Title

Discovering EEG biomarkers of Lennox-Gastaut syndrome through unsupervised time-frequency analysis.

Journal

Epilepsia: Official journal of the International League Against Epilepsy, 66(2)

Authors

Hu, Derek

Pinto-Orellana, Marco

Rana, Mandeep

et al.

Publication Date

2025-02-01

DOI

10.1111/epi.18211

Peer reviewed

Received: 18 June 2024 | Revised: 19 November 2024 | Accepted: 20 November 2024

DOI: 10.1111/epi.18211

Epilepsia®

RESEARCH ARTICLE

Discovering EEG biomarkers of Lennox–Gastaut syndrome through unsupervised time–frequency analysis

Derek K. Hu[1,2,3] | **Marco A. Pinto-Orellana**[1] | **Mandeep Rana**[4] | **Linda Do**[4] | **David J. Adams**[4] | **Shaun A. Hussain**[5] | **Daniel W. Shrey**[4,6] | **Beth A. Lopour**[1]

[1]Department of Biomedical Engineering, University of California, Irvine, California, USA

[2]Department of Biomedical Engineering, California State University, Long Beach, California, USA

[3]Department of Computer Engineering and Computer Science, California State University, Long Beach, California, USA

[4]Division of Neurology, Children's Hospital Orange County, Orange, California, USA

[5]Division of Pediatric Neurology, University of California, Los Angeles, California, USA

[6]Department of Pediatrics, University of California, Irvine, California, USA

Correspondence
Beth A. Lopour, Department of Biomedical Engineering, University of California, 3120 Natural Sciences II, University of California, Irvine, CA 92697-2715, USA.
Email: beth.lopour@uci.edu

Funding information
LGS Foundation; John C. Hench Foundation

Abstract

Objective: The discovery and validation of electroencephalography (EEG) biomarkers often rely on visual identification of waveforms. However, bias toward visually striking events restricts the search space for new biomarkers, and low interrater reliability can limit rigorous validation. We present a data-driven approach to biomarker discovery called scalp EEG Pattern Identification and Categorization (s-EPIC), which enables automated, unsupervised identification of EEG waveforms. S-EPIC is validated on Lennox–Gastaut syndrome (LGS), an epilepsy that is difficult to diagnose and assess due to its variable presentation and insidious evolution of symptoms.

Methods: We retrospectively collected 10-min scalp EEG clips during non–rapid eye movement (NREM) sleep from 20 subjects with LGS and 20 approximately age-matched healthy controls. For s-EPIC, EEG events of interest (EOIs) were detected in all subjects using time-frequency analysis. The 11 705 EOIs were characterized based on 11 features and were collectively grouped using both k-means clustering and feature categorization. To provide clinical context, 1350 EOIs were visually reviewed and classified by three epileptologists.

Results: s-EPIC identified four clusters as candidate biomarkers of LGS, each having significantly more LGS EOIs than control EOIs. Two clusters contained EOIs resembling known LGS biomarkers such as interictal epileptiform discharges and generalized paroxysmal fast activity. The other two LGS-associated EEG clusters contained short bursts of power in beta and gamma frequency bands that were primarily unrecognized by epileptologists. This approach also uncovered significant differences in sleep spindles between LGS and control cohorts.

Significance: s-EPIC provides a quantitative approach to waveform identification that could be broadly applied to EEG from both healthy subjects and those with suspected pathology. s-EPIC can objectively identify and characterize relevant EEG waveforms without visual review or assumptions about the waveform's

Daniel W. Shrey and Beth A. Lopour contributed equally to this work.

morphology and could therefore be a powerful tool for the discovery and refinement of EEG biomarkers.

KEYWORDS

epilepsy, epileptic encephalopathy, generalized paroxysmal fast activity, interictal epileptiform discharge, machine learning, sleep spindle

1 | INTRODUCTION

The discovery of electroencephalography (EEG) biomarkers has historically relied on identifying waveforms during visual review. Initially, descriptions of such waveforms are based on repeated visual observations across multiple patients. This leads to an empirical definition, which can be used to establish potential clinical utility as a biomarker.[1] For example, sleep spindles were first described as waxing-and-waning 10–16 Hz oscillations lasting .5–2 s during sleep.[2–5] They were later associated with memory consolidation and intelligence and found to be altered in patients with epilepsy.[6,7] Similarly, high-frequency oscillations (HFOs) were first observed in rat hippocampus[8] and were later associated with epileptogenesis in humans after visual review of intracranial EEG.[9] Successful validation studies can lead to protracted use of visual review, despite the labor cost and generally low interrater reliability. Sleep spindles,[10] HFOs,[11–13] and interictal epileptiform discharges (IEDs)[14] have all followed this path.

The drawbacks of visual review often spur the creation of automated algorithms. This can enable fast, objective identification of EEG waveforms in large data sets while ensuring consistent detection criteria. Some algorithms analyze the EEG signal in the time domain and aim to match the visually derived empirical definition of the EEG waveform as closely as possible. For example, IED detectors have been designed using the energy of the signal and its slope at various points during the spike[15] and comparison to a template based on averaged IEDs.[16] HFO detectors often use an amplitude threshold,[17] or peak amplitude, as a detection criterion.[18] Time–frequency analysis is a complementary approach that can isolate activity in individual frequency bands and enable rejection of common artifacts or other false-positive detections.[19] This method has been successfully applied to the automated detection of HFOs[20–23] and the identification of interictal spikes in refractory epilepsies.[24,25]

However, the algorithm-development process has critical limitations. The detection criteria for waveforms are often based on empirical definitions derived from their visual appearance rather than the underlying physiology. In these cases, the detection accuracy is often calculated using visual markings as the "ground truth," thus limiting accuracy by the ability of raters to consistently mark the waveforms. For example, the Persyst IED detector (Persyst Development Corporation, USA) currently has the highest reported sensitivity and the highest interrater agreement with experts across commercial automated spike detectors, and has utility as a screening tool for epilepsy,[26] but it also has high variance in IED count compared to human experts and low interrater agreement with clinical markings.[26]

These challenges with biomarker discovery and validation are exemplified by Lennox–Gastaut syndrome (LGS), a severe form of childhood-onset epilepsy characterized by multiple types of refractory seizures, a variety of interictal EEG patterns, and intellectual disability. The LGS population is heterogeneous, varying in seizure type, seizure burden, interictal EEG abnormalities, etiology, and pathology. LGS is often diagnosed years after epilepsy onset, as the variability across seizure types and interictal EEG abnormalities associated with the disease obfuscate the diagnostic criteria.[27,28] Moreover, patients are often seizure-free for several years following the resolution of early life seizures, such as infantile epileptic spasms syndrome, prior to the insidious emergence of LGS.[29] This quiescent period is an attractive

Key points

- Automated, unsupervised time-frequency analysis can identify electroencephalography (EEG) waveforms specific to Lennox–Gastaut syndrome (LGS) and those specific to healthy controls.
- Our method independently identified clusters of waveforms consistent with epileptiform discharges and generalized paroxysmal fast activity.
- Sleep spindles were less frequent in LGS subjects than controls, with shorter duration, lower electrode spread, and lower peak frequency.
- Short bursts of beta and gamma power were strongly associated with LGS and have potential as a novel biomarker.
- This time–frequency approach could more broadly facilitate EEG biomarker discovery in healthy people and those with pathologies.

target for the development of biomarkers to help prevent or mitigate the impact of impending LGS.

