Interictal stereo-electroencephalography features below 45 Hz are sufficient for correct localization of the epileptogenic zone and postsurgical outcome prediction

Petr Klimes, Petr Nejedly, Valentina Hrtonova, Jan Cimbalnik, Vojtech Travnicek, Martin Pail, Laure Peter-derex, Jeffery Hall, Raluca Pana, Josef Halamek, et al.

Received: 19 March 2024 | Revised: 18 July 2024 | Accepted: 25 July 2024

DOI: 10.1111/epi.18081

Epilepsia®

RESEARCH ARTICLE

Interictal stereo-electroencephalography features below 45 Hz are sufficient for correct localization of the epileptogenic zone and postsurgical outcome prediction

Petr Klimes[1] | **Petr Nejedly**[1,2] | **Valentina Hrtonova**[1] | **Jan Cimbalnik**[2,3] | **Vojtech Travnicek**[1,3] | **Martin Pail**[1,2] | **Laure Peter-Derex**[4,5] | **Jeffery Hall**[6] | **Raluca Pana**[6] | **Josef Halamek**[1] | **Pavel Jurak**[1] | **Milan Brazdil**[2,7] | **Birgit Frauscher**[8,9,10]

[1]Institute of Scientific Instruments of the CAS, Brno, Czech Republic

[2]Department of Neurology, Faculty of Medicine, Brno Epilepsy Center, St. Anne's University Hospital, Member of ERN-EpiCARE, Masaryk University, Brno, Czech Republic

[3]International Clinical Research Center, St. Anne's University Hospital, Brno, Czech Republic

[4]Center for Sleep Medicine and Respiratory Diseases, Lyon University Hospital, Lyon 1 University, Lyon, France

[5]Lyon Neuroscience Research Center, INSERM U1028/CNRS UMR5292, Lyon, France

[6]Montreal Neurological Hospital, McGill University, Montreal, Quebec, Canada

[7]Behavioral and Social Neuroscience Research Group, CEITEC Central European Institute of Technology, Masaryk University, Brno, Czech Republic

[8]Analytical Neurophysiology Lab, Montreal Neurological Institute and Hospital, McGill University, Quebec, Canada

[9]Department of Neurology, Duke University Medical Center, Durham, North Carolina, USA

[10]Department of Biomedical Engineering, Duke Pratt School of Engineering, Durham, North Carolina, USA

Correspondence
Petr Klimes, Institute of Scientific Instruments of the CAS, Brno, Czech Republic.
Email: petr.klimes@isibrno.cz

Funding information
Grantová Agentura České Republiky, Grant/Award Number: 22-28784S; The Czech Academy of Sciences, Grant/Award Number: RVO:68081731; Ministerstvo Zdravotnictví Ceské Republiky, Grant/Award Number: NU22-08-00278; Ministerstvo Školství, Mládeže a Tělovýchovy, Grant/Award Number: LM2023053 and LX22NPO5107

Abstract

Objective: Evidence suggests that the most promising results in interictal localization of the epileptogenic zone (EZ) are achieved by a combination of multiple stereo-electroencephalography (SEEG) biomarkers in machine learning models. These biomarkers usually include SEEG features calculated in standard frequency bands, but also high-frequency (HF) bands. Unfortunately, HF features require extra effort to record, store, and process. Here we investigate the added value of these HF features for EZ localization and postsurgical outcome prediction.

Methods: In 50 patients we analyzed 30 min of SEEG recorded during non–rapid eye movement sleep and tested a logistic regression model with three different sets of features. The first model used broadband features (1–500 Hz); the second model used low-frequency features up to 45 Hz; and the third model used HF features above 65 Hz. The EZ localization by each model was evaluated by various metrics including the area under the precision-recall curve (AUPRC) and the positive predictive value (PPV). The differences between the models were tested by the

Wilcoxon signed-rank tests and Cliff's Delta effect size. The differences in outcome predictions based on PPV values were further tested by the McNemar test.

Results: The AUPRC score of the random chance classifier was .098. The models (broad-band, low-frequency, high-frequency) achieved median AUPRCs of .608, .582, and .522, respectively, and correctly predicted outcomes in 38, 38, and 33 patients. There were no statistically significant differences in AUPRC or any other metric between the three models. Adding HF features to the model did not have any additional contribution.

Significance: Low-frequency features are sufficient for correct localization of the EZ and outcome prediction with no additional value when considering HF features. This finding allows significant simplification of the feature calculation process and opens the possibility of using these models in SEEG recordings with lower sampling rates, as commonly performed in clinical routines.

KEYWORDS

EEG, epilepsy, high-frequency oscillations, interictal epileptoform discharges, machine learning

1 | INTRODUCTION

Epilepsy is one of the most common neurological disorders, affecting up to 70 million people worldwide.[1] Unfortunately, ≈40% of patients do not respond to antiseizure medications and the most effective treatment is epilepsy surgery.[2,3] The purpose of epilepsy surgery is to remove the epileptogenic zone (EZ), defined as "the area of cortex that is indispensable for the generation of epileptic seizures."[4] Because it is not possible to know in advance if the removal of the tissue will lead to seizure freedom, the most commonly used approximation of the EZ is the seizure-onset zone (SOZ). The SOZ is marked by epileptologists in stereo-electroencephalography (SEEG) recordings as the area where the seizure starts.[4] Alternatively, for patients who have undergone surgery with a seizure-free outcome, the EZ can be defined as an overlap between the SOZ and resected tissue.[5–7]

The average success rate in epilepsy surgery is ≈50% of well-selected candidates.[8] For the other half, the current clinical gold standard is failing. More recently, various research groups have focused on the development of seizure-independent (interictal) biomarkers to assist in correct localization of the EZ.[9] Interictal epileptic discharges ([IEDs] or "spikes") together with power spectral analysis and connectivity measures can serve as a solid estimator for the area that needs to be resected.[10,11] Other authors have investigated the high-frequency (HF) content of the SEEG signals. IEDs with a preceding increase in gamma power (30–100 Hz),[7,12] ripple-IEDs,[13] HF oscillations (HFOs) detected in ripple (80–250 Hz) and fast ripple (250–600 Hz) bands,[14] as well as HF connectivity features are currently seen as the most promising biomarkers of the EZ.[5,14–16]

Key points

- In multi-marker machine learning models, interictal stereo-electroencephalography (SEEG) features below 45 Hz alone provide sufficient information to correctly localize the epileptogenic zone and predict surgical outcome.
- High-frequency features do not provide any additional value to the results of standard low-frequency EEG analysis.
- This approach simplifies the feature calculation process and opens a possibility to use these models even when the SEEG is recorded at a standard sampling rate as commonly done in clinical routines.

