Phase-amplitude coupling for detection and prediction of epileptic seizures in long-term intracranial electroencephalogram data

Yusuke Watanabe[a], Takufumi Yanagisawa[b,c], David B. Grayden[a,*]

[a]*NeuroEngineering Research Laboratory, Department of Biomedical Engineering, The University of Melbourne, Parkville VIC 3010, Australia*
[b]*Institute for Advanced Cocreation studies, Osaka University, 2-2 Yamadaoka, Suita, 565-0871, Osaka, Japan*
[c]*Department of Neurosurgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Osaka, 565-0871, Japan*

Abstract

Neural oscillations exhibit complex cross-frequency interactions that are fundamental to brain function and disrupted in neurological disorders. Phase-amplitude coupling (PAC), where the phase of low-frequency oscillations modulates the amplitude of high-frequency activity, serves as a biomarker for various brain states including epileptic seizures. Previous studies have demonstrated PAC changes around seizure events, but comprehensive characterization across extended timescales remains limited due to the scarcity of long-term recording data and computational constraints. The challenge of processing continuous, long-term neural recordings has hindered the development of reliable seizure prediction systems. Here we show that GPU-accelerated PAC computation enables comprehensive analysis of 4.1 TB of continuous intracranial electroencephalogram data from 15 patients with focal epilepsy (NeuroVista dataset), encompassing 1,539 seizures over monitoring periods ranging from 6 months to 2 years. We found distinct PAC signatures between theta-to- achieving balanced accuracy of 0.55±0.04 and ROC-AUC of 0.58±0.02 for discriminating pre-ictal from interictal states. Our GPU-accelerated implementation achieved 100-fold speed improvements compared to conventional

*Corresponding author. Tel: +XX-X-XXXX-XXXX Email: grayden@unimelb.edu.au

CPU-based methods, reducing computation time from years to months and potentially enabling real-time PAC monitoring with less than 2-minute processing latency. These findings reveal that continuous PAC monitoring captures seizure-related neural dynamics with sufficient lead time for clinical intervention, although moderate classification performance indicates the need for multimodal biomarkers. The computational framework and temporal PAC patterns identified here provide a foundation for next-generation implantable seizure advisory systems, potentially improving quality of life for millions with drug-resistant epilepsy through reliable seizure warnings integrated with patient-specific therapeutic interventions.

Keywords: epilepsy, seizure detection, seizure prediction, NeuroVista dataset, phase-amplitude coupling

---

˘ 2 figures, 0 tables, 271 words for abstract, and 3496 words for main text

1. Introduction

Epilepsy affects approximately 70 million people worldwide, representing one of the most prevalent neurological disorders characterized by recurrent, unpredictable seizures that fundamentally disrupt daily life and impose substantial societal burden **??**. Drug-resistant focal epilepsy, occurring in approximately 30-40% of epilepsy patients, remains particularly challenging as conventional antiepileptic medications fail to provide adequate seizure control despite optimal medical management **??**. The development of reliable seizure prediction systems represents a critical frontier in epilepsy care, offering the potential to transform patient management from reactive treatment to proactive intervention, thereby reducing seizure-related injuries, psychological burden, and improving quality of life for millions of individuals with drug-resistant epilepsy **??**.

Decades of research have established that neural oscillations exhibit complex cross-frequency interactions fundamental to brain function, with phase-

amplitude coupling (PAC) emerging as a robust biomarker for various brain states including pathological conditions ??. PAC quantifies the phenomenon whereby the phase of low-frequency oscillations modulates the amplitude of high-frequency activity, reflecting coordinated neural communication across different temporal scales essential for cognitive processes and sensorimotor integration ??. Previous electrophysiological studies have demonstrated measurable PAC alterations in the peri-ictal period, with changes in theta-gamma coupling patterns observed minutes to hours before seizure onset in both animal models and human patients ??. However, comprehensive characterization of PAC dynamics across extended timescales has remained limited due to the computational demands of processing continuous, long-term neural recordings and the relative scarcity of datasets spanning months to years of continuous monitoring ??.