One promising EEG biomarker for LGS is generalized paroxysmal fast activity (GPFA), a unique waveform that occurs primarily during non–rapid eye movement (NREM) sleep.[28] Despite the visual descriptions of GPFA[28,30] and its associations with intractable epilepsy[31] and increased seizure counts,[32] the interrater reliability for GPFA using visual identification remains low.[33] An automatic GPFA detector identified EEG waveforms with time–frequency properties similar to manually marked GPFA, but it required individualized patient settings and achieved only 40%–80% agreement with visual markings.[34] This suggests that EEG biomarkers for LGS demonstrate limited clinical utility, in part due to the low reliability of visual identification, similar to IEDs and HFOs.

To address this, we developed an objective, data-driven approach to EEG biomarker discovery that does not rely on visual review of the data or pre-established clinical definitions. This automated, unsupervised approach identifies EEG waveforms in the time–frequency space that are specific to patients with LGS. We show that this process cannot only identify known biomarkers of LGS, such as slow spike–wave and GPFA, but it can also lead to the discovery of novel candidate biomarkers.

2 | MATERIALS AND METHODS

2.1 | Subject information

Approval for this retrospective observational study was obtained from the institutional review boards of the Children's Hospital of Orange County (CHOC) and the University of California, Los Angeles (UCLA), with the requirement for informed consent waived. We identified 20 subjects diagnosed with LGS (7 female, median age 7.4 years, age range 1.0–18.8 years) who received care at CHOC between January 2012 and June 2020. We also identified 50 children who had EEG studies at UCLA between February 2014 and July 2018 but were determined to be neurologically normal.[35] Twenty of these controls (8 female, median age 8.2 years, age range 1.0–17.7 years) were selected to be approximately age-matched to the LGS cohort. See Supplementary Methods—Data S1 for additional details on subject selection.

2.2 | EEG acquisition and preprocessing

Control and LGS EEG data were both recorded using 19 scalp electrodes placed according to the International 10–20 system, sampled at 200 Hz. EEG data were collected overnight and manually sleep scored by a board-certified pediatric electroencephalographer to identify epochs of NREM sleep with no arousals or movement artifacts (see Supplementary Methods—Data S1). For each subject, we analyzed one, 10 min segment of clean, continuous, NREM sleep EEG with no artifacts, detected automatically[35,36] and visually. EEG data were re-referenced to the common average, bandpass filtered from .5–55 Hz with a zero-phase shift digital filter, and analyzed using the Scalp EEG Pattern Identification and Categorization (s-EPIC) method.

2.3 | Scalp EEG Pattern Identification and Categorization (s-EPIC)

2.3.1 | Identifying events of interest (EOIs)

The EEG data were pre-whitened in the time-domain using first-order backward differencing.[37–39] Each channel in every EEG was decomposed into the time–frequency domain using the Stockwell transform for frequencies from 1 to 50 Hz, with a step size of 1 Hz (Figure 1A).[32,40]

For each subject, we used the EEG time–frequency decomposition to identify periods in the Fz EEG channel where the power at any frequency exceeded a threshold p for at least 100 ms. We defined these windows of EEG as "events of interest" (EOIs). The threshold was fixed across all subjects and all frequency bands. EOIs were selected using the Fz electrode for two reasons: (1) the frontocentral location should be minimally impacted by eye movements and muscle artifact, and (2) Fz is maximally sensitive to known EEG waveforms such as sleep spindles and GPFAs.[41] A minimum time of 100 ms was chosen to fully capture the duration of an epileptic spike, which was the shortest EEG waveform we expected to see. Epileptic spikes last 20–70 ms in the time domain but can have longer duration in the time–frequency space, especially at low frequencies.[33]

2.3.2 | EOI features

Each EOI was associated with a time–frequency image (TFI), a binary image in the time–frequency domain (Figure 1A). For each EOI, 11 features were calculated based on the temporal, spectral, and spatial properties (for full mathematical details, see Supplementary Methods—Data S1):

1. Height: The difference between the highest frequency and lowest frequency in the TFI (Figure 1B).
2. Length: The duration of the TFI in seconds (Figure 1C).

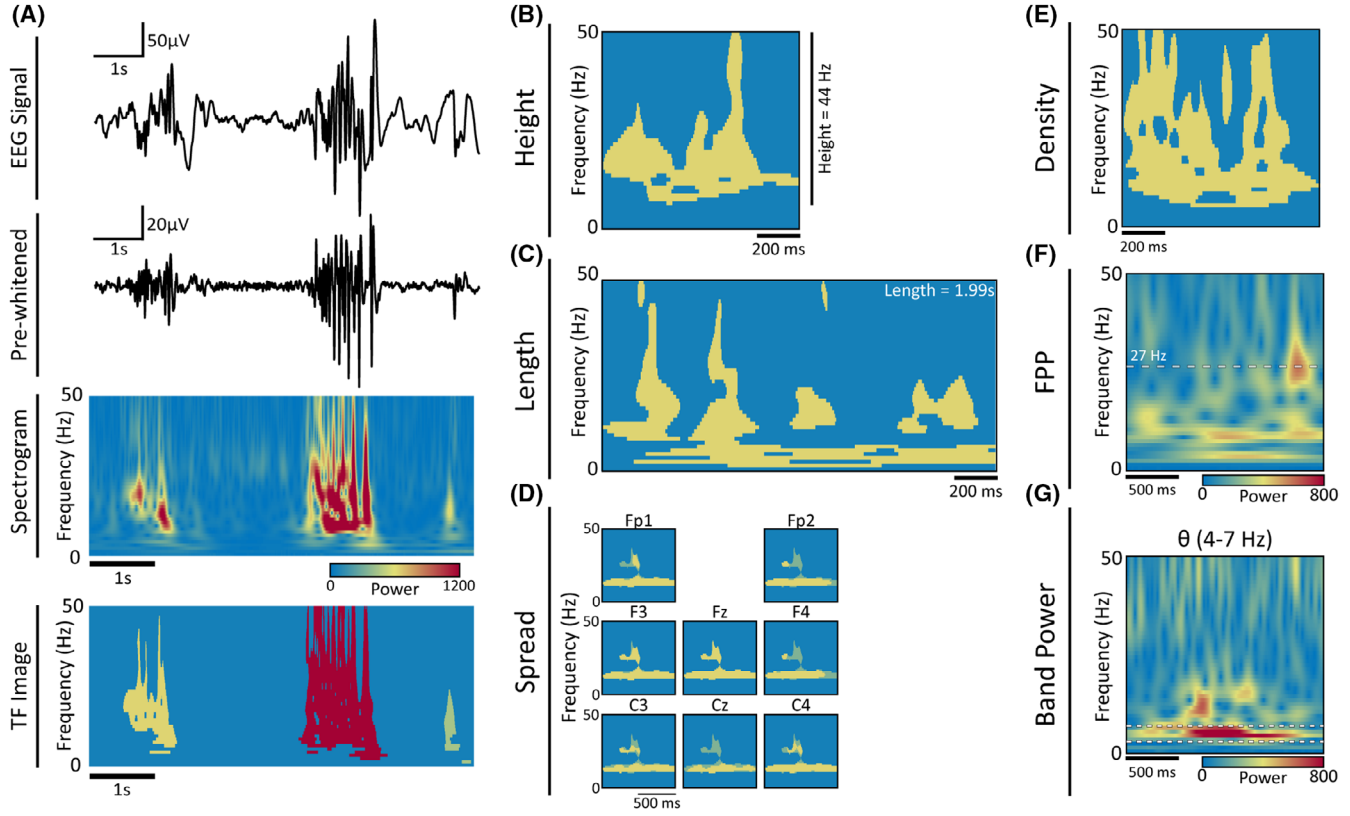

FIGURE 1 Examples of EOI identification and TFI features. (A) The broadband EEG signal is first pre-whitened, followed by time–frequency decomposition using the Stockwell transform for the entire EEG clip. EOIs were defined as consecutive time points in which the EEG power in the time–frequency decomposition exceeded a threshold for a minimum duration. Each EOI then had a corresponding region called a TFI. For example, the yellow, red, and green regions in the bottom subfigure represent three different TFIs, each one associated with an EOI in the EEG. (B) Shows an example TFI with high height, (C) high length, (D) high spread, and (E) high density. The spread is visually represented using electrodes adjacent to Fz, where yellow regions indicate overlap of the TFI in electrode Fz with the TFI for adjacent electrodes and green indicates regions where the TFI for adjacent electrodes does not overlap with the TFI for Fz. The other features used to characterize EOIs were (F) frequency of peak power, indicated by a white dashed line, and (G) mean band power in six frequency bands; a sample TFI with high theta band power (indicated by white dashed lines) is shown. EEG, electroencephalography; EOI, event of interest; TFI, time–frequency image.