Unfortunately, HF features are problematic to record, store, and process, as they require more expensive high-resolution amplifiers and high sampling rates. The Nyquist-Shannon theorem informs us that the sampling rate must be at least twice the bandwidth of the signal.[17] However, to reliably analyze the HF content of the signal (e.g., ripples and fast ripples), a sampling rate of at least 2000 Hz is recommended.[18] Even if the hospital has an acquisition system which is able to record such high sampling rates, it is challenging to record good quality HF signals due to high sensitivity to noise and artifacts.[19]

Until now, the best results in interictal EZ localization were achieved by a combination of the above-mentioned standard frequency band features and HF features in machine learning models.[20–23] However, the question remains whether these HF features are indeed essential. Recent studies show that the standard frequency (low-frequency [LF]) features were sometimes preferred over HF features, and results reported by Travnicek et al. do not show that the feature of relative entropy in the ripple band performs significantly better than relative entropy in standard frequency bands below 20 Hz.[5,20,23] Despite this evidence and the practical need for reduced sampling rates in clinical recordings, there is no systematic study to answer the question how much HF features contribute to correct predictions of the EZ. In this study, we compared three machine learning models, each working with the same recordings but with a different set of features. Model 1 used all available features as done in our previous studies[20,22]; Model 2 was using LF features only (<45 Hz); and Model 3 was using HF features only (>65 Hz). We hypothesized that there would be no statistically significant differences between these three models, in terms of EZ localization and outcome prediction.

2 | METHODS

2.1 | Patients

We analyzed all consecutive adult patients with drug-resistant focal epilepsy who had complete SEEG data sets available for analysis, scalp EEG, electro-oculography, and electromyography or video for sleep scoring, high-resolution three-dimensional (3D) magnetic resonance imaging (MRI) data sets, and subsequent resective surgery at the Montreal Neurological Institute & Hospital (MNI) between January 2010 and December 2015 and at the St. Anne's University Hospital in Brno (SAUH) between April 2017 and October 2019 with post-surgical follow up ≥1 year. The patients undergoing SEEG prior to this period did not have additional scalp EEG with subdermal thin wire electrodes required for sleep staging. Additional inclusion criteria were: (1) less than one seizure per hour; (2) resection cavity overlapped with SOZ region; and (3) availability of ≥24 h of continuous SEEG recording. Patients were divided into two groups according to their post-surgical outcome: (1) seizure-free patients defined as Engel class IA; (2) not seizure-free patients defined as Engel class IB–IVD.

2.2 | Recordings

Standard clinical SEEG electrodes were inserted stereotactically using an image-guided system. SEEG was recorded at 2 kHz (MNI) and 5 kHz (SAUH). All SEEG signals were processed as unipolar signals—an average from all SEEG contacts with a confirmed location inside the brain was used as a reference (bipolar montage was not used for its reduction in the spatial resolution for EZ localization, which would affect the results, especially for connectivity features). The white matter was not excluded. In the MNI data set, scalp EEG for sleep scoring was obtained with subdermal thin wire electrodes.[24] In the SAUH data set, three scalp electrodes covering the frontal, central, and parietal regions, electrodes for electrooculography and electromyography were used. Sleep was scored visually in 30-s epochs.[25]

To test our hypothesis on the best possible data, we focused on recordings of non–rapid eye movement (NREM) sleep, as suggested by previous studies.[22,23] To account for fast-ultradian dynamics, we selected six segments of 5 min long, artifact-free NREM sleep stage (30 min in total) for every patient, randomly from the first, second, and, if available, third nights of recording.[6] The selected segments were at least 1 h from any epileptic seizure. Electrode contact localization was determined as done previously using the IBIS SEEG plugin.[26,27] The SOZ was identified visually by board-certified epileptologists. It comprises contacts exhibiting the earliest changes at seizure onset.[28] An epileptogenicity index >0.3 was not required for labeling a contact as belonging to the SOZ.[29] The SEEG contacts in the resection cavity were identified from pre- and post-surgical imaging. To account for sagging, co-registration error, and the effect of partial contact resection, contacts in the close vicinity of the cavity (<5 mm) were considered as resected as well, as done in the previous study.[30]

2.3 | Feature calculation

Multiple univariate, bivariate, and event-related EEG features were calculated for each SEEG contact (or pair of adjacent SEEG contacts) using our open-source python library EPYCOM, as used previously.[5,20,22] The list of all calculated features can be seen in Figure S1. For feature details please refer to the EPYCOM library (https://gitlab.com/bbeer_group/development/epycom/epycom) or directly to the feature documentation (https://gitlab.com/bbeer_group/development/epycom/epycom/-/blob/develop/doc/feature_extraction.rst). The features were calculated in the following frequency bands: 1–4 Hz, 4–8 Hz, 8–12 Hz, 12–20 Hz, 20–45 Hz, 65–80 Hz, 80–250 Hz, 250–600 Hz and for the raw unfiltered recordings 1–600 Hz for MNI, 1–1200 Hz for SAUH.

Event-related features were calculated for single SEEG contacts, namely IED rates, detected by a validated detector,[31] gamma-IEDs,[7,12] HFO rates in ripple (80–250 Hz), and fast ripple (250–600 Hz) bands, detected by the MNI detector.[32]

Univariate and bivariate features were calculated in 1 s windows with zero overlap. Then, a median value was

calculated across 30 min of analyzed recording. Rates of detected IEDs, gamma-IEDs, and HFOs were calculated per second.

