0,167	0,074	0,093
Right_Frontal_Lobe	0,429	0,895	0,119	0,167	0,087	0,101
Left_Frontal_Lobe	0,429	0,895	0,150	0,167	0,163	0,163
Parietal_Lobe	0,792	0,914	0,156	0,167	0,075	0,093
Right_Parietal_Lobe	0,792	0,914	0,110	0,167	0,056	0,084
Left_Parietal_Lobe	0,931	0,997	0,151	0,167	0,108	0,116
Occipital_Lobe	0,537	0,895	0,055	0,153	0,016	0,073
Right_Occipital_Lobe	0,247	0,895	0,056	0,153	0,027	0,073
Left_Occipital_Lobe	0,247	0,895	0,046	0,153	0,034	0,073
Temporal_Lobe	0,247	0,895	0,025	0,153	0,007	0,073
Right_Temporal_Lobe	1,000	1,000	0,271	0,271	0,039	0,073
Left_Temporal_Lobe	0,126	0,895	0,085	0,160	0,047	0,078

Appendix B

Longitudinal evolution of wPLI

This appendix summarizes the findings regarding the longitudinal analyses performed to investigate the evolution of wPLI following VNS implantation and its relationship with clinical response.

Figures B.1–B.3 present heatmaps of FDR-corrected p-values obtained for the interaction between the six-months visit (V5) and clinical response. Each heatmap summarizes all combinations of the seven frequency bands and the fifteen predefined brain regions. Numerical values within each cell correspond to FDR-corrected p-values, while colors indicate significance levels. Blue cells represent non-significant results, orange cells indicate significant findings ($p_{FDR} < 0.05$), and red cells indicate highly significant findings ($p_{FDR} < 0.01$). These visualizations provide an overview of the spatial and spectral distribution of longitudinal associations.

Tables B.1–B.3 summarize statistically significant results identified by the mixed-effects analyses. Reported values include model estimates (β), standard errors, raw p-values, and FDR-corrected p-values. The singularity status of each model is additionally reported.

Figure B.4 illustrate significant longitudinal associations between hyperventilation wPLI values and binary response. Figures B.5–B.9 illustrate longitudinal associations between hyperventilation wPLI values and seizure reduction percentage for all significant region $\times$ band combinations. Each panel corresponds to a visit (V1, V3, V4, and V5) and displays individual patient observations together with fitted linear regression lines (dashed) and associated 95% confidence intervals (shaded areas). The coefficient of determination (R^2) is displayed for each visit.

Figure B.1: Heatmap — Visit V5 $\times$ VNS response interaction — Binary response — Hyperventilation