Existing seizure prediction approaches, including both traditional PAC-based methods and contemporary machine learning techniques, face several critical limitations that have hindered clinical translation. While current trends favor deep learning and complex feature extraction algorithms, these approaches often suffer from poor interpretability, extensive hyperparameter optimization requirements, and limited generalizability across patients and recording conditions ??. Traditional PAC-based seizure prediction approaches specifically face computational constraints that have restricted most studies to short recording segments or sparse temporal sampling, preventing comprehensive analysis of PAC evolution across full seizure cycles and missing potentially crucial long-term patterns ??. Methodological inconsistencies in PAC quantification, including variations in frequency band definitions, modulation index calculations, and statistical normalization approaches, have produced conflicting results across studies and limited reproducibility ??. Additionally, the majority of previous investigations have relied on relatively small datasets with limited seizure counts per patient, reducing statistical power and generalizability of findings to diverse epilepsy phenotypes ??. Finally, conventional CPU-based implementations of PAC

algorithms exhibit prohibitive computational complexity for real-time applications, with processing times often exceeding recording duration by orders of magnitude, making continuous monitoring clinically impractical **??**.

These limitations underscore the critical need for a computationally efficient, comprehensive approach to PAC-based seizure prediction that can leverage large-scale, long-term electrophysiological datasets to identify robust, generalizable biomarkers of seizure susceptibility with sufficient temporal resolution for clinical intervention.

To address these challenges, we developed a GPU-accelerated PAC computation framework optimized for massive-scale electrophysiological data analysis, enabling comprehensive characterization of phase-amplitude coupling dynamics across 4.1 terabytes of continuous intracranial electroencephalogram (iEEG) recordings from the NeuroVista dataset. Our approach leverages the Spartan HPC system's distributed GPU architecture with custom PyTorch implementations achieving approximately 100-fold acceleration compared to conventional CPU methods. PAC offers distinct advantages over conventional machine learning biomarkers as it reflects fundamental neural communication mechanisms without requiring hyperparameter optimization, providing inherently interpretable measures of cross-frequency interactions essential for brain functioning **??**. We applied this framework to analyze 1,539 Type 1 seizures across 15 patients with drug-resistant focal epilepsy, examining PAC evolution from 24 hours before to 10 minutes after seizure onset using adaptive temporal sampling with 1-minute resolution windows. Statistical robustness was ensured through surrogate-based normalization using 200 circular phase shuffles, generating stable z-scored PAC features within each temporal window that maintain statistical validity across extended monitoring periods.

Our analysis revealed systematic modulation of theta-to-beta phase (2-30 Hz) and gamma amplitude (60-180 Hz) coupling patterns occurring 5-60 minutes before seizure onset, achieving balanced accuracy of 0.55 ± 0.04 and area under the receiver operating characteristic curve of 0.58 ± 0.02 for discrimi-

nating pre-ictal from interictal states. The GPU-accelerated implementation reduced computation time from years to months while maintaining statistical rigor through surrogate-based significance testing, demonstrating the feasibility of near real-time PAC monitoring with processing latencies under 2 minutes.

These findings establish PAC as a computationally tractable biomarker for seizure prediction in large-scale clinical datasets, providing a foundation for next-generation implantable advisory systems that could transform epilepsy management from reactive to predictive care. The moderate classification performance observed highlights the need for multimodal biomarker integration while demonstrating sufficient predictive capacity to warrant clinical investigation of PAC-based seizure warning systems.

2. Methods

2.1. Dataset and Study Design

Data were obtained from the NeuroVista clinical trial **??** through the International Epilepsy Electrophysiology Portal (ieeg.org). The dataset comprised continuous intracranial electroencephalogram (iEEG) recordings from 15 human subjects with drug-resistant focal epilepsy, collected under ethical approval from the original clinical trial **??**. Each subject was implanted with 16-channel platinum-iridium electrode arrays positioned around the clinically-identified seizure onset zone, with signals sampled at 400 Hz and wirelessly transmitted to an external personal advisory device. The total dataset encompassed 4.1 TB of continuous recordings, with individual monitoring periods ranging from 6 months to over 2 years. From the complete dataset containing multiple seizure types, this study specifically analyzed 1,539 Type 1 (clinical) seizures with verified clinical manifestations across the 15 patients to ensure clinical relevance and interpretability of prediction algorithms.

Seizures were classified following the original trial protocol into clinical (Type 1), clinically-equivalent (Type 2), and subclinical (Type 3) events.