All calculated features were further divided into two groups, LF group and HF group, by the following rules: If the feature was calculated on signals filtered below 45 Hz, it was assigned to the LF group. If the feature was calculated on signals filtered above 65 Hz, it was assigned to the HF group. Otherwise, if the feature was calculated on raw, unfiltered signals, it was not assigned to any group. The thresholds of 45 and 65 Hz were chosen to avoid line noise (50 Hz Europe, 60 Hz North America) and to fully cover standard electrophysiological frequencies up to the beta band.

2.4 | Machine learning

Features were normalized by a *z*-score within each patient. We built three logistic regression models, each based on a different set of features. Model 1 used all calculated features; Model 2 used LF features only; and Model 3 used HF features only.

Optimal hyperparameters of the logistic regression models were found by standard grid search, using all available features and *F1* score as a performance metric. The hyperparameters that performed the best across all cross-validation folds were further used for all three models. Each model was then trained and tested separately, with zero knowledge of other models or features. A feature selection algorithm was applied for each leave-one-patient-out iteration, resulting in a unique set of features for each run. The imbalance between classes was handled in the logistic regression model by the "class_weight" parameter, which automatically adjusts weights inversely proportional to class frequencies in the input data.

In general, a logistic regression model might not benefit from potentially highly correlated features. However, it is important not to discard any complementary features that might contribute to the model's stability and robustness. For the feature selection, the features were grouped by a hierarchical clustering on the Spearman rank-order correlations by a parameter of 1.5, set manually after visual assessment of dominant clusters. These clusters were chosen to avoid unnecessary features in the model, while keeping the features from different feature families which might complement each other. The best-performing features within each cluster were identified by outliers of the analysis of variance (ANOVA) F-score values, as done in a previous study.[20] From each cluster, only the best performing feature was selected.

The models were then trained on data from seizure-free patients with the localization target defined as the SOZ contacts. Resected contacts, not marked as the SOZ, were ignored in the training. Evaluation was done using a leave-one-patient-out cross-validation—omitting the tested patient from the training set (in case of not seizure-free patients, all seizure-free patients were used for training).

2.5 | Evaluation of target localization

The target was defined as the resected area because it allows us to perform the outcome prediction in the most intuitive way.[33] Therefore, the predictions of the models were summarized as True Positives (TP), resected contacts marked by the model as a target; False Positives (FP), not resected contacts marked as target; True Negatives (TN), not resected contacts not marked as target; and False Negatives (FN), resected contacts not marked as target. The performance of the models was then evaluated using seizure-free patients in terms of accuracy, sensitivity (also known as recall), specificity, *F1* score, positive predictive value (PPV; also known as precision) and additionally by the area under the receiver-operating characteristic (ROC) curve (AUROC) and area under the precision-recall curve (AUPRC). Every leave-one-patient-out iteration produced one score of each metric. The differences between these scores across different models were statistically tested by paired, non-parametric Wilcoxon signed-rank tests with Bonferroni correction for three comparisons and Cliff's Delta effect size, interpreted as negligible, small, medium, or large effect size, respectively, by thresholds .147, .33, or .474. Any differences with a p-value <.05 and medium or large effect size >.33 were considered significant.[34]

2.6 | Evaluation of outcome prediction

The outcome prediction was done as in our previous study.[33] Briefly, a PPV was calculated for each patient as PPV = TP/(TP + FP), reflecting how many from all predicted electrodes were resected. Subsequently, seizure freedom was predicted for patients with a PPV higher than .5 (more than 50% of predicted contacts were resected).

To test whether HF features bring any additional value to the outcome prediction, we compared outcome predictions made by Model 2 (LF features only) and by Model 1 (all features) using contingency tables and the McNemar test.[35] In addition, to test whether HF features alone perform reasonably well, we compared the outcome predictions made by Model 3 (HF features only) to the two other models. Any differences in the McNemar test with a p-value <.05 were considered significant. The power of the McNemar test was calculated for the proportions of discordant pairs and alpha levels found by our study.[36]

2.7 | Alternative models and targets

To test for robustness of the results we further tested different machine learning models potentially suitable for this type of data. Namely support vector machine (SVM) with linear kernel, SVM with radial basis function (RBF) kernel, and, to eliminate any potential bias in the feature selection strategy, random forests without any previous feature selection. Furthermore, we calculated the feature importance using mean decrease in impurity (MDI), which is a measure of the homogeneity of the labels in a node. Lower impurity signifies a more homogeneous node. The primary goal of using impurity measures is to guide the decision tree algorithm in creating the most informative and pure splits. MDI was calculated across decision trees on the full data set of seizure-free patients (no patient was left out for this calculation). The alternative models were trained and tested on the same targets as the logistic regression model.

To account for scenarios with smaller targets, which might be more challenging for less-focal (less-specific) markers, we tested the logistic regression model on different training and testing targets: (1) SEEG contacts marked as SOZ; (2) resected SEEG contacts; and (3) resected SEEG contacts marked as SOZ.

Calculated features and python codes for subsequent machine learning and statistics are available at https://gitlab.com/bbeer_group/research/neurology/hflf

3 | RESULTS

3.1 | Patients

The initial patient cohort consisted of 57 patients. A total of seven patients were excluded as per selection criteria: two patients had ≥1 seizure per hour, one patient did not have an overlap between the SOZ and the resection cavity, and four patients did not have the availability of ≥24 h of continuous SEEG recording. The final patient cohort consisted of 50 patients fulfilling the selection criteria (SAUH = 17, MNI = 33), with 23 female. The median age of the group was 32.5 (interquartile range [IQR] = 11) years. The median number of implanted contacts was 88 (IQR = 92) and the median number of SOZ contacts was 9 (IQR = 6.75) (class imbalance ratio was .098 (IQR = .131)). The SOZ region was completely removed in 26 patients. The complete or partial removal of the SOZ correctly predicted the post-surgical outcome in 38 patients (individual patient predictions are available online at the GitLab repository). A total of 44% had temporal epilepsy; 64% were lesional cases. The two most frequent pathologies were focal cortical dysplasia (FCD) (50%) and gliosis (24%). Eighteen patients had Engel class IA post-surgical outcome. (For further information please see Table S1).