3. Spread: The percentage overlap of the Fz TFI with TFIs in other electrodes, in time and frequency (Figure 1D).
4. Density: TFI area (number of data points within the TFI) divided by its length and the maximum height (50 Hz) (Figure 1E).
5. Frequency of Peak Power (FPP): The frequency at which the EOI had maximum power (Figure 1F).
6. Band Power (BP): The mean band power across all time points in the EOI within the delta ($2 \leq \delta < 4$ Hz), theta ($4 \leq \theta < 7$ Hz), alpha ($8 \leq \alpha < 12$ Hz), sigma ($12 \leq \sigma < 15$ Hz), beta ($15 \leq \beta < 30$ Hz), and gamma ($30 \leq \gamma < 50$ Hz) frequency bands (Figure 1G).

2.3.3 | Clustering and feature categorization

Prior to clustering, each feature value was normalized by dividing by the range of feature values across all EOIs.[42] All EOIs were clustered using k-means based on their 11 normalized features.[43] For visualization purposes, clusters were sorted in descending order based on the sum of the 11 z-scores, such that Cluster 1 had the highest sum of features.

One limitation of clustering is the variability of the results using a different dataset or new subjects. As a complement to clustering, we implemented a second method that categorized each EOI based on the median feature values of the height, length, spread, and density. Each of these features was classified as weak (lower than the median) or strong (higher than the median). Each EOI was then categorized using a five-character string starting with the frequency of highest BP (δ, θ, α, σ, β, γ), followed by characters indicating weak/strong height (h/H), length (l/L), spread (s/S), and density (d/D). For example, an EOI with peak power in the alpha band and height and spread above the median was categorized as "αHlSd."

2.4 | Visual classification and labelling of EOIs

To provide a clinical interpretation of each cluster, we selected a subset of 1350 EOIs for visual analysis by three board-certified pediatric epileptologists from two different institutions (CHOC and UCLA).[44] The selection procedure, detailed in Hu et al.,[44] aimed to represent the broad range of EOIs, with approximately equal selection from LGS and control subjects. Each rater classified 900 EOIs using a 15 s clip of the surrounding EEG and was blinded to subject type, subject number, and the time at which the EEG was recorded. Each EOI was independently marked by at least two different raters. Raters classified each EOI as one of the following: (1) IED, (2) trains of IEDs (IEDTs), (3) GPFA, (4) seizure, (5) sleep spindle, (6) vertex sharp, (7) muscle, (8) artifact, (9) other waveform, or (10) nothing.

2.5 | Automated sleep spindle detection

To complement the time–frequency analysis, we also applied an automated sleep spindle detector[45] to the EEG for all subjects. (See Supplementary Methods—Data S1 for details.)

3 | RESULTS

3.1 | LGS EOIs have higher values of spatial and spectral time–frequency features than control EOIs

A power threshold of $p = 250$ was selected using histograms of all power values in each cohort (Figure S1). Overall, the EEG recordings of patients with LGS had higher power than controls. This threshold was chosen to select ~5% of time–frequency values, balanced between the two cohorts, enabling detection of both pathological and physiological EOIs. Using $p = 250$ resulted in the identification of 11 705 EOIs across all 40 subjects. A total of 6743 EOIs were from 20 controls ($n_{controls} = 365.0$ [259.0–442.5]; reported as the median [Q1–Q3] for all results) and 4962 were from 20 LGS subjects ($n_{LGS} = 232.0$ [156.0–324.5]).

The height, spread, FPP, density, sigma BP, beta BP, and gamma BP of EOIs from LGS subjects were significantly greater than those from controls (Figure S2; $p < .05$, Mann–Whitney U test, Bonferroni corrected, $n = 11$). To test the impact of the threshold on EOI selection, we repeated the analysis with $p = 200$ and $p = 300$, and we found qualitatively similar results (Figure S3).

3.2 | EOIs unique to LGS subjects were isolated by clustering

To generate 12 clusters, k-means was applied to all EOIs. On average, subjects had EOIs in 9.5 clusters; each cluster contained EOIs from multiple subjects (Figure 2A). Each cluster's centroid was characterized by a different combination of the 11 features (Figure 3). The results were not dependent on the power threshold, as thresholds of $p = 200$ and $p = 300$ produced qualitatively similar cluster centroids (Figure S4). These results were also robust to the number of clusters, as using 6, 9, 15, and 18 clusters similarly resulted in LGS-dominated groups of EOIs with similar mean centroids (Figure S5).

3.3 | s-EPIC identified known and novel candidate biomarkers of LGS

Clusters 1, 2, 4, and 7 had significantly more LGS EOIs than control EOIs, and thus represent potential biomarkers for LGS (Figure 2B; $p < .05$, permutation test, Bonferroni corrected, $n = 12$). Representative control and LGS EOIs with features most similar to each cluster centroid are shown in Figure 4.

3.3.1 | EOIs in Cluster 1 were identified as GPFA in LGS subjects and sleep spindles in healthy controls

Cluster 1 consisted of 427 EOIs, 71 from 5 controls and 356 from 15 LGS subjects. These EOIs lasted ~1.2 s and had high height, electrode spread, density, and high alpha, sigma, and beta BP (Figure 3). Of the 37 visually labeled control EOIs, 83.8% of events were agreed to be sleep spindles by both raters; of the 124 labeled LGS EOIs, 58.1% were labeled as GPFA by at least one rater (Figure 4A). In total, 85.7% of all GPFAs with two-rater agreement were contained in Cluster 1.

3.3.2 | EOIs in Cluster 2 were identified as IEDTs

Cluster 2 had 490 EOIs, 18 from 9 controls and 472 from 17 LGS subjects. Events in Cluster 2 had the highest mean EOI length (4.9 s), high height, length, electrode spread, and high delta, theta, and alpha BP (Figure 3). The EOIs in Cluster 2 had consistent rater agreement, with 96.9% of the 195 marked LGS events labeled as IEDTs by at least one rater and 79.0% as IEDTs with