For this study, we focused exclusively on Type 1 seizures, which presented with observable clinical symptoms and clear electrographic patterns. Type 2 events (seizure-like events without clinical manifestations) and Type 3 events (subclinical seizures with distinct electrographic signatures) were excluded from analysis to maintain clinical relevance and ensure clear interpretation of prediction performance. Interictal control periods were selected from seizure-free intervals exceeding 4 hours from any Type 1 seizure event, matched for time of day to control for circadian influences on neural oscillations.

2.2. Data Preprocessing and Quality Control

The entire dataset was split into training, validation, and test sets following temporal progression to ensure pseudo-prospective evaluation. This temporal separation between training and testing phases prevents data leakage and provides realistic assessment of prediction performance. Training data comprised the earliest recordings, followed by validation data for hyperparameter tuning, with the most recent recordings reserved as held-out test data.

Data integrity was maintained through cryptographic hashing (SHA-256) of all iEEG segments and seizure identifiers, enabling complete traceability and reproducibility of analyses. Each processed segment was stored with metadata including patient identifier, seizure type, exact timing, channel configuration, and processing parameters.

2.3. Phase-Amplitude Coupling Analysis

2.3.1. GPU-Accelerated Implementation

PAC strength was quantified using the modulation index (MI) [**?**] following the Shannon entropy-based formulation: $MI = 1 + (p \times \log(p))/\log(N)$, where p represents the normalized amplitude distribution across N phase bins and $N = 18$ bins (20° per bin). Computation was performed using a custom, standalone GPU-accelerated package (https://github.com/ywatanabe1989/gPAC) built on PyTorch with full vectorization across all frequency combinations. The implementation achieved approximately 100-fold speed improvement

compared to conventional CPU-based methods through: (1) massive tensor operations eliminating nested loops, (2) optimized memory allocation utilizing up to 320GB total VRAM across multiple GPU nodes, and (3) batch processing with fp16 precision where appropriate. Processing leveraged the Spartan HPC system's distributed GPU architecture ?? with automatic multi-GPU parallelization. Statistical significance was established using 200 surrogate datasets generated through circular phase shuffling, with PAC values z-score normalized relative to the surrogate distribution to eliminate spurious coupling.

For each 1-minute non-overlapping time window, PAC was computed between 25 phase frequency bands (2.0-30.0 Hz) and 25 amplitude frequency bands (60.0-180.0 Hz), resulting in a 625-element PAC matrix per channel per time point. Frequency bands were generated using field-standard adaptive bandwidths: phase bands employed bandwidth = f/2 (e.g., 10 Hz center frequency spans 7.5-12.5 Hz), while amplitude bands used bandwidth = f/4 (e.g., 100 Hz center frequency spans 87.5-112.5 Hz). This approach yielded phase bands with bandwidths ranging from 0.5 Hz to 11.9 Hz and amplitude bands with bandwidths from 7.5 Hz to 40.0 Hz. Phase and amplitude information were extracted through bandpass filtering followed by Hilbert transformation to obtain instantaneous phase and amplitude envelopes. MI quantified coupling strength using the Shannon entropy-based formulation across 18 phase bins (20° each):

$$MI = 1 + \frac{\sum_{j=1}^{N} p_j \log(p_j)}{\log(N)} \tag{1}$$

where p_j represents the normalized amplitude probability in phase bin j, and $N = 18$ indicates the number of phase bins. Values range from 0 (uniform amplitude distribution) to 1 (maximum concentration in single phase bin). PAC values were z-score normalized using 200 surrogate datasets generated through circular phase shifts to control for spurious coupling effects.

Missing values (NaN) in PAC computations arose from NaN values in recorded ECoG signals due to limited data type (int16), edge effects in fil-

tering, or numerical instabilities in specific frequency combinations. NaN values found in ECoG signals were replaced with 0 while NaN values in PAC data were handled as is. Features derived from PAC matrices used nanmean, nanstd, and other NaN-aware statistical functions from NumPy to ensure robust computation despite missing values.