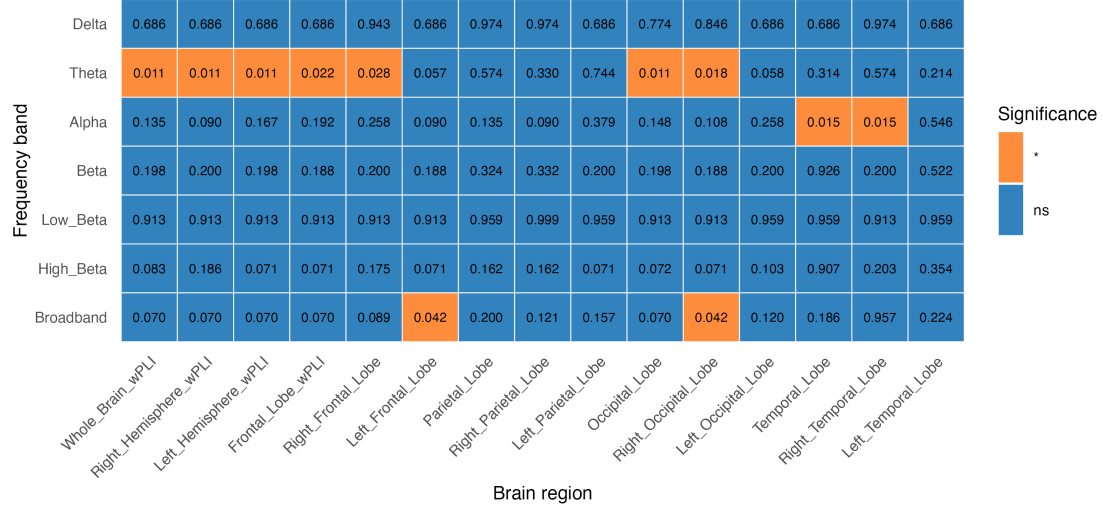


Table B.1: Significant mixed-effects model results — Binary response

Region	Band	singular	Estimate (β)	SE	raw p -value	p_{FDR}
Whole_Brain_wPLI	Theta	FALSE	0,081	0,023	0,002	0,011
Right_Hemisphere_wPLI	Theta	FALSE	0,084	0,025	0,003	0,011
Left_Hemisphere_wPLI	Theta	FALSE	0,086	0,026	0,003	0,011
Occipital_Lobe	Theta	FALSE	0,126	0,036	0,002	0,011
Temporal_Lobe	Alpha	FALSE	0,116	0,034	0,002	0,015
Right_Temporal_Lobe	Alpha	FALSE	0,196	0,053	0,001	0,015
Right_Occipital_Lobe	Theta	FALSE	0,113	0,038	0,006	0,018
Frontal_Lobe_wPLI	Theta	TRUE	0,116	0,042	0,009	0,022
Right_Frontal_Lobe	Theta	TRUE	0,126	0,048	0,013	0,028
Left_Frontal_Lobe	Broadband	TRUE	0,079	0,026	0,005	0,042
Right_Occipital_Lobe	Broadband	FALSE	0,072	0,024	0,006	0,042

Figure B.2: Heatmap — Visit V5 $\times$ VNS response interaction — Seizure reduction percentage — Hyperventilation



Table B.2: Significant mixed-effects model results — Seizure reduction percentage

Region	Band	singular	Estimate (β)	SE	raw p -value	p_{FDR}
Right_Temporal_Lobe	Alpha	FALSE	0,0036	0,0008	0,000	0,003
Right_Occipital_Lobe	Broadband	FALSE	0,0013	0,0004	0,002	0,032
Left_Hemisphere_wPLI	High_Beta	FALSE	0,0011	0,0004	0,006	0,036
Left_Frontal_Lobe	High_Beta	FALSE	0,0016	0,0006	0,009	0,036
Left_Parietal_Lobe	High_Beta	FALSE	0,0016	0,0006	0,010	0,036
Right_Occipital_Lobe	High_Beta	FALSE	0,0016	0,0005	0,004	0,036
Whole_Brain_wPLI	Broadband	FALSE	0,0009	0,0003	0,015	0,040
Right_Hemisphere_wPLI	Broadband	FALSE	0,0009	0,0004	0,012	0,040
Left_Hemisphere_wPLI	Broadband	FALSE	0,0010	0,0004	0,017	0,040
Frontal_Lobe_wPLI	Broadband	FALSE	0,0011	0,0004	0,023	0,040
Right_Frontal_Lobe	Broadband	FALSE	0,0011	0,0005	0,024	0,040
Left_Frontal_Lobe	Broadband	TRUE	0,0013	0,0004	0,008	0,040
Left_Parietal_Lobe	Broadband	TRUE	0,0010	0,0004	0,017	0,040
Occipital_Lobe	Broadband	FALSE	0,0010	0,0004	0,024	0,040
Whole_Brain_wPLI	Theta	FALSE	0,0012	0,0004	0,006	0,044
Right_Hemisphere_wPLI	Theta	FALSE	0,0013	0,0004	0,006	0,044
Left_Hemisphere_wPLI	Theta	FALSE	0,0013	0,0005	0,010	0,045
Right_Frontal_Lobe	Theta	TRUE	0,0021	0,0008	0,012	0,045

Figure B.3: Heatmap — Visit V5 $\times$ VNS response interaction — CRRS — Hyperventilation