3.2 | Model hyperparameters and selected features

The winning hyperparameters for the logistic regression model were the Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (LBFGS) solver, with "l2" penalty and C = 1. The best performing features in Model 1 were IED rates, relative entropy in the beta band, and spectral power in the beta band. These three features were selected in 92% of leave-one-patient-out iterations. In one case (patient 82 left out), the low-frequency ratio was selected as an additional feature to the previous three; in two cases (patients 89 and 717 left out), HFO rates in the ripple band were selected instead of IED rates; and in another two cases (patients 717 and 1630 left out), the spectral power in the fast ripple band was selected instead of the spectral power in the beta band. The best-performing features in Model 2 were IED rates, relative entropy in the beta band, and spectral power in the beta band. These three features were selected in all leave-one-patient-out iterations. The best-performing features in Model 3 were HFO rates in the ripple band and relative entropy in the ripple band. These three features were selected in all leave-one-patient-out iterations. All calculated and selected features are provided in Table S2.

3.3 | Target localization results

On average, Model 1 achieved a median (IQR) AUPRC of .608 (.407), Model 2 a AUPRC of .582 (.407), and Model 3 a AUPRC of .522 (.57) (Figure 1 and Table 1). There were no statistically significant differences in the AUPRC or any other metric, when tested by the paired, non-parametric Wilcoxon signed-rank tests after Bonferroni correction (all p's > .05) and Cliff's Delta (all effect sizes <.33). (The complete results of statistical tests can be found in the Table S3).

3.4 | Outcome prediction results

On average, Model 1 achieved a median (IQR) PPV of .414 (.381) and .211 (.237) in seizure-free and not seizure-free patients, respectively. Model 2 achieved a PPV of .406 (.401) and .211 (.237), and Model 3 a PPV of .347 (.446) and .194 (.274) (Figure 2). When the threshold of .5 was applied for the outcome prediction, Model 2 (LF features

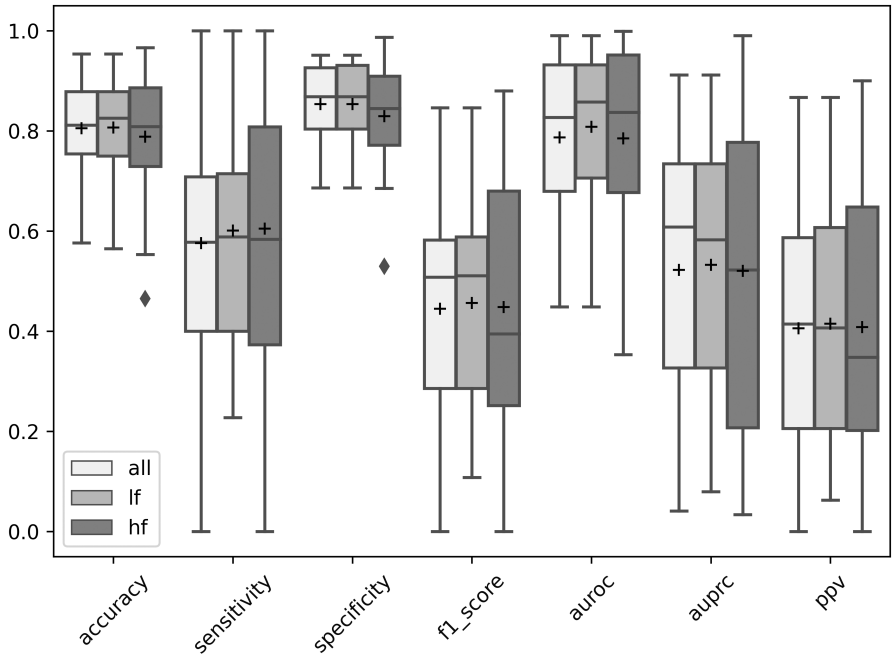

FIGURE 1 Scores of various metrics for three analyzed models in all seizure-free patients ($N = 18$) visualized by box plots. The box shows the quartiles of the data set, whereas the whiskers extend to show the rest of the distribution, except for diamond points that are determined to be outliers. The middle line shows the median, whereas the + symbol represents the mean. There were no statistically significant differences between the models (all p's > .05, all effect sizes <.33).

TABLE 1 Scores of various metrics for three analyzed models.

Performance of the models (values from the Figure 1)			
	Model 1 (All)	**Model 2 (LF)**	**Model 3 (HF)**
Accuracy			
Mean	.805	.807	.788
Median	.811	.825	.809
Sensitivity			
Mean	.576	.601	.605
Median	.578	.588	.583
Specificity			
Mean	.854	.853	.829
Median	.868	.868	.845
F1			
Mean	.445	.456	.448
Median	.508	.511	.394
AUROC			
Mean	.787	.808	.785
Median	.827	.858	.837
AUPRC			
Mean	.522	.533	.521
Median	.608	.583	.522
PPV			
Mean	.406	.415	.408
Median	.414	.407	.348

Abbreviations: AUPRC, area under the precision-recall curve; AUROC, area under the receiver-operating characteristic (ROC) curve; PPV, positive predictive value.

only) correctly predicted 38 of 50 patients. After adding HF features to the model (using Model 1 with all features), the model correctly predicted the same 38 of 50 patients (Table S4). The McNemar test did not show any difference between these models ($p = 1$).

Furthermore, predictions made by Model 3 (HF features only) were correct in 33 of 50 patients. Thirty-two of these patients were correctly predicted by all models. Eleven patients were wrongly predicted by all models. Six patients were wrongly predicted by Model 3 only (Tables S5 and S6). The McNemar test did not show any statistically significant difference between Model 3 and Models 1 and 2 (both ps = .125). The power of the McNemar test was .68, which means there is a 32% chance of a false-negative result (type II error).

In our data set, we identified an optimal PPV threshold for the outcome prediction .4, resulting in a lower number of seizure-free patients identified as not seizure-free. However, we decided to keep the original 50% to be consistent with the previous work.

3.5 | Alternative models and targets

Feature importance calculated by the mean decrease in impurity identified IED rates, HFO ripple rates, and gamma-IED rates as the most important features all results are available in Figure S1. Subsequent feature grouping by a hierarchical clustering on the Spearman rank-order correlations are visualized in Figure S2.