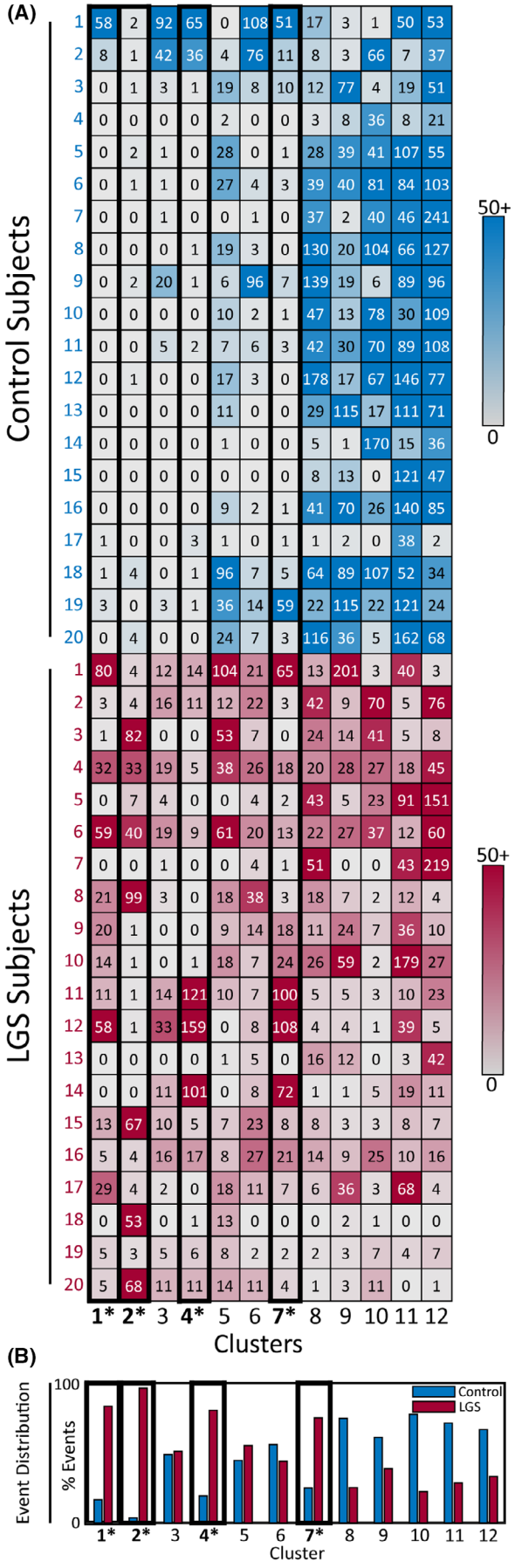

FIGURE 2 Four clusters had EOIs primarily originating from LGS subjects. (A) Heatmap of the number of EOIs in each cluster for all controls (*top*) and LGS subjects (*bottom*). (B) The percentage of control and LGS EOIs in each cluster. Clusters 1, 2, 4, and 7 had a significantly greater number of EOIs from LGS subjects compared to controls and are formatted in bold text with asterisks. Significance level is $p < .05$, with *p*-values corrected using the Bonferroni method. EOI, event of interest; LGS, Lennox–Gastaut syndrome.

two-rater agreement. In contrast, there were no consistent labels across the 18 reviewed control events (Figure 4B).

3.3.3 | EOIs in Clusters 4 and 7 have high beta and gamma power and are novel candidate biomarkers of LGS

In Cluster 4, 111 EOIs were found across 9 controls and 461 across 13 LGS subjects; in Cluster 7, 156 EOIs were found across 13 controls and 469 EOIs across 17 LGS subjects. These clusters are distinguished by their low length, with high gamma BP and FPP of 36.1 Hz in Cluster 4, and high beta BP and FPP of 24.9 Hz in Cluster 7 (Figure 3). Events in Cluster 7 have a significantly lower height, spread, and density compared to Cluster 4 ($p < .05$, Mann–Whitney *U* test, Bonferroni corrected, $n = 11$), although these features have low values relative to those in Clusters 1 and 2. Rater labels for Clusters 4 and 7 were inconclusive for both LGS and control EOIs. In Cluster 4, 88.2% of LGS EOIs and 94.1% of control EOIs were labeled as "nothing" or were given mismatched labels by the two raters; similarly, this was true for 96.4% of LGS EOIs and 87.5% control EOIs in Cluster 7 (Figure 4C,D). The frequent occurrence of these EOIs in LGS subjects compared to controls, combined with a lack of recognition by clinicians, suggests that they may be novel candidate biomarkers of LGS.

3.4 | Alterations to sleep spindles in LGS subjects compared to controls

The substantially reduced number of sleep spindles in LGS subjects warranted further investigation, as only five EOIs had two-rater agreement on sleep spindles in LGS, compared to 113 in controls. Within those EOIs, LGS spindles had a shorter length, lower height, lower electrode spread, and a lower FPP compared to spindles in healthy controls ($p < .05$, Mann–Whitney *U* test, Bonferroni corrected, $n = 4$). To independently validate this finding, we ran an automatic sleep spindle detector[45] on the EEG

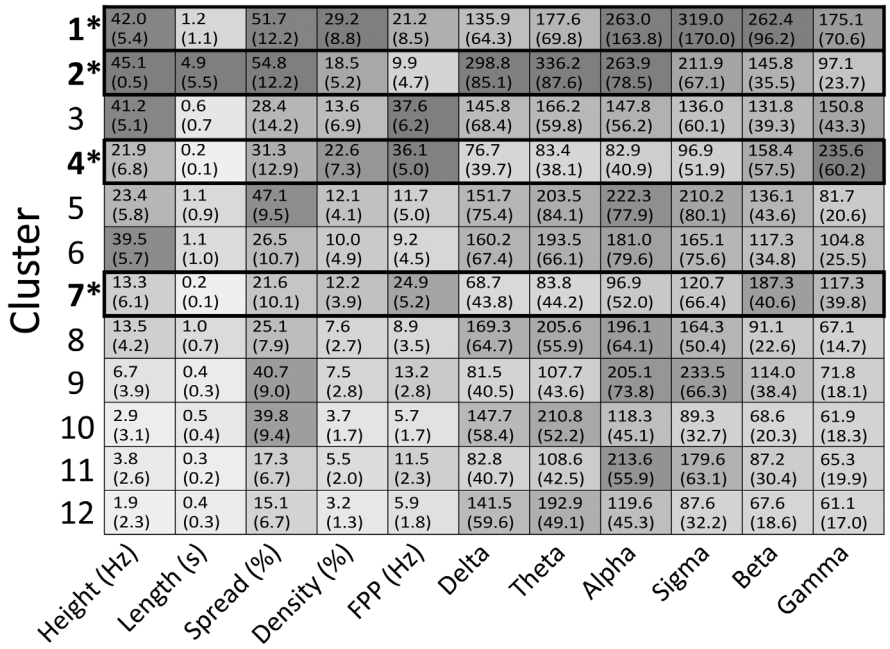

FIGURE 3 Mean feature values for each cluster. The opacity of the cells indicates the feature value relative to the maximum in each column, with a value of zero appearing white. Clusters with a significantly greater number of LGS EOIs than control EOIs are formatted in bold text and indicated by asterisks. Numbers are reported as mean (standard deviation [SD]). Significance level is p<0.05, with p-values corrected using the Bonferroni method. EOI, event of interest; LGS, Lennox–Gastaut syndrome.

recordings from all subjects. Sleep spindles were detected more frequently in healthy controls ($n = 96.5$ [64.5–123.0]) than LGS subjects ($n = 25.5$ [18.5–44.0]) (Figure 5). Moreover, the mean LGS spindle was significantly lower in length, global spread, and peak frequency compared to that of healthy controls ($p < .05$, Mann–Whitney U test, Bonferroni corrected, $n = 4$) (Figure 5). These results mirror the s-EPIC results, suggesting sleep spindle occurrence and characteristics as candidate biomarkers of LGS.

3.5 | Analysis of EOIs using feature categorization

To complement the clustering analysis, we also placed each EOI into one of 96 categories, as described in the Methods section on Feature Categorization. The median EOI feature values were a height of 7 Hz, length of 360 ms, spread of 26.3%, and density of 6.4%. Eighty-three of the 96 categories had at least one EOI, with a median of 43.5 [6–167.0] EOIs per category.