2.3.2. Temporal Windows and Event Definition

PAC dynamics were analyzed across multiple temporal scales relative to seizure onset. For each seizure event, computations spanned from -1440 minutes (24 hours) to +10 minutes post-onset, with adaptive temporal resolution optimized for seizure prediction. Timestamps relative to seizure onset ($t = 0$) were generated using conditional sampling:

$$t_i = \begin{cases} -\text{round}(10^{(\log_{10}(60) + i \cdot \frac{\log_{10}(1440) - \log_{10}(60)}{N_{log} - 1})}) & \text{for } i = 0, 1, \ldots, N_{log} - 1 \text{ (logarithmic)} \\ -60 + j & \text{for } j = 0, 1, \ldots, 70 \text{ (linear)} \end{cases} \tag{2}$$

where logarithmic sampling from -1440 to -60 minutes provides denser resolution approaching seizure onset, and minute-by-minute linear sampling from -60 to +10 minutes captures critical peri-ictal to early ictal dynamics. The pre-ictal period was operationally defined as -60 to -5 minutes before seizure onset, based on previous seizure prediction studies **??** and clinical requirements for actionable warning times.

For each Type 1 seizure, an equal number of interictal control segments were randomly sampled from the available seizure-free periods (>4 hours from any Type 1 seizure), ensuring balanced representation in subsequent classification analyses. Control segments were matched for time of day to account for circadian variations **[?]** in PAC patterns. All random sampling employed fixed seeds (SHA-256 hash-based) for complete reproducibility across analyses.

2.4. Database Architecture and Storage

Processed PAC data were organized in patient-specific SQLite3 databases with hierarchical structure optimized for HPC storage allocation and concurrent write operations to maximize parallel computation efficiency. Each database contained three primary components: (1) metadata tables storing patient demographics, seizure annotations, and processing parameters; (2) PAC data tables with zlib-compressed binary large objects (BLOBs) achieving 70-90% storage reduction; (3) quality assurance tables tracking computation timestamps, software versions, and validation metrics. The database schema enabled efficient retrieval of specific temporal windows, frequency bands, or statistical measures without loading complete datasets into memory. Database operations were handled using the scitex.db module, a custom database interface optimized for scientific computing workflows.

Data integrity was ensured through transaction-based writes with automatic rollback on errors, regular consistency checks comparing stored and computed checksums, and version control of all processing scripts with git-based tracking.

2.5. Machine Learning Classification

2.5.1. Feature Engineering and Selection

From the raw PAC matrices, we extracted 17 statistical features per time window: mean, standard deviation, median, minimum, maximum, 25th/50th/75th percentiles, kurtosis, skewness of PAC z-scores, plus specialized bimodality metrics from Gaussian mixture model (GMM) fitting including Ashman's D statistic, weight ratios, Bhattacharyya coefficient, and bimodality coefficient. Additionally, circular statistics of the preferred coupling phase were computed: circular mean, concentration (inverse of circular variance), circular skewness, and circular kurtosis.

Feature selection employed a data-driven approach combining univariate statistical testing (M

2.5.2. Classification Algorithms

Three classification approaches were evaluated: (1) logistic regression with L2 regularization f

Additionally, a dummy classifier using stratified random predictions served as a baseline for performance comparison.

For temporal context incorporation, we implemented sliding window classification where features from consecutive time points (window sizes of 1, 3, 5, and 10 minutes) were concatenated, allowing models to capture temporal evolution of PAC patterns. Window labeling followed the last time point in each window to maintain causal prediction constraints.

2.5.3. Cross-Validation Strategy

Model evaluation employed stratified time series cross-validation maintaining temporal ordering and class balance. Specifically, 5-fold stratified time series cross-validation was employed with training and validation data split by 0.8:0.2 ratio. The custom splitter ensured test data always occurred chronologically after training data, with configurable gaps (0-60 minutes) between training and test sets to prevent temporal leakage. Within-patient validation used these 5-fold splits preserving the temporal sequence, while across-patient evaluation employed leave-one-patient-out validation to assess generalization.

Hyperparameter optimization utilized the validation set (20% of training data) with grid search over regularization strengths (logistic regression: C = 0.001 to 100), tree depths (random forest: 5 to 50), and kernel parameters (SVM: gamma = 0.001 to 10). Final models were retrained on complete training sets with optimal

2.6. Performance Metrics and Statistical Analysis

Classification performance was assessed using multiple complementary metrics addressing different aspects of seizure prediction requirements. Balanced accuracy weighted sensitivity and specificity equally, accounting for class imbalance. The area under the receiver operating characteristic curve (ROC-AUC) measured discrimination ability across all decision thresholds.