Logistic regression model with different training and testing targets showed results similar to those of the benchmark logistic regression model in both EZ localization and outcome prediction. Median (IQR) number of SEEG contacts within each target and their class imbalance ratios were: 9 (6.75) SEEG contacts with .098 (.131)

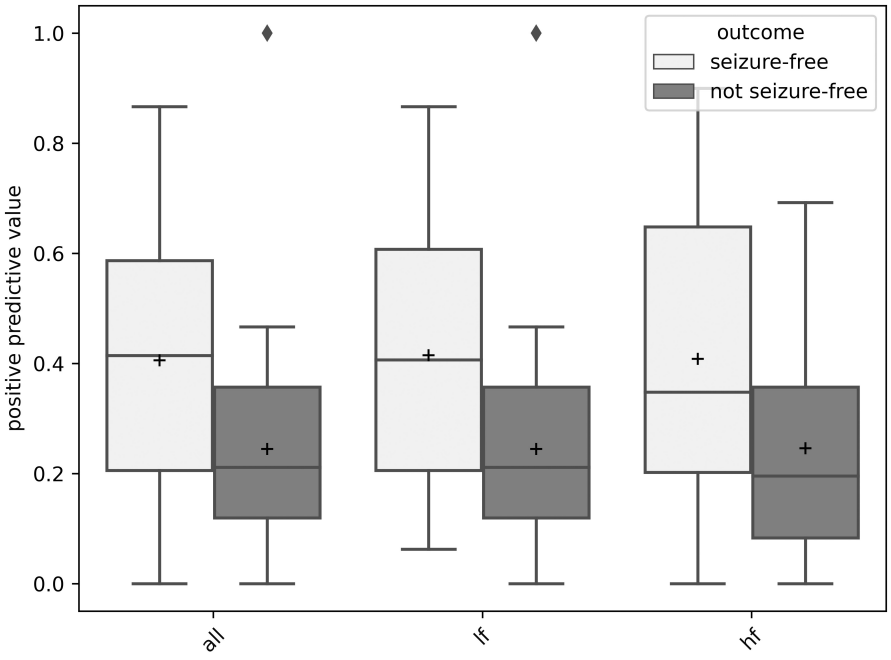

FIGURE 2 Positive predictive value scores for three analyzed models in all patients ($N = 50$) visualized by box plots. The box shows the quartiles of the data set, whereas the whiskers extend to show the rest of the distribution, except for diamond points that are determined to be outliers. The middle line shows the median, whereas the + symbol represents the mean.

class imbalance ratio for the target defined as SOZ; 11 (11.75) SEEG contacts with .108 (.182) class imbalance ratio for the target defined as resected contacts; and 5.5 (5) SEEG contacts with .057 (.059) class imbalance ratio for the target defined as resected SOZ. All results are reported in Tables S7–S21.

Alternative models showed results similar to those of the benchmark logistic regression model in both EZ localization and outcome prediction. All results are reported in Tables S22–S30.

4 | DISCUSSION

In this study we analyzed 50 patients with drug-resistant focal epilepsy and tested different models for EZ localization and outcome prediction to evaluate if HF SEEG features allow us to add any additional information to standard SEEG features. Our results showed that HF features do not provide any additional value to the results of standard LF EEG analysis.

Evaluation of the feature selection showed that the LF model was driven exclusively by IED rates, relative entropy[5] in the beta band, and spectral power in the beta band. The HF model used HFO rates in the ripple band and relative entropy in the ripple band. Of interest, HFOs in the fast ripple band have not been selected by the HF model, even though they had been reported as a more specific biomarker of the EZ.[37] However, they are not present in all patients and are more sensitive to noise.[33] At the same time, given the size of the target (resected volume), less-localized markers (IEDs, ripples) seem to provide sufficient information. The first model, which had all features available, selected most often IED rates together with relative entropy and spectral power in the beta band, similar to the LF model. Other features selected by this model were the LF ratio,[38] power in the fast ripple band, and HFO rates in the ripple band.

The decision trees, which used the full data set, identified IEDs, HFOs in the ripple band, and gamma-IEDs as the best features. This is in agreement with the feature selection used for the logistic regression. However, due to the high correlation between these three features, the model would not benefit from including all of them. Previous clustering of features and selecting the best-performing feature from each cluster help to preserve different physiological phenomena of selected features in the multi-marker approach.

When evaluating the EZ localization performance, the models achieved similar results with no statistically significant differences between them (even without any correction for multiple comparisons). The larger variability of the model, which used only the HF features, might be caused by the lower number of features selected for the model, advocating for our strategy of using a more stable and robust multi-marker approach with features from different feature clusters. EZ localization can be evaluated in seizure-free patients only, since we do not have confirmed location of the EZ in patients who continued to have seizures. Biomarkers of EZ and evaluation of their EZ localization potential are problematic in general.[39] Therefore, in the next step, we performed outcome prediction on the full data set with 50 patients, based on individual models, and compared the results of this prediction. When using the model with LF features only, the model achieved results similar to

those of the previous study.[22] After HF features were added, the predictions did not change and the McNemar test did not show any difference between these models ($p = 1$). Subsequently, predictions made by the model that used only the HF features showed that HFOs in the ripple band together with relative entropy in the ripple band work reasonably well. This model did not differ significantly from the other two models, even though it wrongly classified six more patients (705, 909, 1002, 1159, 1162, and 1299). All of these patients had temporal lobe epilepsy, five of them with gliosis in histopathology and persisting seizures. Of interest, the HF model correctly classified patient 717, which was misclassified by the other two models. Patient number 717 was a male treated in Montreal with temporal lobe epilepsy, with gliosis as the main histopathological finding and seizure-free outcome.

Alternative machine learning models and targets achieved similar results. SVM with linear kernel worked, in some cases, even slightly better than the logistic regression model (not tested for statistical significance). Logistic regression is, however, less complex and well poised for this type of data, and it is therefore preferred over the SVM. Random forest did not perform optimally in terms of sensitivity. On the other hand, thanks to its robustness to the high number of potentially strongly correlated features, we were able to use this model without any previous feature selection, thereby eliminating any potential bias in our default feature selection strategy.