3.5.1 | EOIs with high beta and gamma band power are potential biomarkers of LGS

Twenty-two categories had significantly more EOIs from LGS subjects than controls (Table 1A; $p < .05$, permutation test, Bonferroni corrected, $n = 96$). Seven significant categories had peak BP in delta, theta, alpha, or sigma, with high height and density; most also had high spread and low length. In contrast, the nine categories with peak beta BP and six with peak gamma BP included combinations of high and low feature values (Table 1A). The wide variety of feature combinations suggests a broader relationship between LGS and beta/gamma band EOIs. For example, in the beta band, nine of 16 categories were significantly associated with LGS, and the remaining seven categories each contained fewer than two EOIs. Similarly, only 5 of 16 gamma band categories had more than one EOI, with all 5 containing significantly more LGS than control EOIs (Table 1A). This suggests that simply identifying EOIs with beta or gamma peak BP, without any further clustering or categorization, can distinguish LGS subjects from controls. Doing this, we find that across all EOIs, those with peak beta BP or peak gamma BP originated from LGS subjects 84.1% and 78.2% of the time, respectively.

3.5.2 | Feature categorization identified EOIs strongly associated with the control group

Twenty-three feature categories had significantly more control EOIs than LGS EOIs (Table 1B). Most of these categories were characterized by peak alpha or sigma BP; 52.0% of all EOIs in the 23 categories had peak BP in the alpha or sigma frequency. Events in these categories generally had low feature values, with 82.2% of them having low height, 60.8% having low length, 64.2% having low spread, and 78.0% having low density. Only three control-associated categories had peak beta or gamma BP, and each one contained a single EOI (Table 1B).

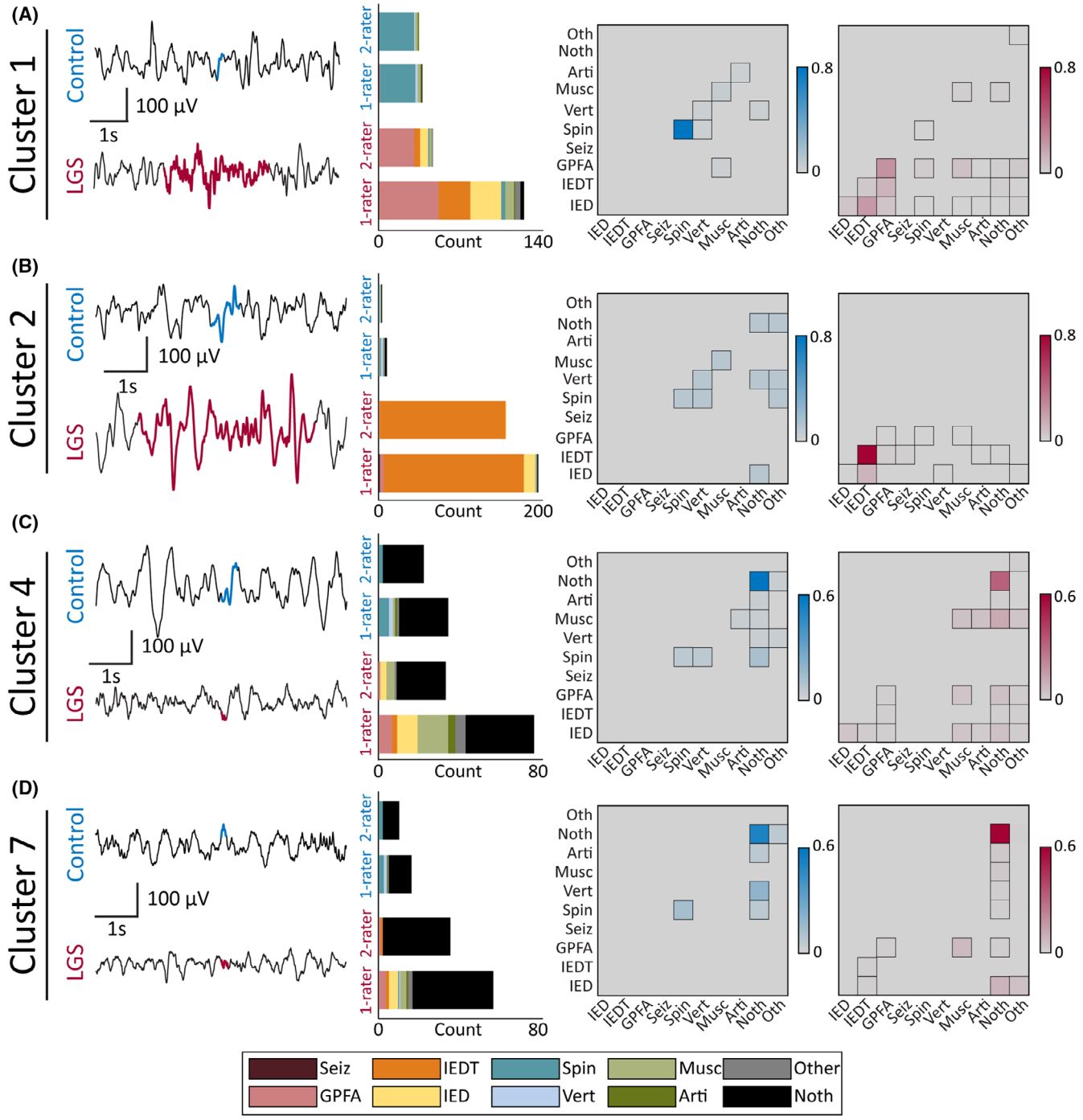

FIGURE 4 Breakdown of reviewer classifications in clusters with significantly more LGS EOIs than control EOIs: (A) Cluster 1, (B) Cluster 2, (C) Cluster 4, and (D) Cluster 7. For each cluster, the left column shows a representative EEG waveform for controls (blue) and LGS subjects (red), selected using the minimum Euclidean distance from the cluster centroid. The bar graph in the middle shows the occurrence of each label for control and LGS EOIs marked by a single rater or two raters, with the latter indicating rater agreement on the EOI label. The right column shows confusion matrices for rater labels applied to control (blue) and LGS (red) EOIs. EEG, electroencephalography; EOI, event of interest; LGS, Lennox–Gastaut syndrome.

4 | DISCUSSION

Here, we presented s-EPIC, a novel technique to discover EEG biomarkers using time–frequency analysis. When applied to patients with LGS, s-EPIC identified EOIs consistent with known clinical biomarkers, specifically IEDTs and GPFA. It also uncovered significant differences in sleep spindles between LGS subjects and healthy controls and suggested beta/gamma band EOIs as novel candidate biomarkers of LGS, as they occurred