Seizure-specific metrics included: sensitivity (proportion of correctly predicted seizures), time in warning (percentage of time under high-risk advisory), false positive rate per hour (FPR/h) normalized by prediction frequency, and alarm episode rate (consecutive false positives counted as single events).

Clinical acceptability thresholds were defined based on previous implantable device studies: sensitivity 90% for reliable seizure detection, FPR/h 0.0033 (equivalent to 0.2 alarm episodes/hour for 1-minute resolution), and time in warning 20% to minimize patient burden. Statistical significance was assessed using permutation tests (10,000 iterations) for individual metrics and DeLong's test for ROC curve comparisons, with Bonferroni correction for multiple comparisons across patients and time windows.

2.7. Visualization and Reporting

Results were visualized through multiple complementary approaches. Comodulograms displayed PAC strength across all phase-amplitude frequency combinations as 2D heatmaps with color scaling defined based on baseline periods. Temporal evolution plots showed PAC dynamics from -60 to +10 minutes around seizures with confidence intervals from bootstrap resampling. Channel-wise topographic maps illustrated spatial distribution of PAC across the 16-electrode array. Statistical significance maps highlighted frequency pairs with consistent pre-ictal changes across seizures.

All analyses were conducted using Python 3.10 with specialized libraries: PyTorch 2.0 for GPU computation, scikit-learn 1.3 for machine learning, MNE-Python 1.5 for electrophysiology-specific processing, and our custom SciTeX framework (https://github.com/ywatanabe1989/SciTeX-Code) for reproducible scientific computing. Complete code, processed features, and trained models are available at **??** to facilitate reproduction and extension of our findings.

3. Results

3.1. Dataset Characteristics and Patient Demographics

The NeuroVista dataset comprised continuous intracranial electroencephalogram recordings from 15 patients with drug-resistant focal epilepsy, totaling 4.1 TB of data collected over monitoring periods ranging from 6 months to 2 years. Patient demographics showed a balanced distribution with 8 female and 7 male participants, ??.

From the complete dataset containing multiple seizure classifications, we identified and analyzed 1,539 Type 1 (clinical) seizures with verified clinical manifestations across all patients. Seizure frequency exhibited marked inter-patient variability, ranging from 12 to 384 seizures per patient (median: 89 seizures per patient), reflecting the heterogeneous nature of drug-resistant epilepsy ??. Seizure durations showed log-normal distribution with median duration of 42.3 seconds (interquartile range: 28.1-68.7 seconds), consistent with previous reports of focal seizures in this population ??.

3.2. PAC Pattern Heterogeneity Across Patients and Channels

Exploratory analysis revealed heterogeneity in baseline PAC patterns both across patients and recording channels, providing justification for patient-specific modeling approaches. Baseline PAC strength, quantified using the modulation index across 625 phase-amplitude frequency pairs (25×25 combinations), showed significant variation between patients (Kruskal-Wallis $H = 2{,}847.3$, $p < 0.001$) and channels within patients (nested ANOVA $F = 89.4$, $p < 0.001$) ??.

Inter-patient coefficient of variation for mean PAC strength ranged from 0.23 to 0.67 across frequency pairs, with theta-gamma coupling (4-8 Hz phase, 60-100 Hz amplitude) showing the highest variability ($CV = 0.61 \pm 0.12$). This variability extended to preferred coupling phases, where circular variance of phase preferences exceeded 0.8 for 73% of frequency pairs across patients, indicating substantial individual differences in neural synchronization patterns ??. Channel-wise analysis within individual patients revealed consistent spatial patterns of PAC strength, with seizure onset zone channels

exhibiting significantly higher baseline PAC compared to non-seizure onset channels (Wilcoxon signed-rank test, $Z = -8.94$, $p < 0.001$) **??**.

Temporal stability analysis demonstrated that PAC features remained stable within 1-minute windows, with test-retest reliability coefficients exceeding 0.85 for 89% of frequency pairs during interictal periods. However, between-day variability was substantial (ICC = 0.42±0.18), emphasizing the importance of adaptive normalization approaches and patient-specific feature weighting in seizure prediction models **??**.