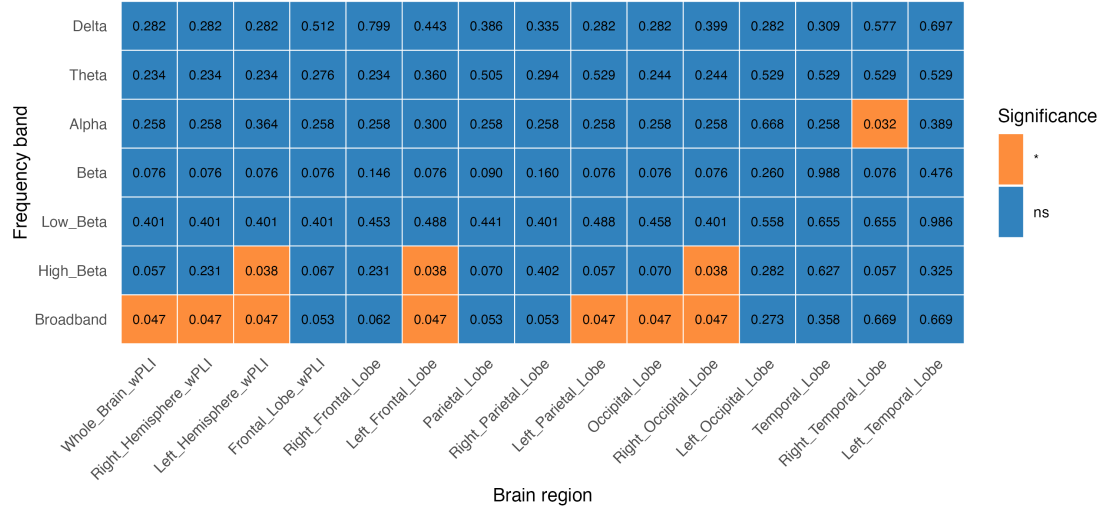


Table B.3: Significant mixed-effects model results — CRRS

Region	Band	singular	Estimate (β)	SE	raw p -value	p_{FDR}
Right_Temporal_Lobe	Alpha	FALSE	0,018	0,005	0,002	0,032
Left_Hemisphere_wPLI	High_Beta	FALSE	0,006	0,002	0,006	0,038
Left_Frontal_Lobe	High_Beta	FALSE	0,009	0,003	0,008	0,038
Right_Occipital_Lobe	High_Beta	FALSE	0,009	0,003	0,006	0,038
Whole_Brain_wPLI	Broadband	FALSE	0,005	0,002	0,014	0,047
Right_Hemisphere_wPLI	Broadband	FALSE	0,005	0,002	0,013	0,047
Left_Hemisphere_wPLI	Broadband	FALSE	0,005	0,002	0,022	0,047
Left_Frontal_Lobe	Broadband	TRUE	0,006	0,003	0,021	0,047
Left_Parietal_Lobe	Broadband	TRUE	0,006	0,002	0,018	0,047
Occipital_Lobe	Broadband	FALSE	0,006	0,002	0,019	0,047
Right_Occipital_Lobe	Broadband	FALSE	0,007	0,002	0,004	0,047

Figure B.4: Longitudinal evolution of wPLI values during hyperventilation recordings following VNS implantation according to binary clinical response. Mean values across visits are shown separately for responders (blue) and non-responders (red). Error bars indicate standard errors of the mean.