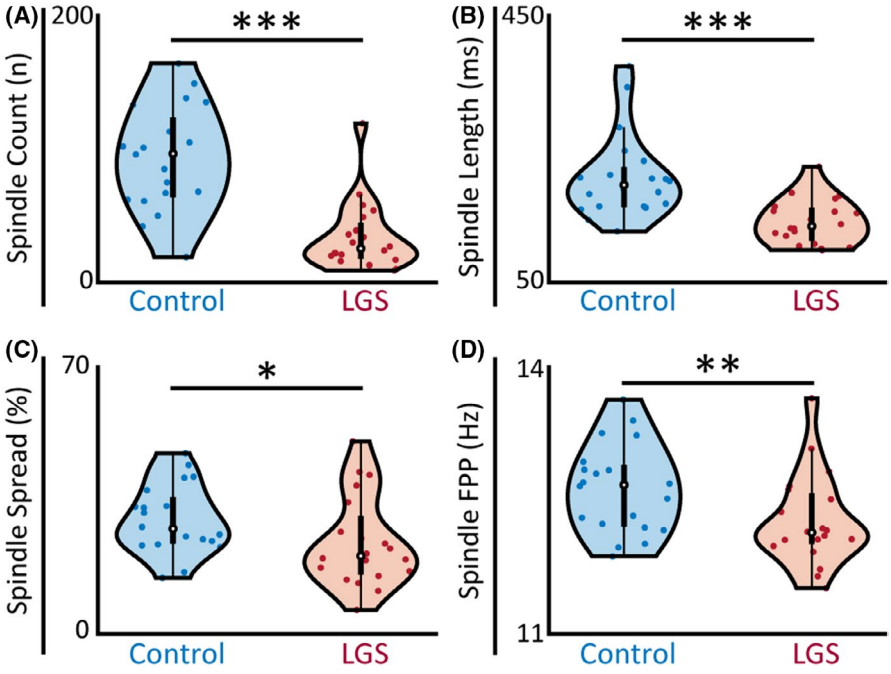

FIGURE 5 Mean features of automatically detected sleep spindles across all control and LGS subjects. Sleep spindles in healthy controls had a significantly greater (A) spindle count, (B) spindle length, (C) spindle spread, and (D) peak frequency compared to spindles in LGS subjects. Significance levels are $*p < .05$, $**p < .01$, and $***p < .001$, with p-values modified using the Bonferroni method. LGS, Lennox–Gastaut syndrome.

predominantly in LGS subjects but were unrecognized by epileptologists. These results, encompassing both known and previously unknown EEG waveforms, speak to the potential power of this approach.

4.1 | Comparison to prior time–frequency image analyses

Prior implementations of time–frequency analysis have been used primarily to create automated detectors for specific EEG waveforms, such as GPFA,[34] IED,[24] or HFOs.[20,22] One study used time–frequency image features, such as area, entropy, width, and height,[20] to quantify HFOs; another HFO study used features of the time–frequency spectrogram, such as flux, energy concentration, and skewness.[22] Time–frequency analysis has also been applied to EEG from children with focal epilepsies to detect IEDs using power spectrum changes.[24] Our work improves on these prior studies. First, we broadly detect all anomalies in the time–frequency domain without requiring empirical definitions of waveforms, preset parameters, or annotated EEG data. Second, our method can equally analyze EEG of both epilepsy patients and healthy controls, rather than being tailored to epilepsy-related event detection. This enabled us to identify clusters of EOIs that were pathological (occurring mostly in LGS subjects) and physiological (occurring mostly in controls). Third, our method is the first to enable the detection of EOIs of variable length, rather than relying on the assumption of a fixed event duration.

4.2 | Time–frequency characteristics of GPFA

Visually reviewed EOIs labeled as GPFA by at least one clinician were found in 40% of LGS subjects[33]; this is lower than a prior study, which reported generalized paroxysmal fast rhythms in 66% of patients with genetic generalized epilepsy using a 24 h EEG recording.[46] The relative rarity of GFPA in our data set was expected, given that we analyzed only 10 min of sleep EEG, compared to the multiple hours of EEG often needed to see the first generalized polyspike activity.[47]

There is currently no consensus on GPFA characteristics. One study of scalp EEG reported GPFA to have a mean amplitude of 293 uV, length of 1.6 s, and frequency of 11.1 Hz[48]; another study reported a mean amplitude of 88.3 uV, length of 1–4 s, and frequencies of 11–20 Hz.[41] An automated GPFA detector found bursts as low as 3 Hz in some subjects and up to 16–18 Hz in others.[34] GPFA in Cluster 1 had a mean length of 1.9 s, FPP of 19.6 Hz, alpha BP of 467.8, and sigma BP of 500.0. This characterization may be a useful benchmark for future studies, including differentiating GPFA from generalized polyspike trains, which have overlapping features.[31,47]

4.3 | Aberration of sleep spindles in LGS

The lower spindle rate in LGS subjects compared to healthy controls is consistent with lower global rates of sleep spindles during N2 sleep reported in other epilepsies.[49,50]

TABLE 1 Feature categories with significantly more LGS or control EOIs.

(A) Significant LGS categories			(B) Significant control categories		
No. EOIs	No. LGS EOI (%)	Category	No. EOIs	NO. LGS EOI (%)	Category
393	291 (74.1%)	δHLSD	180	48 (26.7%)	δhLSd
57	38 (66.7%)	θHlSD	839	248 (29.6%)	θhlsd
711	463 (65.1%)	θHLSD	371	87 (23.4%)	θhlSd
20	17 (85.0%)	αHlSd	421	71 (16.9%)	θhLSd
109	80 (73.4%)	αHlSD	256	82 (32.0%)	θHLsD
147	83 (56.5%)	σHlsD	144	33 (22.9%)	θHLSd
214	168 (78.5%)	σHlSD	862	217 (25.2%)	αhlsd
77	61 (79.2%)	βhlsD	143	34 (23.8%)	αhlsD
7	7 (100%)	βhlSd	255	79 (31.0%)	αhlSd
39	35 (89.7%)	βhlSD	107	10 (9.4%)	αhLsD
245	194 (79.2%)	βHlsD	74	13 (17.6%)	αhLSd
1	1 (100%)	βHlSd	82	7 (8.5%)	αhLSD
221	194 (87.8%)	βHlSD	62	14 (22.6%)	αHLsd
1	1 (100%)	βHLsd	254	53 (20.9%)	αHLsD
17	17 (100%)	βHLsD	327	94 (28.8%)	σhlsd
68	65 (95.6%)	βHLSD	27	1 (3.7%)	σhLsD
6	6 (100%)	γhlsD	35	2 (5.7%)	σhLSd
1	1 (100%)	γhlSD	65	5 (7.7%)	σhLSD
259	180 (69.5%)	γHlsD	78	16 (20.5%)	σHLsD
281	231 (82.2%)	γHlSD	22	3 (13.6%)	σHLSd
13	12 (92.3%)	γHLsD	1	0 (0%)	βhLsD
56	53 (94.6%)	γHLSD	1	0 (0%)	γHlsd
			1	0 (0%)	γHlSd

Note: The five-character strings indicate the frequency of highest band power (δ, θ, α, σ, β, γ), followed by characters indicating low/high height (h/H), length (l/L), spread (s/S), and density (d/D), based on a comparison to the median value across all subjects.

Abbreviations: EOI, event of interest; LGS, Lennox–Gastaut syndrome.

However, LGS subjects have significantly less stage 2 sleep compared to age-matched controls,[50] which could reduce the likelihood of spindles occurring within the NREM EEG clips in our study. The presence of sleep spindles in LGS can also be masked by concurrent epileptiform activity, which confounds visual classification.

4.4 | Candidate high frequency EEG biomarkers for LGS

Our automated time–frequency analysis identified beta and gamma band EOIs as candidate EEG biomarkers of LGS. These events, found in Clusters 4 and 7, had generally low feature values, similar to controls, yet 78% of the events in these categories originated from LGS subjects. The EOIs in both clusters were unrecognizable to epileptologists, as 64.6% of all EOIs in Cluster 4 and 83.3% of EOIs in Cluster 7 were labeled as "nothing" by at least one rater.