In general, machine learning models achieve results that are similar to the current clinical gold standard of the SOZ while working with interictal data. They are fast (hours vs weeks) and objective. Whether they bring additional value to SOZ in terms of the precise localization of the EZ is an ongoing discussion in the community and challenging to test, since they are trained on SOZ. However, we strongly believe that they provide important information. In our data set, 10 patients with complete SOZ resection (patients 74, 80, 84, 88, 90, 1149, 1159, 1162, 1218, and 1246) and not seizure-free outcome and 1 patient (1041) with incomplete resection and seizure-free outcome were correctly classified by the machine learning model. On the other hand, 11 different patients were wrongly classified by the machine learning model and correctly classified by the "complete SOZ resection" model (patients 61, 63, 71, 77, 82, 93, 473, 657, 717, 965, and 1043). A higher proportion of seizure-free patients wrongly classified by the machine learning model points to its lower specificity, which needs to be addressed before its clinical utilization.

The above-presented results suggest that the HF features of SEEG are not necessary for correct EZ localization and outcome prediction. We arrived at these conclusions by objective and data-driven analysis, which used the most common expert features of SEEG, such as IED and HFO rates, well-defined feature selection algorithms, and standard machine learning models, suitable for this type of data. Our results are further supported by other studies that used IEDs (LF feature) as the only one or one of the main SEEG features to localize the EZ or prognosticate epilepsy surgery outcome.[40–44] Furthermore, Lundstrom et al. recently published a novel LF interictal marker based on the ratio of infra-slow activity and faster frequencies such as delta (2–4 Hz) and beta-gamma (20–50 Hz).[38] This feature was actually selected by our first model in one iteration of the leave-one patient out process. Another study by Vasickova et al. identified shadows of HF SEEG components in lower frequency bands.[45] Studies by Thomas et al. and Shi et al. show potential promise in the combination of LF and HF features of IEDs with gamma power or ripples and their better specificity when compared to HFO rates alone, while showing that regular IEDs achieve a similar performance.[7,13] Furthermore, as mentioned earlier in the introduction, the study by Cimbalnik et al. shows that the LF features of SEEG were, in some cases, preferred over HF features in the EZ localization algorithm, and results by Travnicek et al. do not show that the feature of relative entropy performs significantly better in higher frequencies, compared to standard frequency bands below 20 Hz.[5,20] A study by Roehri et al. shows that the performance of HFOs is weakened by the presence of strong physiological HFO generators and that fast ripples are not sufficiently sensitive to be the unique biomarker of epileptogenicity.[46] Finally, a randomized, single-blind, adaptive non-inferiority trial reported that HFO-guided tailoring of epilepsy surgery was not non-inferior to spike-guided tailoring on intraoperative electrocorticography.[47]

We, however, strongly believe that the HF component of the SEEG signal is important for our understanding of the basic mechanisms of human brain electrophysiology. The phenomenon of HFOs (80–500 Hz),[14] very-fast ripples (500–1000 Hz), ultra-fast ripples (1000–2000 Hz),[37] and ultra-fast oscillations recorded in vivo in human brain by microelectrodes (>2 kHz)[48] together with mathematical modeling, can lead to better understanding of the generation mechanisms of these events, signatures of epileptic disorders, seizures, and interictal markers.[49–51] In general, HF features are believed to be more local and, therefore, more specific for the EZ. However, this appears to be true under the condition that the SEEG electrode is correctly placed.[33] If the electrode is placed further from the source, dispersed IEDs and thresholding of their rates by a well-trained machine learning

model seems to provide enough information to correctly localize the region that needs to be resected. We might hypothesize that when the EZ is missed, the LF features have an advantage over the HF ones. Unfortunately, our results cannot confirm or disprove this statement. Further analysis aiming at how each marker is focal is needed to properly answer this question.

A potential limitation of this study might be the size of our data set and the imbalance of pathologies, especially caused by a high representation of FCD cases, which are typically over 40% across different centers and SEEG studies.[52,53] Higher number of FCD cases could overestimate the importance of the IED rates in the model. However, the main focus of the study was solely on LF v HF features. The performance of individual features was not tested. In addition, the length of analyzed signals, which is 30 min, can miss important variability in the data. However, the data segments were selected from the first, second, and, if available, third night of SEEG recording, further supported by the recent works by Dellavale et al. and Chybowski et al., which show that even shorter recordings might provide sufficient information.[6,23] Another potential limitation of this study is utilization of a single model of logistic regression and single feature selection strategy. Even though we tested different models potentially suitable for this type of data and different feature selection strategies, it cannot be excluded that specific settings of the feature selection algorithm and machine learning model might find HF features useful, as well as the inclusion of the anatomic localization of individual SEEG contacts, which is ignored in the current model. Furthermore, the localization target in this study was defined as all resected SEEG contacts. This target might be, in some cases, unnecessarily big, favoring less-specific markers. Tests on different, smaller targets did not show any statistically significant differences between the models, but a more precise delineation of EZ, for example, in seizure-free patients after thermocoagulation of small EZs as present for some patients during clinically indicated SEEG, might potentially find the HF features more relevant.

5 | CONCLUSION

In multi-marker machine learning models, interictal SEEG features below 45 Hz alone provide enough information to correctly localize the EZ and predict surgical outcome. HF features did not provide any additional value. This finding is important, as it can significantly simplify the feature calculation process and opens the possibility to use these models even when the SEEG is recorded at a standard sampling rate as commonly done in clinical routines.

AUTHOR CONTRIBUTIONS

Conception and design of the study: P.K. and B.F. Acquisition and analysis of the data: P.K., P.N., V.H., J.C., V.T., M.P., L.P.D., Je.H., R.P., and Jo.H. Drafting of the manuscript and preparing the figures: P.K., P.J., M.B., and B.F.