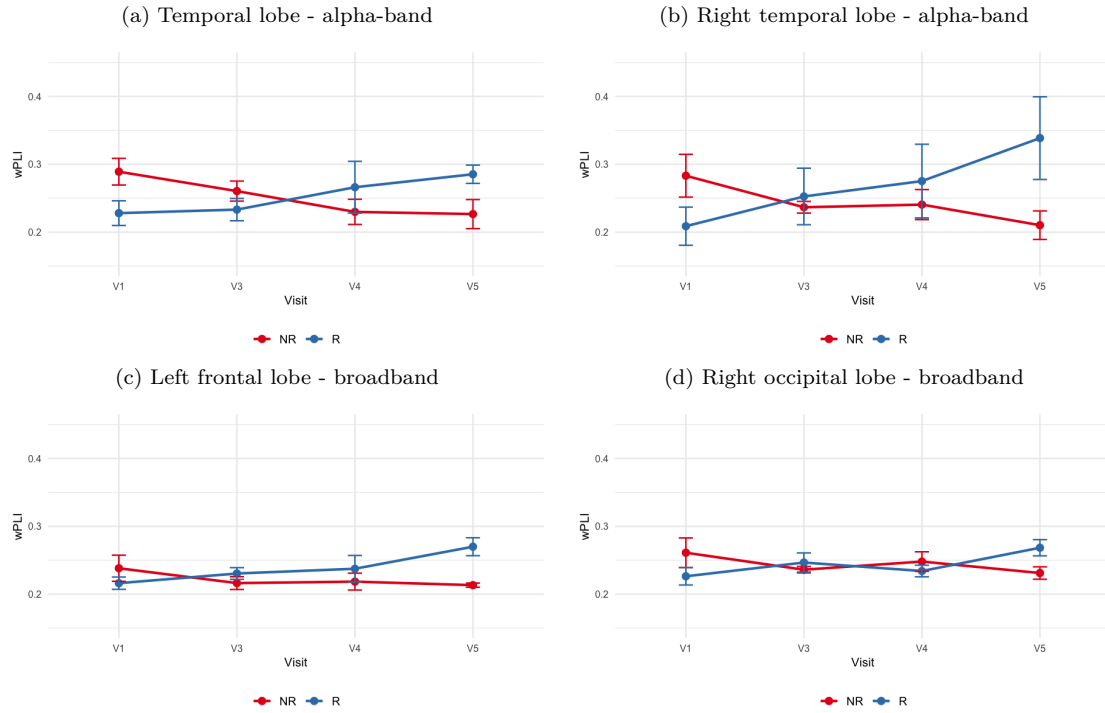


Figure B.5: Longitudinal associations between hyperventilation wPLI values and seizure reduction percentage (High-beta band).

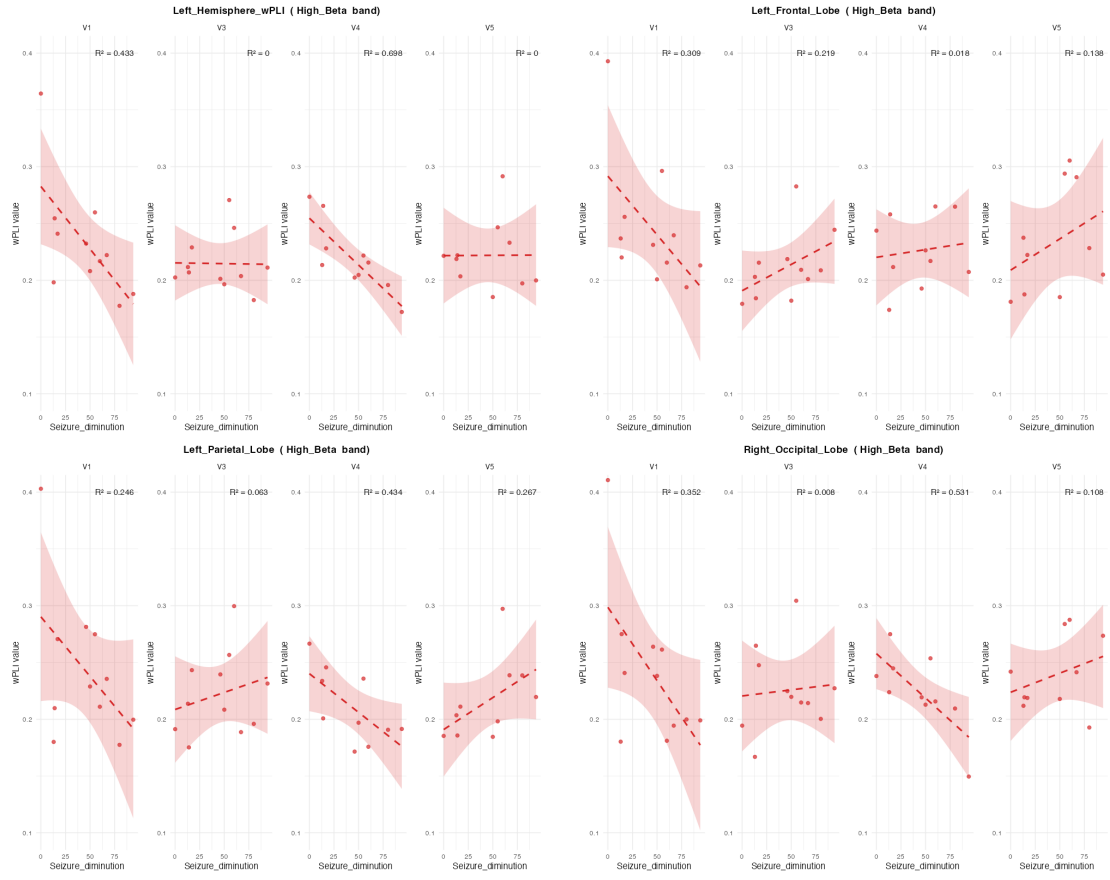


Figure B.6: Longitudinal associations between hyperventilation wPLI values and seizure reduction percentage (Broadband band, part I).