It is noteworthy that the remaining EOIs in Clusters 4 and 7 were not interpreted to be muscle artifacts. Muscle artifact has outlying high amplitude, long duration, and high electrode spread in the temporal and frontopolar channels, in contrast to the features described in these two clusters.[51,52] However, the EOIs in Clusters 4 and 7 do share characteristics with paroxysmal EEG waveforms suggested to be a biomarker of epileptogenesis in children, which have peak frequencies in the beta and gamma bands and durations >200 ms.[53]

4.5 | Limitations and future directions

There are several limitations to our study. EOIs were detected solely in the Fz channel, although our time–frequency method could be adapted to multi-channel detection, with additional criteria to avoid double-counting EOIs that span multiple channels. This expanded approach

would be supported by the use of high-resolution EEG, which would also improve the estimates of EOI spread. In addition, the 10 min of EEG analyzed for each subject is a relatively short duration for characterizing paroxysmal activity, as prior studies used recordings of up to 24 h.[46,47] The amount of EEG data available for analysis can be increased through automated removal of artifact, such as muscle activity.[54] This would enable analysis of EEG recorded during wakefulness and automated selection of EEG epochs to reduce the bias associated with manual selection. The use of high sampling rate EEG (≥1000 Hz) would also facilitate this approach. The computational efficiency of this work can also be improved through implementation of the fast S-transform, enabling analysis of longer EEG recordings with a higher sampling rate and a larger subject population.[55] The EEG visual analysis procedure was also different from standard clinical review, as raters viewed 15 s of isolated EEG surrounding the EOI, rather than scrolling through continuous recordings.

Future work may focus on improving the clustering techniques. The k-means provided broad separation between features but was ineffective for differentiating EEG waveforms with overlapping properties, such as sleep spindles and GPFA. Other clustering methods, such as hierarchical clustering or decision trees, may be able to address this limitation. Further validation of our findings with an independent data set comprising healthy controls, children with LGS, and patients with other types of pediatric epilepsy, is paramount.

5 | CONCLUSION

s-EPIC is a robust computational approach to biomarker discovery based on the time–frequency features of EEG. This method can be applied to normal or abnormal EEG recordings to help establish quantitative definitions of EEG biomarkers without visual review. Ultimately, this can reduce the reliance on empirical definitions of EEG waveforms, increase the accuracy of known EEG biomarkers, and facilitate the discovery of novel biomarkers of health and disease.

AUTHOR CONTRIBUTIONS

Derek K. Hu has first authorship of the study and oversaw the conceptualization, methodology, software, verification, formal analysis, and the draft manuscript preparations for the study. Marco A. Pinto-Orellana helped design the methodology and review and edit the manuscript. Mandeep Rana and Linda Do contributed to the data collection. David J. Adams, Shaun A. Hussain, and Daniel W. Shrey provided clinical expertise, helped review and edit the manuscript, and served as blinded raters in the study. Beth A. Lopour is the corresponding author of the study and helped conceptualize, administer, and supervise the project, and reviewed and edited the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank the clinical epileptologists at Children's Hospital of Orange County for their contributions to this study. This work was supported by the Lennox–Gastaut Syndrome Foundation and the John C. Hench Foundation.

CONFLICT OF INTEREST STATEMENT

None of the authors has any conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

Anonymized data that support the findings of this study are available from the corresponding author upon request.

ETHICS STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Derek K. Hu https://orcid.org/0000-0003-4720-2622
Marco A. Pinto-Orellana https://orcid.org/0000-0001-6495-1305
Shaun A. Hussain https://orcid.org/0000-0001-6947-8852
Daniel W. Shrey https://orcid.org/0000-0002-3163-4773
Beth A. Lopour https://orcid.org/0000-0003-4233-4802

REFERENCES

1. Lai N, Li Z, Xu C, Wang Y, Chen Z. Diverse nature of interictal oscillations: EEG-based biomarkers in epilepsy. Neurobiol Dis. 2022;2023(177):105999.
2. Andrillon T, Nir Y, Staba RJ, Ferrarelli F, Cirelli C, Tononi G, et al. Sleep spindles in humans: insights from intracranial EEG and unit recordings. J Neurosci. 2011;31(49):17821–34.
3. Loomis AL, Harvey EN, Hobart G. Potential rhythms of the cerebral cortex during sleep. Science. 1935;81(2111):597–8.
4. Gibbs FA, Gibbs EL. Atlas of electroencephalography. Boston City Hospital: F. A. Gibbs; 1941.
5. Wallant DCT, Maquet P, Phillips C. Sleep spindles as an electrographic element: description and automatic detection methods. Neural Plast. 2016;2016:1–19.
6. Fernandez LMJ, Lüthi A. Sleep spindles: mechanisms and functions. Physiol Rev. 2020;100(2):805–68.
7. Latka M, Kozik A, Jernajczyk J, West BJ, Jernajczyk W. Wavelet mapping of sleep spindles in young patients with epilepsy. J Physiol Pharmacol an off J Polish Physiol Soc. 2005;56 (Suppl 4):15–20.