ACKNOWLEDGMENTS

We wish to express our gratitude to Prof. Jean Gotman, PhD and people from his lab for providing excellent feedback on methods and results interpretation. We further thank Nicolas von Ellenrieder, PhD, and John Thomas, PhD, for providing their codes for HFO and gamma-IED detections. This work was supported by project no. LX22NPO5107 (MEYS): Financed by European Union—Next Generation EU; Czech Science Foundation, project 22-28784S; Ministry of Health of the Czech Republic, project NU22-08-00278; The CAS project RVO:68081731; and by the Ministry of Education, Youth and Sport of the Czech Republic (EATRIS-CZ, LM2023053). Open access publishing facilitated by Ustav pristrojove techniky Akademie ved Ceske republiky, as part of the Wiley - CzechELib agreement.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interst.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The study was approved by the MNI and SAUH Ethics Review Board. 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.

CONSENT

All patients granted written informed consent.

ORCID

Petr Klimes https://orcid.org/0000-0002-0232-9518
Vojtech Travnicek https://orcid.org/0000-0002-5846-6722
Laure Peter-Derex https://orcid.org/0000-0002-9938-9639
Jeffery Hall https://orcid.org/0000-0001-8403-8986
Milan Brazdil https://orcid.org/0000-0001-7979-2343
Birgit Frauscher https://orcid.org/0000-0001-6064-1529

REFERENCES

1. Thijs RD, Surges R, O'Brien TJ, Sander JW. Epilepsy in adults. Lancet. 2019;393:689–701.
2. Jehi L. The epileptogenic zone: concept and definition. Epilepsy Curr. 2018;18:12–6.

3. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: a 30-year longitudinal cohort study. JAMA Neurol. 2018;75:279–86.
4. Rosenow F, Lüders H. Presurgical evaluation of epilepsy. Brain. 2001;124:1683–700.
5. Travnicek V, Klimes P, Cimbalnik J, Halamek J, Jurak P, Brinkmann B, et al. Relative entropy is an easy-to-use invasive electroencephalographic biomarker of the epileptogenic zone. Epilepsia. 2023;64:962–72.
6. Dellavale D, Bonini F, Pizzo F, Makhalova J, Wendling F, Badier JM, et al. Spontaneous fast-ultradian dynamics of polymorphic interictal events in drug-resistant focal epilepsy. Epilepsia. 2023;64:2027–43.
7. Thomas J, Kahane P, Abdallah C, Avigdor T, Zweiphenning WJEM, Chabardes S, et al. A subpopulation of spikes predicts successful epilepsy surgery outcome. Ann Neurol. 2023;93:522–35.
8. Krucoff MO, Chan AY, Harward SC, Rahimpour S, Rolston JD, Muh C, et al. Rates and predictors of success and failure in repeat epilepsy surgery: a meta-analysis and systematic review. Epilepsia. 2017;58:2133–42.
9. Bernabei JM, Li A, Revell AY, Smith RJ, Gunnarsdottir KM, Ong IZ, et al. Quantitative approaches to guide epilepsy surgery from intracranial EEG. Brain. 2023;146:2248–58.
10. van Mierlo P, Papadopoulou M, Carrette E, Boon P, Vandenberghe S, Vonck K, et al. Functional brain connectivity from EEG in epilepsy: seizure prediction and epileptogenic focus localization. Prog Neurobiol. 2014;121:19–35.
11. Lagarde S, Roehri N, Lambert I, Trebuchon A, McGonigal A, Carron R, et al. Interictal stereotactic-EEG functional connectivity in refractory focal epilepsies. Brain. 2018;141:2966–80.
12. Ren L, Kucewicz MT, Cimbalnik J, Matsumoto JY, Brinkmann BH, Hu W, et al. Gamma oscillations precede interictal epileptiform spikes in the seizure onset zone. Neurology. 2015;84:602–8.
13. Shi W, Shaw D, Walsh KG, Han X, Eden UT, Richardson RM, et al. Spike ripples localize the epileptogenic zone best: an international intracranial study. Brain. 2024;147(7):2496–506. https://doi.org/10.1093/brain/awae037
14. Frauscher B, Bartolomei F, Kobayashi K, Cimbalnik J, van 't Klooster MA, Rampp S, et al. High-frequency oscillations: the state of clinical research. Epilepsia. 2017;58:1316–29.
15. Warren CP, Hu S, Stead M, Brinkmann BH, Bower MR, Worrell GA. Synchrony in normal and focal epileptic brain: the seizure onset zone is functionally disconnected. J Neurophysiol. 2010;104:3530–9.
16. Klimes P, Duque JJ, Brinkmann B, van Gompel J, Stead M, St. Louis EK, et al. The functional organization of human epileptic hippocampus. J Neurophysiol. 2016;115:3140–5.
17. Shannon C, Weaver W. The Mathematical Theory of Communication.
18. Gliske SV, Irwin ZT, Chestek C, Stacey WC. Effect of sampling rate and filter settings on high frequency oscillation detections. Clin Neurophysiol. 2016;127:3042–50.
19. Bénar CG, Chauvière L, Bartolomei F, Wendling F. Pitfalls of high-pass filtering for detecting epileptic oscillations: a technical note on 'false' ripples. Clin Neurophysiol. 2010;121:301–10.
20. Cimbalnik J, Klimes P, Sladky V, Nejedly P, Jurak P, Pail M, et al. Multi-feature localization of epileptic foci from interictal, intracranial EEG. Clin Neurophysiol. 2019;130:1945–53.
21. Varatharajah Y, Berry B, Cimbalnik J, Kremen V, van Gompel J, Stead M, et al. Integrating artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy. J Neural Eng. 2018;15:046035.
22. Klimes P, Cimbalnik J, Brazdil M, Hall J, Dubeau F, Gotman J, et al. NREM sleep is the state of vigilance that best identifies the epileptogenic zone in the interictal electroencephalogram. Epilepsia. 2019;60:2404–15.
23. Chybowski B, Klimes P, Cimbalnik J, Travnicek V, Nejedly P, Pail M, et al. Timing matters for accurate identification of the epileptogenic zone. Clin Neurophysiol. 2024;161:1–9.
24. Ives JR. New chronic EEG electrode for critical/intensive care unit monitoring. J Clin Neurophysiol. 2005;22:119–23.
25. Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, et al. AASM scoring manual updates for 2017 (Version 2.4). J Clin Sleep Med. 2017;13:665–6.
26. Frauscher B, von Ellenrieder N, Zelmann R, Doležalová I, Minotti L, Olivier A, et al. Atlas of the normal intracranial electroencephalogram: neurophysiological awake activity in different cortical areas. Brain. 2018;141:1130–44.
27. Zelmann R, Frauscher B, Aro RP, Gueziri H-E, Collins DL. SEEGAtlas: a framework for the identification and classification of depth electrodes using clinical images. J Neural Eng. 2023;20:036021.
28. Spanedda F, Cendes F, Gotman J. Relations between EEG seizure morphology, interhemispheric spread, and mesial temporal atrophy in Bitemporal epilepsy. Epilepsia. 1997;38:1300–14.
29. Bartolomei F, Chauvel P, Wendling F. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. Brain. 2008;131:1818–30.
30. Cuello Oderiz C, von Ellenrieder N, Dubeau F, Eisenberg A, Gotman J, Hall J, et al. Association of cortical stimulation-induced seizure with surgical outcome in patients with focal drug-resistant epilepsy. JAMA Neurol. 2019;76:1070–8.
31. Janca R, Jezdik P, Cmejla R, Tomasek M, Worrell GA, Stead M, et al. Detection of Interictal Epileptiform discharges using signal envelope distribution modelling: application to epileptic and non-epileptic intracranial recordings. Brain Topogr. 2015;28:172–83.
32. von Ellenrieder N, Frauscher B, Dubeau F, Gotman J. Interaction with slow waves during sleep improves discrimination of physiologic and pathologic high-frequency oscillations (80-500 Hz). Epilepsia. 2016;57:869–78.
33. Nevalainen P, von Ellenrieder N, Klimeš P, Dubeau F, Frauscher B, Gotman J. Association of fast ripples on intracranial EEG and outcomes after epilepsy surgery. Neurology. 2020;95:e2235–e2245.
34. Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. Psychol Bull. 1993;114:494–509.
35. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika. 1947;12:153–7.
36. Miettinen OS. The matched pairs design in the case of all-or-none responses. Biometrics. 1968;24:339–52.