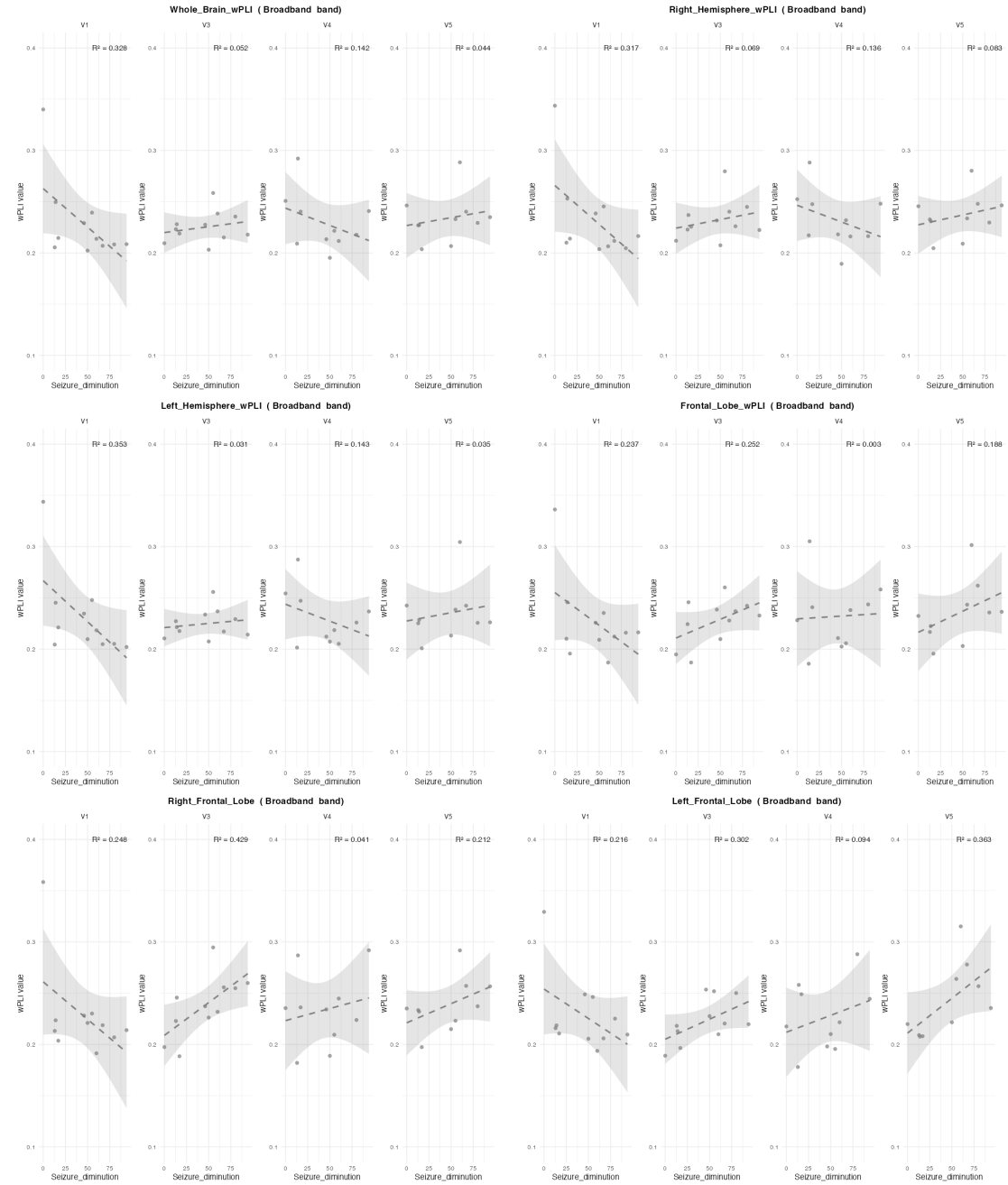


Figure B.7: Longitudinal associations between hyperventilation wPLI values and seizure reduction percentage (Broadband band, part II).