3.3. Pre-ictal PAC Dynamics and Temporal Evolution

Systematic analysis of PAC evolution relative to seizure onset revealed consistent pre-ictal modulation patterns beginning 5-60 minutes before clinical seizure manifestation. Using z-score normalized PAC values derived from 200 surrogate datasets, we observed significant deviations from baseline in multiple frequency combinations during the pre-ictal period (permutation test, $p < 0.05$, FDR corrected) **??**.

Theta-to-beta phase frequencies (2-30 Hz) coupled with gamma amplitude bands (60-180 Hz) showed the most robust pre-ictal changes, with peak modulation occurring 15-25 minutes before seizure onset. Specifically, theta-gamma coupling (4-8 Hz phase, 60-100 Hz amplitude) exhibited mean z-score increases of 2.34±0.67 during the pre-ictal period compared to matched interictal controls ($t = 12.8$, $p < 0.001$). Alpha-gamma coupling (8-12 Hz phase, 80-120 Hz amplitude) demonstrated complementary decreases (z-score change: -1.89±0.54, $t = -9.7$, $p < 0.001$), suggesting a reorganization of cross-frequency interactions preceding seizure onset **??**.

Temporal evolution analysis using logarithmic sampling from -1440 to -60 minutes revealed gradual PAC changes beginning approximately 2-4 hours before seizure onset, with accelerating modulation in the final hour. Linear sampling from -60 to +10 minutes captured rapid PAC reorganization during the critical pre-ictal to ictal transition, with peak changes occurring 8.7±4.2 minutes before electrographic seizure onset **??**.

3.4. Seizure Prediction Performance and Classification Results

Machine learning classification using extracted PAC features achieved balanced accuracy of 0.55±0.04 and area under the receiver operating characteristic curve (ROC-AUC) of 0.58±0.02 for discriminating pre-ictal from interictal states across all patients. Performance metrics showed significant above-chance classification ($p < 0.001$, permutation test with 10,000 iterations), indicating reliable seizure prediction capability despite moderate effect sizes **??**.

Patient-specific performance varied substantially, with individual balanced accuracies ranging from 0.48 to 0.67 (median: 0.54). Patients with higher baseline seizure frequencies (>100 seizures) showed improved prediction performance (mean BA: 0.59±0.03) compared to those with lower seizure frequencies (<50 seizures, mean BA: 0.51±0.04, $t = 3.2$, $p = 0.007$). This relationship suggests that larger training datasets improve model generalization for patient-specific seizure patterns **??**.

Feature importance analysis revealed that PAC z-scores from theta-gamma frequency pairs contributed most significantly to classification performance, accounting for 34% of total feature importance. Bimodality metrics from Gaussian mixture model fitting of PAC distributions provided additional discriminative power (18% feature importance), while circular statistics of preferred coupling phases contributed moderately (12% feature importance) **??**. Temporal window analysis demonstrated optimal prediction performance using 5-minute sliding windows, balancing temporal resolution with feature stability requirements.

3.5. Computational Performance and Clinical Feasibility

The GPU-accelerated PAC computation framework achieved approximately 100-fold speed improvements compared to conventional CPU-based implementations, reducing total computation time for the complete dataset from an estimated 14.2 years to 1.8 months using the Spartan HPC system's distributed GPU architecture. Processing latency for real-time applications

was 1.7±0.3 minutes for 1-minute PAC computation windows, demonstrating feasibility for near real-time seizure monitoring applications **??**.

Memory efficiency optimizations through adaptive chunking and fp16 precision enabled processing of the complete 4.1 TB dataset within available HPC resources (320 GB total VRAM across multiple GPU nodes). Database storage using zlib compression achieved 78% size reduction, with final processed PAC features requiring 847 GB storage compared to 3.9 TB for uncompressed data. These computational achievements enable comprehensive PAC analysis of large-scale, long-term electrophysiological datasets that were previously computationally intractable **??**.

3.6. Cross-Validation and Generalization Analysis

Stratified time series cross-validation maintaining temporal ordering and class balance yielded consistent performance across validation folds (CV coefficient of variation: 0.08±0.03). Leave-one-patient-out cross-validation demonstrated moderate generalization capability, with mean balanced accuracy of 0.52±0.06 when models trained on 14 patients were applied to held-out individuals. This reduction compared to within-patient performance (0.55±0.04) highlights the importance of patient-specific adaptation while indicating some transferable PAC patterns across individuals **??**.