8. Buzsáki G, Horváth Z, Urioste R, Hetke J, Wise K. High-frequency network oscillation in the hippocampus. Science. 1992;256(5059):1025–7.
9. Bragin A, Engel J, Wilson CL, Fried I, Buzsáki G. High-frequency oscillations in human brain. Hippocampus. 1999;9(2):137–42.
10. Wendt SL, Welinder P, Sorensen HBD, Peppard PE, Jennum P, Perona P, et al. Inter-expert and intra-expert reliability in sleep spindle scoring. Clin Neurophysiol off J Int Fed Clin Neurophysiol. 2015;126(8):1548–56.
11. Nariai H, Wu JY, Bernardo D, Fallah A, Sankar R, Hussain SA. Interrater reliability in visual identification of interictal high-frequency oscillations on electrocorticography and scalp EEG. Epilepsia Open. 2018;3(2):127–32.
12. Spring AM, Pittman DJ, Aghakhani Y, Jirsch J, Pillay N, Bello-Espinosa LE, et al. Interrater reliability of visually evaluated high frequency oscillations. Clin Neurophysiol off J Int Fed Clin Neurophysiol. 2017;128(3):433–41.
13. Spring AM, Pittman DJ, Aghakhani Y, Jirsch J, Pillay N, Bello-Espinosa LE, et al. Generalizability of high frequency oscillation evaluations in the ripple band. Front Neurol. 2018;9(June):1–11.
14. Jing J, Herlopian A, Karakis I, Ng M, Halford JJ, Lam A, et al. Interrater reliability of experts in identifying Interictal Epileptiform discharges in electroencephalograms. JAMA Neurol. 2020;77(1):49–57.
15. Gaspard N, Alkawadri R, Farooque P, Goncharova II, Zaveri HP. Automatic detection of prominent interictal spikes in intracranial EEG: validation of an algorithm and relationsip to the seizure onset zone. Clin Neurophysiol off J Int Fed Clin Neurophysiol. 2014;125(6):1095–103.
16. Karoly PJ, Freestone DR, Boston R, Grayden DB, Himes D, Leyde K, et al. Interictal spikes and epileptic seizures: their relationship and underlying rhythmicity. Brain. 2016;139(4):1066–78.
17. Charupanit K, Sen-Gupta I, Lin JJ, Lopour BA. Detection of anomalous high-frequency events in human intracranial EEG. Epilepsia Open. 2020;5(2):263–73.
18. Charupanit K, Lopour BA. A simple statistical method for the automatic detection of ripples in human intracranial EEG. Brain Topogr. 2017;30(6):724–38.
19. Jiang X, Bin BG, Tian Z. Removal of artifacts from EEG signals: a review. Sensors (Switzerland). 2019;19(5):1–18.
20. Migliorelli C, Bachiller A, Alonso JF, Romero S, Aparicio J, Jacobs-le van J, et al. SGM: a novel time-frequency algorithm based on unsupervised learning improves high-frequency oscillation detection in epilepsy. J Neural Eng. 2020;17(2):026032.
21. Burnos S, Hilfiker P, Sürücü O, Scholkmann F, Krayenbühl N, Grunwald T, et al. Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis. PLoS One. 2014;9(4):e94381.
22. Krikid F, Karfoul A, Chaibi S, Kachenoura A, Nica A, Kachouri A, et al. Classification of high frequency oscillations in intracranial EEG signals based on coupled time-frequency and image-related features. Biomed Signal Process Control. 2022;73:103418.
23. Donos C, Mîndruţă I, Barborica A. Unsupervised detection of high-frequency oscillations using time-frequency maps and computer vision. Front Neurosci. 2020;14:183.
24. Jabran Y, Mahmoudzadeh M, Martinez N, Heberlé C, Wallois F, Bourel-Ponchel E. Temporal and spatial dynamics of different Interictal epileptic discharges: a time-frequency EEG approach in pediatric focal refractory epilepsy. Front Neurol. 2020;11:1–13.
25. Jacobs J, Kobayashi K, Gotman J. High-frequency changes during interictal spikes detected by time-frequency analysis. Clin Neurophysiol. 2011;122(1):32–42.
26. Reus EEM, Cox FME, van Dijk JG, Visser GH. Automated spike detection: which software package? Seizure. 2022;95:33–7.
27. Kim HJ, Kim HD, Lee JS, Heo K, Kim DS, Kang HC. Long-term prognosis of patients with Lennox–Gastaut syndrome in recent decades. Epilepsy Res. 2015;110:10–9.
28. Mastrangelo M. Lennox—Gastaut Syndrome: A State of the Art Review. 2017.
29. Berg AT, Levy SR, Testa FM. Evolution and course of early life developmental encephalopathic epilepsies: focus on Lennox–Gastaut syndrome. Epilepsia. 2018;59(11):2096–105.
30. Kane N, Acharya J, Benickzy S, Caboclo L, Finnigan S, Kaplan PW, et al. A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017. Clin Neurophysiol Pract. 2017;2:170–85.
31. Cerulli Irelli E, Barone FA, Mari L, Morano A, Orlando B, Salamone EM, et al. Generalized fast discharges along the genetic generalized epilepsy Spectrum: clinical and prognostic significance. Front Neurol. 2022;13:1–10.
32. Dalic LJ, Warren AEL, Spiegel C, Thevathasan W, Roten A, Bulluss KJ, et al. Paroxysmal fast activity is a biomarker of treatment response in DBS for Lennox–Gastaut syndrome. Epilepsia. 2022;63:1–14.
33. Hu DK, Rana M, Adams DJ, Do L, Shrey DW, Hussain SA, et al. Interrater Reliability of Interictal EEG Waveforms in Lennox–Gastaut Syndrome. Epilepsia open. 2024;9:176–86.
34. Omidvarnia A, Warren AEL, Dalic LJ, Pedersen M, Jackson G. Automatic detection of generalized paroxysmal fast activity in interictal EEG using time-frequency analysis. Comput Biol Med. 2021;133:104287.
35. Smith RJ, Hu DK, Shrey DW, Rajaraman R, Hussain SA, Lopour BA. Computational characteristics of interictal EEG as objective markers of epileptic spasms. Epilepsy Res. 2021;176:106704.
36. Hu DK, Goetz PW, To PD, Garner C, Magers AL, Skora C, et al. Evolution of cortical functional networks in healthy infants. Front Netw Physiol. 2022;2:1–12.
37. Usui S, Amidror I. Digital differentiation filters for biological signal processing. Int Ser Biomech. 1983;4(B10):1207–14.
38. Roehri N, Lina JM, Mosher JC, Bartolomei F, Benar CG. Time-frequency strategies for increasing high-frequency oscillation detectability in intracerebral EEG. IEEE Trans Biomed Eng. 2016;63(12):2595–606.
39. Gardner AB, Worrell GA, Marsh E, Dlugos D, Litt B. Human and automated detection of high-frequency oscillations. Clin Neurophysiol. 2007;118(118):1134–43.
40. Stockwell RG. Localization of the complex spectrum: the s transform. IEEE Trans Signal Process. 1996;44(4):993–1001.
41. Bansal L, Vargas Collado L, Pawar K, Nagesh D, Ilyas M, Hall A, et al. Electroclinical features of generalized paroxysmal fast activity in typical absence seizures. J Clin Neurophysiol. 2019;36(1):36–44.
42. Milligan GW, Cooper MC. A study of standardization of variables in cluster analysis. J Classif. 1988;5(2):181–204.

43. Lloyd SP. Least squares quantization in PCM. IEEE Trans Inf Theory. 1982;28(2):129–37.
44. Hu DK, Rana M, Adams DJ, do L, Shrey DW, Hussain SA, et al. Interrater reliability of interictal EEG waveforms in Lennox–Gastaut syndrome. Epilepsia Open. 2023;9:1–11.
45. Mensen A, Poryazova R, Huber R, Bassetti CL. Individual spindle detection and analysis in high-density recordings across the night and in thalamic stroke. Sci Rep. 2018;8(1):1–11.
46. Seneviratne U, Cook MJ, D'Souza WJ. Electroencephalography in the diagnosis of genetic generalized epilepsy syndromes. Front Neurol. 2017;8:499.
47. Conrad EC, Chugh N, Ganguly TM, Gugger JJ, Tizazu EF, Shinohara RT, et al. Using generalized Polyspike train to predict drug-resistant idiopathic generalized epilepsy. J Clin Neurophysiol. 2022;39(6):459–65.
48. Mohammadi M, Okanishi T, Okanari K, Baba S, Sumiyoshi H, Sakuma S, et al. Asymmetrical generalized paroxysmal fast activities in children with intractable localization-related epilepsy. Brain and Development. 2015;37(1):59–65.
49. Schiller K, Avigdor T, Abdallah C, Sziklas V, Crane J, Stefani A, et al. Focal epilepsy disrupts spindle structure and function. Sci Rep. 2022;12(1):1–11.
50. Eisensehr I, Parrino L, Noachtar S, Smerieri A, Terzano MG. Sleep in Lennox–Gastaut syndrome: the role oft the cyclic alternating pattern (CAP) in the gate control of clinical seizures and generalized polyspikes. Epilepsy Res. 2001;46(3):241–50.
51. Brunner DP, Vasko RC, Detka CS, Monahan JP, Reynolds CF, Kupfer DJ. Muscle artifacts in the sleep EEG: automated detection and effect on all-night EEC power spectra. J Sleep Res. 1996;5(3):155–64.
52. Kobayashi K, Inoue T, Watanabe Y, Oka M, Endoh F, Yoshinaga H, et al. Spectral analysis of EEG gamma rhythms associated with tonic seizures in Lennox–Gastaut syndrome. Epilepsy Res. 2009;86(1):15–22.
53. Wu JY, Koh S, Sankar R, Mathern GW. Paroxysmal fast activity: an interictal scalp EEG marker of epileptogenesis in children. Epilepsy Res. 2008;82(1):99–106.
54. Chen X, Liu Q, Tao W, Li L, Lee S, Liu A, et al. ReMAE: user-friendly toolbox for removing muscle artifacts from EEG. IEEE Trans Instrum Meas. 2020;69(5):2105–19.
55. Brown RA, Frayne R. A fast discrete S-transform for biomedical signal processing. Proc 30th Annu Int Conf IEEE Eng Med Biol Soc EMBS'08 - "Personalized Healthc through Technol 2008 2586–2589.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hu DK, Pinto-Orellana MA, Rana M, Do L, Adams DJ, Hussain SA, et al. Discovering EEG biomarkers of Lennox–Gastaut syndrome through unsupervised time–frequency analysis. Epilepsia. 2025;66:541–553. https://doi.org/10.1111/epi.18211