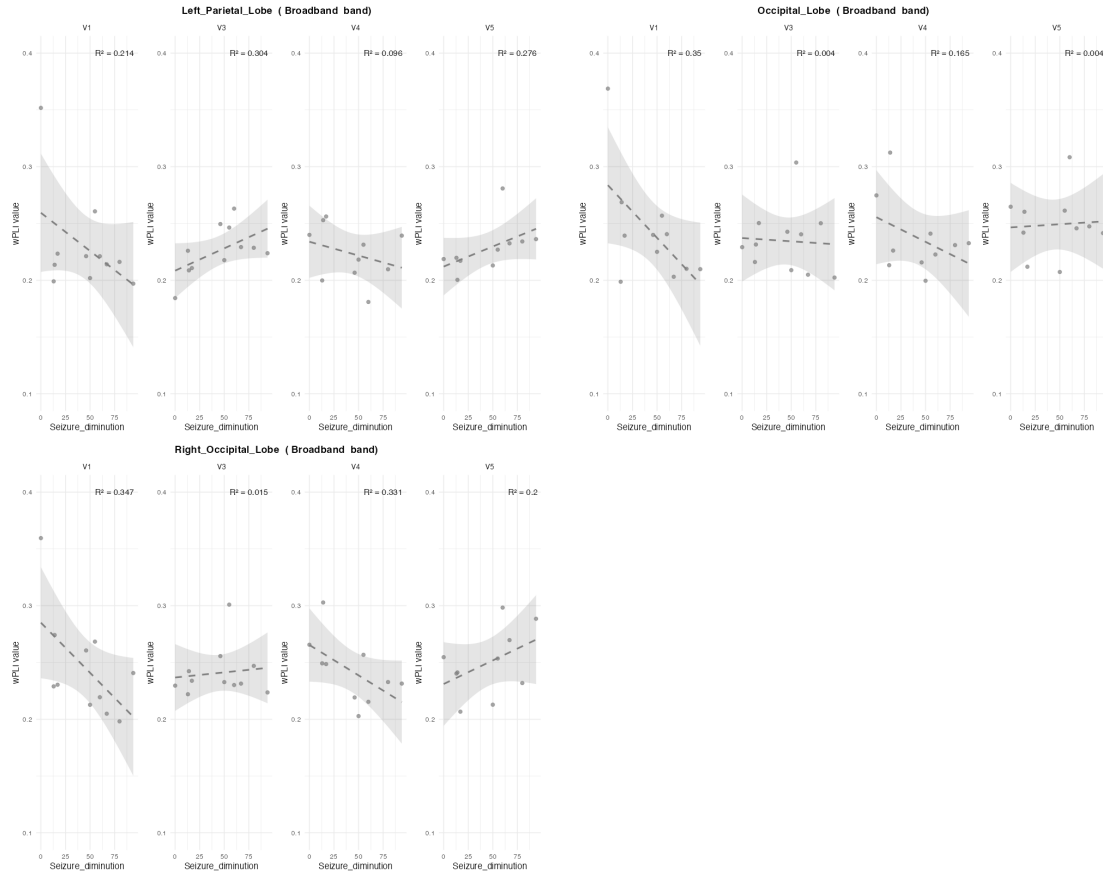


Figure B.8: Longitudinal associations between hyperventilation wPLI values and seizure reduction percentage (Theta band).