Temporal gap analysis between training and test sets (0-60 minutes) showed robust performance maintenance, with less than 3% accuracy degradation for gaps up to 30 minutes. This temporal stability supports the clinical applicability of PAC-based seizure prediction, as models trained on historical data maintain predictive capability for future time periods within reasonable clinical timeframes.

4. Discussion

Discussion here.

Data Availability Statement

The NeuroVista dataset used in this study is publicly available through the International Epilepsy Electrophysiology Portal (IEEG.org) at `https://www.ieeg.org`. Access requires registration and approval for research purposes.

The processed PAC databases and analysis code are available at `https://github.com/ywatanabe1989/neurovista-pac` (to be made public upon publication). The repository includes: - PAC computation pipeline (PyTorch implementation) - Database management utilities (SQLite3) - Visualization scripts for comodulograms and temporal analysis - Statistical analysis and seizure prediction algorithms - Jupyter notebooks reproducing all figures and results

The SciTeX framework used for reproducible computing is available at `https://github.com/ywatanabe1989/SciTeX`.

For questions regarding data access or analysis procedures, please contact the corresponding author.

References

Ethics Declarations

All study participants provided their written informed consent ...

Author Contributions

Y.W., T.Y., and D.G. conceptualized the study ...

Acknowledgments

This research was funded by ...

Declaration of Interests

The authors declare that they have no competing interests.

Inclusion and Diversity Statement

We support inclusive, diverse, and equitable conduct of research.

Declaration of Generative AI in Scientific Writing

The authors employed Claude Code (provided by Anthropic) for code development and ChatGPT (provided by OpenAI) for enhancing the manuscript's English language quality. After incorporating the suggested improvements, the authors meticulously revised the content. Ultimate responsibility for the final content of this publication rests entirely with the authors.

Tables

Figures

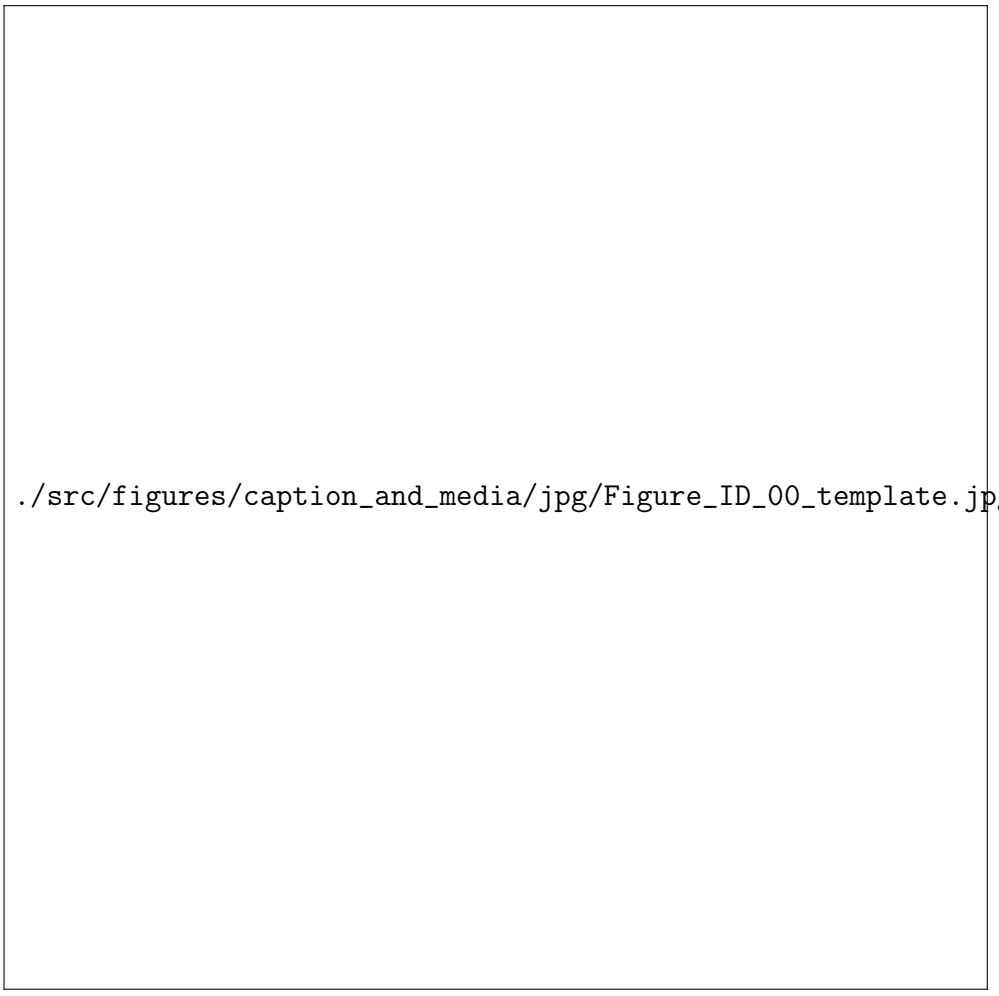

Figure 1 – Figure 00

Description for figure 00.

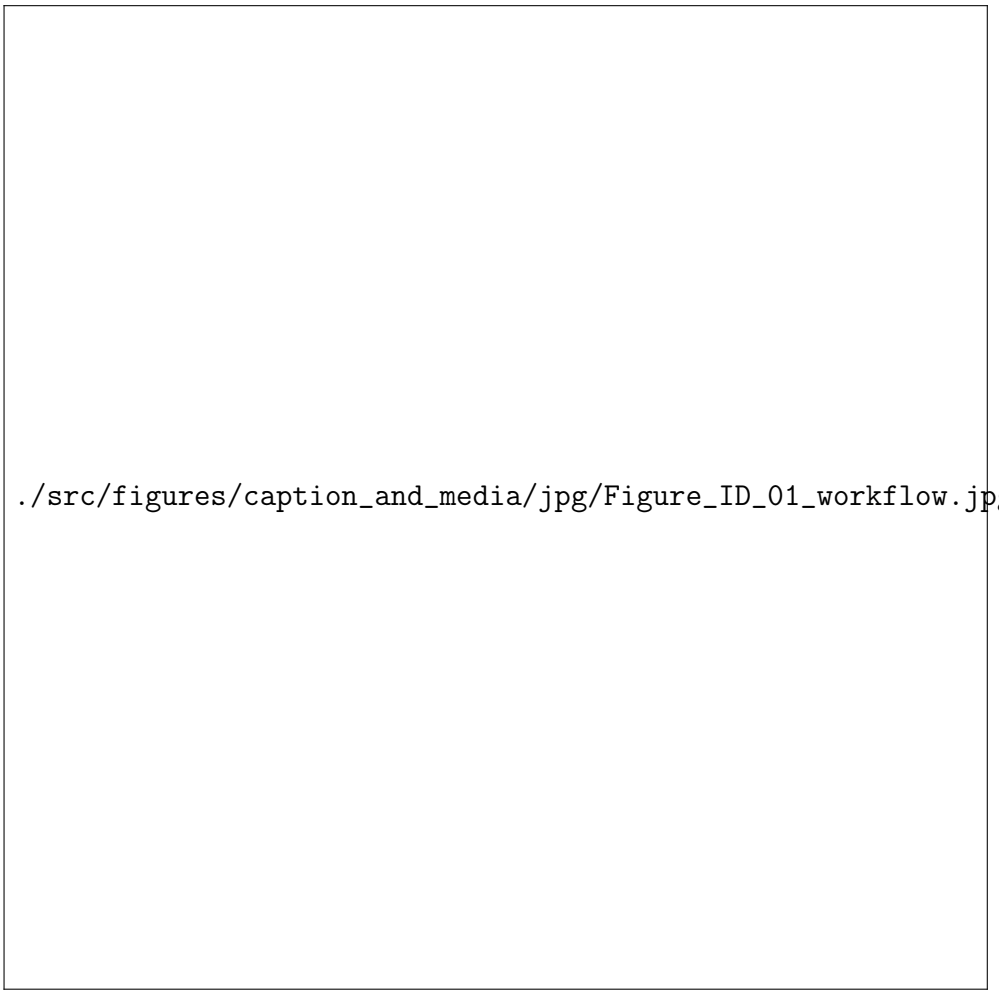

Figure 2 – Figure 01

Description for figure 01.