37. Brázdil M, Pail M, Halámek J, Plešinger F, Cimbálník J, Roman R, et al. Very high-frequency oscillations: novel biomarkers of the epileptogenic zone. Ann Neurol. 2017;82:299–310.
38. Lundstrom BN, Brinkmann BH, Worrell GA. Low frequency novel interictal EEG biomarker for localizing seizures and predicting outcomes. Brain Commun. 2021;3:fcab231.
39. Gliske SV, Stacey WC. The BEST conceivable way to talk about epilepsy biomarkers. Epilepsy Curr. 2023;23:175–8.
40. Klimes P, Peter-Derex L, Hall J, Dubeau F, Frauscher B. Spatiotemporal spike dynamics predict surgical outcome in adult focal epilepsy. Clin Neurophysiol. 2022;134:88–99.
41. Tomlinson SB, Porter BE, Marsh ED. Interictal network synchrony and local heterogeneity predict epilepsy surgery outcome among pediatric patients. Epilepsia. 2017;58:402–11.
42. Azeem A, von Ellenrieder N, Hall J, Dubeau F, Frauscher B, Gotman J. Interictal spike networks predict surgical outcome in patients with drug-resistant focal epilepsy. Ann Clin Transl Neurol. 2021;8:1212–23.
43. 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;139:1066–78.
44. Conrad EC, Revell AY, Greenblatt AS, Gallagher RS, Pattnaik AR, Hartmann N, et al. Spike patterns surrounding sleep and seizures localize the seizure-onset zone in focal epilepsy. Epilepsia. 2023;64:754–68.
45. Vasickova Z, Klimes P, Cimbalnik J, Travnicek V, Pail M, Halamek J, et al. Shadows of very high-frequency oscillations can be detected in lower frequency bands of routine stereoelectroencephalography. Sci Rep. 2023;13:1065.
46. Roehri N, Pizzo F, Lagarde S, Lambert I, Nica A, McGonigal A, et al. High-frequency oscillations are not better biomarkers of epileptogenic tissues than spikes. Ann Neurol. 2018;83:84–97.
47. Zweiphenning W, Klooster MAV', van Klink NEC, et al. Intraoperative electrocorticography using high-frequency oscillations or spikes to tailor epilepsy surgery in The Netherlands (the HFO trial): a randomised, single-blind, adaptive non-inferiority trial. Lancet Neurol. 2022;21:982–93.
48. Brázdil M, Worrell GA, Trávníček V, Pail M, Roman R, Plešinger F, et al. Ultra fast oscillations in the human brain and their functional significance. medRxiv. 2023. https://doi.org/10.1101/2023.02.23.23285962
49. Jiruska P, Alvarado-Rojas C, Schevon CA, Staba R, Stacey W, Wendling F, et al. Update on the mechanisms and roles of high-frequency oscillations in seizures and epileptic disorders. Epilepsia. 2017;58:1330–9.
50. Sklenarova B, Zatloukalova E, Cimbalnik J, Klimes P, Dolezalova I, Pail M, et al. Interictal high-frequency oscillations, spikes, and connectivity profiles: a fingerprint of epileptogenic brain pathologies. Epilepsia. 2023;64:3049–60.
51. Přibylová L, Ševčík J, Eclerová V, Klimeš P, Brázdil M, Meijer HGE. Weak coupling of neurons enables very high-frequency and ultra-fast oscillations through the interplay of synchronized phase shifts. Netw Neurosci. 2024;8:293–318. https://doi.org/10.1162/netn_a_00351
52. Di Giacomo R, Uribe-San-Martin R, Mai R, et al. Stereo-EEG ictal/interictal patterns and underlying pathologies. Seizure. 2019;72:54–60.
53. Cardinale F, Rizzi M, Vignati E, et al. Stereoelectroencephalography: retrospective analysis of 742 procedures in a single centre. Brain. 2019;142:2688–704.

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: Klimes P, Nejedly P, Hrtonova V, Cimbalnik J, Travnicek V, Pail M, et al. Interictal stereo-electroencephalography features below 45 Hz are sufficient for correct localization of the epileptogenic zone and postsurgical outcome prediction. Epilepsia. 2024;65:2935–2945. https://doi.org/10.1111/epi.18081