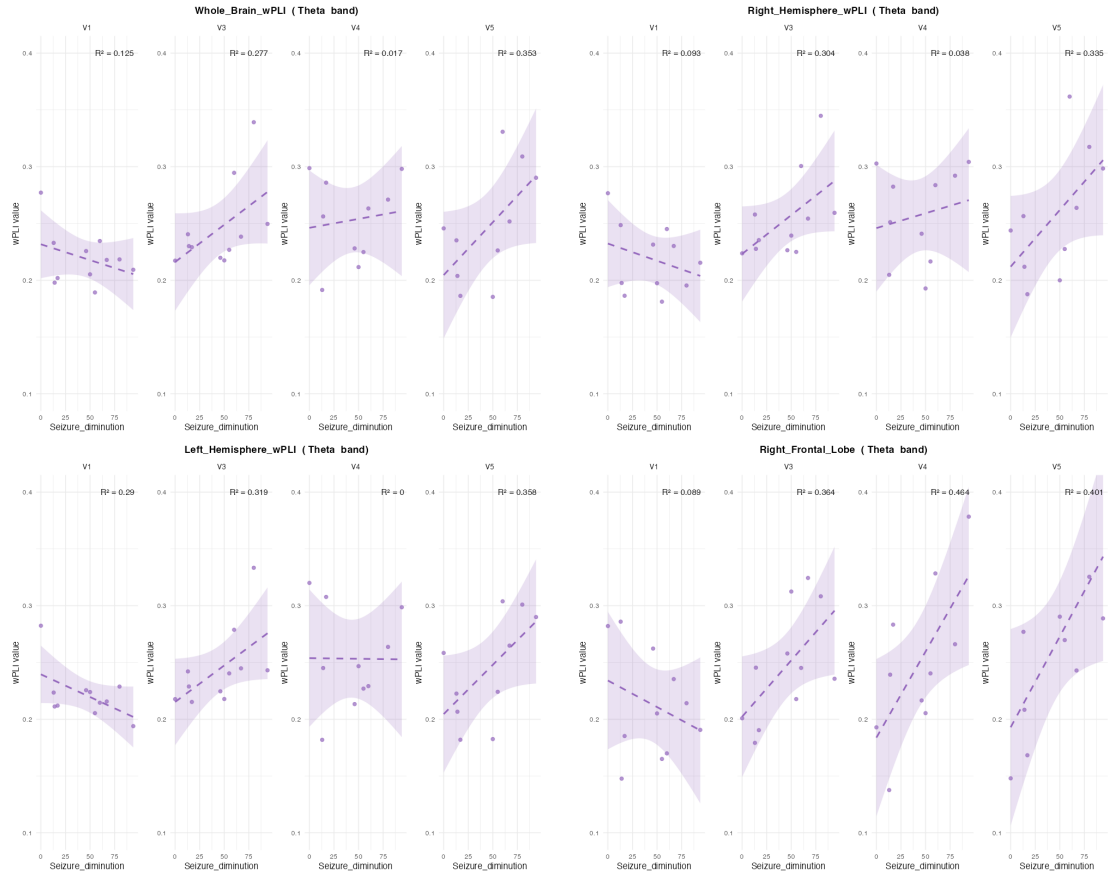
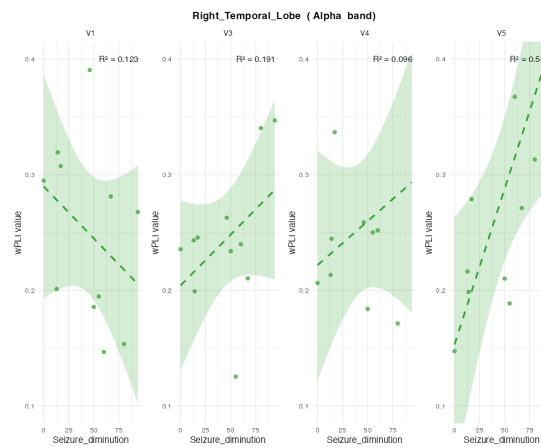


Figure B.9: Longitudinal associations between hyperventilation wPLI values and seizure reduction percentage (Alpha band).



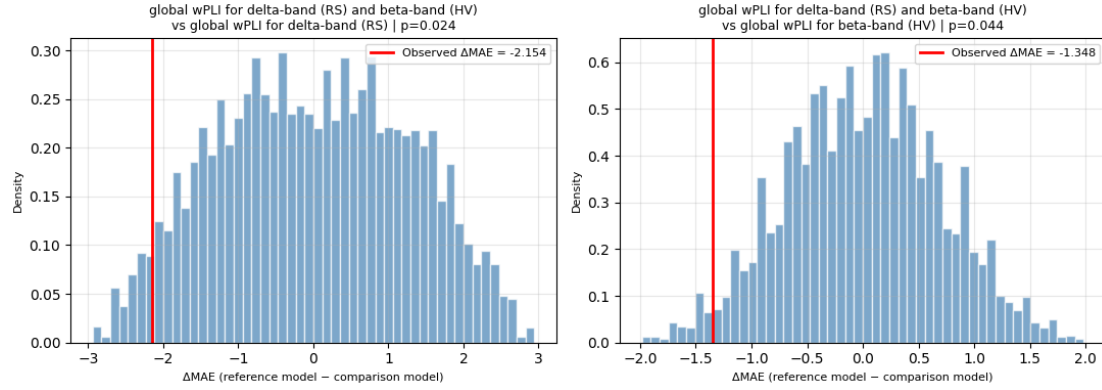
Appendix C

Permutation-test distributions

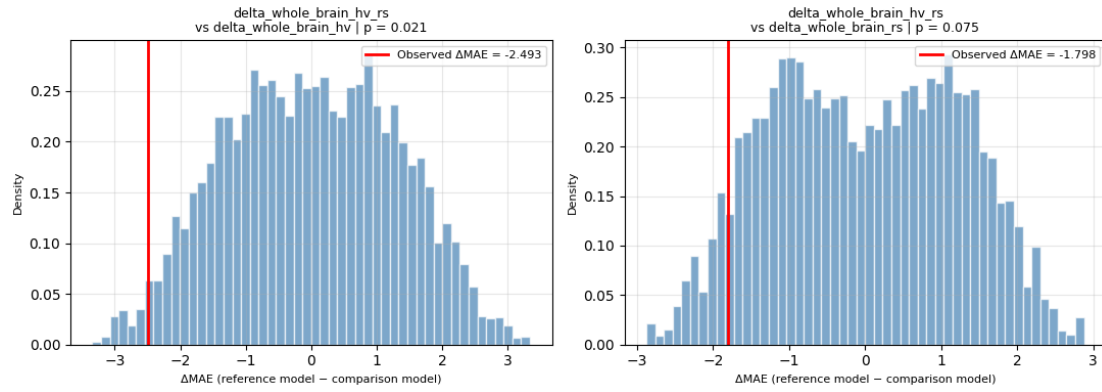
Figure C.1 represents paired permutation-test distributions comparing the prediction errors of the three selected models with those of their corresponding single-condition models. For each patient, absolute prediction errors from the selected model and the comparison model were randomly exchanged to generate the null distribution of ΔMAE , defined as the difference between the mean absolute error of the selected model and that of the corresponding single-condition model ($\text{MAE}_{\text{selected}} - \text{MAE}_{\text{single}}$). Negative observed ΔMAE values therefore indicate lower prediction error for the selected model relative to the comparison model.

Figure C.1: Paired permutation-test distributions for the three selected models. For each subfigure, comparisons with the resting-state-only model are shown on the left and comparisons with the hyperventilation-only model on the right. Red vertical lines indicate the observed Δ MAE values.

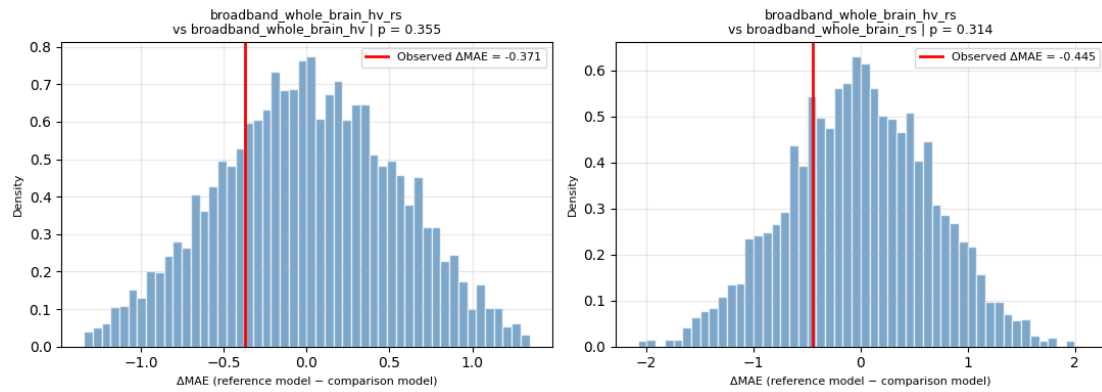
(a) Combined delta-beta model



(b) Delta-band model



(c) Broadband model



UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl