Patient-specific optimization of automated detection improves seizure onset zone localization based on high frequency oscillations

Casey L. Trevino[1], Jack J. Lin[1,2], Indranil Sen-Gupta[2]*, Beth A. Lopour[1]*

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

[2] Comprehensive Epilepsy Program, Department of Neurology, University of California, Irvine, Irvine, CA, USA

Correspondence:

Beth A. Lopour (beth.lopour@uci.edu)

Department of Biomedical Engineering

University of California, Irvine

* These authors contributed equally to this work.

Highlights

- Optimal detection of high frequency oscillations (HFOs) improves demarcation of seizure onset zone (SOZ) channels.
- Optimal detection parameters for HFOs are patient-specific and vary over a wide range.
- Assessment of SOZ localization accuracy can be impacted by an imbalanced number of SOZ and non-SOZ channels.

Keywords: Intracranial electroencephalogram, Parameter optimization, Algorithm, Epilepsy surgery, Ripple, Epileptogenic zone

Abstract

Objective: High frequency oscillations (HFOs) are a promising biomarker of epileptogenicity, and automated algorithms are critical tools for their detection. However, it is not uncommon for a previously validated algorithm to work poorly when applied to a new data set. There is no consensus on whether (or how) parameters should be optimized. Here we evaluate the impact of parameter selection on seizure onset zone (SOZ) localization using automatically detected HFOs.

Methods: We detected HFOs in intracranial EEG from twenty medically refractory epilepsy patients with seizure free surgical outcomes using an automated algorithm. For each patient, we assessed classification accuracy of channels inside/outside the SOZ using a wide range of detection parameters and identified the parameters associated with maximum classification accuracy.

Results: Only three of twenty patients achieved maximal localization accuracy using conventional HFO detection parameters, and optimal parameter ranges varied significantly across patients. The use of individualized optimal parameters led to substantial improvements in localization accuracy, particularly in reducing HFO rates in non-SOZ channels.

Conclusion: Optimal HFO detection parameters are patient-specific and often differ from conventional parameters. The use of optimal parameters significantly improves SOZ localization.

Significance: Individual variability should be considered when implementing automatic HFO detection, and novel methods for patient-specific optimization are needed.

1. Introduction

For patients with medically refractory epilepsy, surgical resection of the epileptogenic zone (EZ) offers the potential to eliminate the occurrence of seizures (Noachtar & Borggraefe, 2009; Schuele & Lüders, 2008). However, the EZ remains a largely theoretical construct, as there are no biomarkers that enable identification of this brain region (Kovac et al., 2017; Ryvlin et al., 2014). Current clinical practice aims to identify the seizure onset zone (SOZ), currently the best proxy of the EZ, during the patients' pre-surgical evaluation. As a complementary biomarker, epileptiform discharges are considered indicative of the SOZ due to their high degree of association with seizures (Westmoreland, 1996); however, they are not specific to the SOZ (Bautista et al., 1999). Furthermore, patients with spikes occurring in multiple brain areas are less likely to become seizure free after surgery than patients with well localized spikes (Bautista et al., 1999). Overall, only 50-60% of patients are seizure free following resective surgery, emphasizing the need to improve localization of the EZ (Edelvik et al., 2013).

Over the past two decades, there has been growing interest in studying electrographic events called high-frequency oscillations (HFOs) as potential biomarkers of the EZ. HFOs are generally defined as spontaneous EEG patterns that consist of at least four oscillations with frequency > 80 Hz that clearly stand out from the background activity (Bragin et al., 1999). Interictal HFOs have a demonstrated association with the SOZ, as they occur more frequently at the site of seizure onset (Jacobs et al., 2012; Zijlmans et al., 2012). They were first discovered in rodent models of epilepsy, but numerous studies have demonstrated the measurement of HFOs in humans using clinical macro electrodes (Crépon et al., 2010; Worrell et al., 2008). Because HFOs occur interictally, their use in surgical planning could reduce risk and discomfort for patients by minimizing recording time and the necessity of ictal recordings (Migliorelli et al., 2017). It has also been suggested that HFOs are more specific to the SOZ than spikes (Crépon et al., 2010; Jacobs et al., 2008; Staba & Bragin, 2011), although another study found that their co-occurrence with spikes was the best predictor of the SOZ (Roehri et al., 2018). The removal of HFO-generating regions has also been correlated with good post-surgical outcome (Akiyama et al., 2011; Jacobs et al., 2010; J. Y. Wu et al., 2010). Moreover, it was demonstrated that HFOs

are not specific to brain lesions, further strengthening its case as a marker of epileptogenic tissue rather than other pathologic tissue changes (Jacobs et al., 2009). Despite its potential as a biomarker of the EZ, HFO analysis has not yet proven to be reliable in individual patients (Jacobs et al., 2018; Roehri et al., 2018).

One possible reason for the lack of individual patient reliability is the method of detection and analysis. HFOs are sometimes detected through visual identification by expert reviewers (Ferrari-Marinho et al., 2015; Jacobs et al., 2014); however, visual marking is time-consuming, subjective, and is insufficiently reproducible (Spring et al., 2017, 2018; Zelmann et al., 2012). Therefore, automated methods of HFO detection have been developed to address these shortcomings (Sindhu et al., 2020), and this is seen as a critical step to translating HFO analysis into clinical practice (Worrell et al., 2012). Automatic HFO detection algorithms are generally designed to be applied to a specific frequency band, location of the brain, and/or electrode type (Zelmann et al., 2012). When these algorithms are subsequently applied to a new dataset, it is preferable to use the parameter values for the algorithm defined in the original publication, in order to provide independent validation of the algorithm's utility, promote reproducibility of results, and to avoid overfitting to the data. However, the use of such standard configurations can lead to suboptimal performance when used for a different data set or frequency range. For example, studies implementing previously published algorithms without optimizing parameters often produce results worse than originally reported (Burnos et al., 2014; Gardner et al., 2007; M. Wu et al., 2018; Zelmann et al., 2012). Furthermore, studies typically apply the same detection parameters to all patients (Gliske et al., 2016; Jacobs et al., 2018; Zelmann et al., 2012). There is evidence that optimization of HFO detection parameters can improve sensitivity (Zelmann et al., 2012) and localization accuracy (Chaibi et al., 2013; Charupanit & Lopour, 2017; Dümpelmann et al., 2012), however, the optimization procedures in these studies were often limited to small parameter ranges, were performed only to compare performance across detectors, and did not account for patient variability. Thus, it remains unknown whether patient-specific optimization is necessary, which parameter ranges are relevant for HFOs, and how this optimization should be accomplished to support feasibility in clinical settings.

Therefore, we evaluated the significance of parameter selection for automatic HFO detection using two independent datasets. Using a frequently cited automatic detector, we tested a wide range of algorithm parameter sets and determined the SOZ localization accuracy for each one. This allowed us to estimate the best possible localization accuracy using HFOs, assuming that an optimization procedure had been used to select the detection parameters. We also measured the improvement in accuracy when using individualized detection parameters versus the conventional set. Our goal was to determine optimal parameter ranges for HFO detection, assess the variation in results across patients, and investigate the impact of optimization on the accuracy of SOZ localization.

2. Methods

2.1. Clinical data collection

Twenty patients with medically refractory epilepsy from two medical centers were identified retrospectively and included in this study. Seven of these patients underwent implantation of intracranial electrodes for presurgical evaluation from June 2015 to March 2017 at the University of California, Irvine (UCI) Medical Center. We will refer to this as the UCI dataset. We included only patients with a postoperative outcome of Engel Class I, which suggests that the clinical SOZ localization was successful. For each patient, the SOZ was defined as electrodes involved within the first three seconds of seizure onset, based on visual assessment of the ictal intracranial EEG (iEEG). If patients had multiple seizures with different regions of onset during recording, all SOZ channels were considered SOZ. The demographic and clinical characteristics of these patients are shown in Table 1. Collection and analysis of retrospective patient data for this study was approved by the Institutional Review Board of the University of California, Irvine.

Table 1. *Patient demographics of the UCI dataset. Abbreviations: AM = amygdala; AH = anterior hippocampus; depth = depth electrode; FLE = frontal lobe epilepsy; grid = subdural grid electrode; HH = head of hippocampus; HP = hippocampus; L = left; MF = mesial frontal; PFC = pre-frontal cortex; TH = tail of hippocampus; TLE = temporal lobe epilepsy; R = right*

Patient	Age, Gender	Seizure semiology	MRI	Diagnosis	Electrodes	SOZ channels	Surgery	Engel outcome	Postoperative follow-up (months)
1	44, M	Complex Partial Seizures	Cortical Dysplasia in PFC	FLE	1 grid 8x8 2 grid 2x8 1 grid 4x8	RMF28-32	R Frontal Lobectomy	IA	47
2	50, M	Complex Partial Seizures	White matter changes in frontal lobe; possible right mesial temporal atrophy	TLE	5 depth 1x16	RAH5-8	R Temporal Lobectomy	IB	22
3	46, M	Complex Partial Seizures	Right Hippocampal Sclerosis	TLE	5 depth 1x16	RAH4	R Temporal Lobectomy	IA	41
4	34, M	Complex Partial Seizures	Bilateral hippocampal abnormalities with mild left-sided atrophy	TLE	2 depth (1x14) 3 depth (1x16) 2 grid (2x6) 1 depth (1x10)	RAM1-2, RHP1-3	R Temporal Lobectomy	IA	2
5	57, F	Complex Partial Seizures	None	TLE	11 depth (1x10) 1 depth (1x12)	LHH1-3, LTH2-3	L Lateral Temporal Lobectomy, L Amygdala and Hippocampal Resection	IA	23
6	53, F	Complex Partial Seizures	Left Mesial Temporal Sclerosis	TLE	10 depth (1x10)	LTH1	L Temporal Lobectomy	IA	32
7	54, F	Complex Partial Seizures	None	TLE	10 depth, (1x10)	RHH1-4, RAM1-2	R Temporal Lobectomy	IB	26

The remaining thirteen patients included in this study were obtained from the freely available online database associated with Fedele et al., 2017 at iEEG.org (http://crcns.org/data-

sets/methods/ieeg-1/about-ieeg-1). These patients underwent invasive EEG recordings with subdural and/or depth electrodes from March 2012 to April 2016 at the University Hospital Zurich as part of their presurgical evaluation. We refer to these thirteen patients as the ETH Zurich dataset. We included all patients that reported good clinical outcomes (class 1) at least 1 year following resective surgery based on the International League Against Epilepsy (ILAE) scale. Information regarding electrode types, data acquisition, and sleep scoring can be found in Fedele et al., 2017. Patient demographics for the ETH Zurich dataset are in Table 2.

Table 2. *Patient demographics of the ETH Zurich dataset. Abbreviations: depth = depth electrode; ETE = extratemporal epilepsy; ILAE = International League Against Epilepsy; Les = lesionectomy; sAHE = selective amygdala hippocampectomy; strip = strip electrode; TLE = temporal lobe epilepsy.*

Patient	Age, Gender	Epilepsy	Type of electrodes	Number of channels in resected tissue	Number of channels in non-resected tissue	Surgery	ILAE outcome	Postoperative follow-up (months)
8	25, M	TLE	5 depth 1 strip 4 x 1 1 strip 6 x 1	9	14	sAHE; Les	I	12
9	33, M	TLE	8 depth	12	12	sAHE; Les	I	29
10	20, F	TLE	5 depth	12	3	sAHE	I	13
11	20, F	TLE	8 depth	12	12	sAHE	I	41
12	40, M	TLE	8 depth	12	12	sAHE	I	14
13	48, M	TLE	8 depth	12	12	sAHE	I	11
14	37, M	ETE	1 grid 8 x 4 2 strips 4 x 1	3	31	Les	I	36
15	36, M	ETE	1 grid 8 x 8 1 depth	3	62	Les	I	37
16	49, M	ETE	1 grid 8 x 4 1 depth	21	16	Les	I	25
17	17, M	ETE	1 grid 8 x 8 1 depth	3	50	Les	I	25
18	46, F	ETE	2 grids 8 x 2 1 strip 6 x 1 1 strip 4 x 1 1 depth	1	29	Les	I	10

19	31, F	ETE	1 grid 8 x 4 2 strips 4 x 1	4	23	Les	I	25
20	17, F	ETE	1 grid 8 x 4 1 depth	3	34	Les	I	19

2.2. Data acquisition

Intracranial EEG was recorded for each patient using a combination of subdural electrocorticogram (ECoG) grids and strips, as well as depth electrodes. Recordings from the UCI dataset were collected using a Nihon Kohden JE-120A amplifier with a minimum sampling frequency of 2000 Hz for all patients. Recorded data were re-referenced to a bipolar montage for analysis. All bipolar re-referenced channel pairs that included an SOZ channel were deemed as SOZ (for example, channels RAH3-4 and RAH4-5 of UCI patient 3 were considered SOZ). Five 3-minute epochs from one night of iEEG recording were randomly selected for each UCI patient, for a total of 15 minutes per patient. Each epoch was chosen from data recorded between 8pm and 8am to improve the likelihood of analyzing sleep data, as HFO rates are increased and the occurrence of muscle artifacts is reduced during slow-wave sleep compared to wakefulness (Bagshaw et al., 2009; Clemens et al., 2003; Staba et al., 2004). We confirmed that all epochs were interictal data, recorded at least 1 hour away from seizures to reduce the influence of seizures on HFOs (Pearce et al., 2013). We analyzed all implanted electrodes, which ranged from 80 to 128 contacts for each UCI patient.

All interictal recordings from the ETH Zurich dataset were obtained at least 3 hours away from seizure activity, and we analyzed five 5-minute epochs of slow wave sleep from one night of recording. We included the same channels used for analysis by the original authors; the included and excluded channels can be found in the supplementary information provided by Fedele et al., 2017. Of note, the number of analyzed channels for each patient ranged from 15 to 53. For the analysis in the current study, we will define the “SOZ channels” to be the bipolar re-referenced channels within the resected area for each patient; the specific sites of seizure onset within the resected area were not provided. Defining the SOZ based on the resected volume will likely overestimate the true SOZ in the ETH Zurich dataset, while the method used

for the UCI dataset is likely an underestimate due to limited sampling of brain tissue using iEEG electrodes. These datasets are therefore complementary and represent the full range of possible SOZ definitions.

2.3. Data processing

2.3.1 Automatic HFO detection

Automated HFO detection was performed using an algorithm based on the root-mean-square (RMS) amplitude of bandpass filtered data (Staba et al., 2002). We will refer to this as the RMS detector. We used the RMS detector because it is referenced often as a benchmark for new algorithms (Charupanit et al., 2020; Gardner et al., 2007; Zelmann et al., 2012), serves as the core of many other detectors (Blanco et al., 2010; Burnos et al., 2014; Chaibi et al., 2013; Gliske et al., 2016; M. Wu et al., 2018; Zelmann et al., 2010), and is algorithmically simple to implement and examine. In this algorithm, broadband data is bandpass filtered in the 100-500Hz frequency range, and candidate events are identified when the RMS of the bandpass filtered signal exceeds a threshold for a minimum duration (which we will abbreviate as "min_dur;" see Figure 1A). The RMS signal is calculated using a moving window ("RMS_win"), and the first threshold ("nSD1") is defined as five standard deviations above the mean RMS signal. Consecutive events separated by less than a predefined gap time ("gap") are combined into one event. Candidate events are retained when a minimum number of peaks ("min_pk") in the rectified filtered data exceed a second threshold ("nSD2") defined as 3 standard deviations above the mean rectified data.

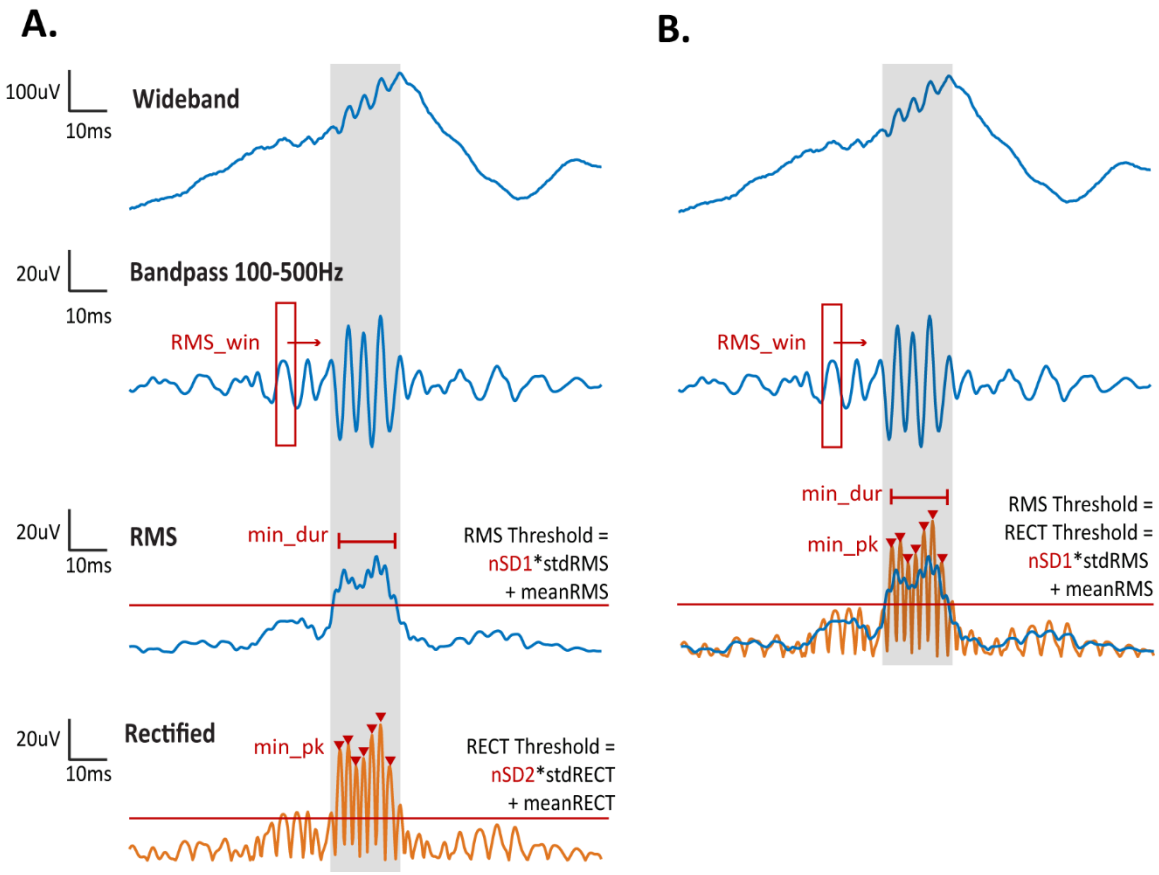

Figure 1. *Automatic HFO detection algorithm. (A) Full RMS detector and (B) Reduced RMS detector used in our analysis. The shaded gray region represents the window containing the identified candidate event. Detection thresholds are indicated by horizontal red lines, and each peak in the candidate event is marked by a red triangle. Abbreviations: RMS = root-mean-square, std = standard deviation*

To reduce algorithmic complexity and minimize the number of parameters to be optimized, we simplified the RMS detector algorithm by modifying the threshold applied to the rectified data (Figure 1B). We set threshold two equal to threshold one (nSD2 = nSD1), as they both ensure that the signal's energy exceeds a threshold determined from baseline activity. As in the original algorithm, we required that a minimum number of peaks in the rectified filtered data exceed this threshold, to promote rejection of fast transients. All other steps in the original algorithm were maintained.

The detection algorithm used in our analysis thus contains five parameters that must be considered during optimization: RMS window size (RMS_win), minimum event duration (min_dur), number of standard deviations above the mean RMS signal (nSD1), minimum gap time (gap), and minimum number of peaks (min_pk). From these parameters, we varied the three that directly affect initial detection of candidate events: RMS window size, minimum event duration, and threshold. Because the minimum gap time is a post-processing step to join candidate events and the minimum number of peaks has little impact as a criterion due to the redundancy with minimum event duration, we kept these variables constant at their default values. The default values of each parameter and values tested during the optimization procedure are described in Table 3.

***Table 3**. Detection parameters varied during optimization and corresponding ranges of values. Default values are underlined. The default value for minimum gap time is 10ms and minimum number of oscillations is 6.*

Parameter	Values
Threshold (nSD1)	1, 2, 3, 4, <u>5</u>, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
RMS window size (RMS_win) (ms)	2, <u>3</u>, 5, 7, 9, 11, 15, 17, 20
Minimum event duration (min_dur) (ms)	<u>6</u>, 12

2.3.2 Automatic artifact rejection

Because the RMS detector is highly sensitive (Gardner et al., 2007), we implemented two artifact rejection methods to improve specificity: PopDet and BkgStabaDet. Both methods were introduced by Gliske et al., 2016 for the purpose of creating a generalized HFO detection algorithm for long-term intracranial EEG recordings, such that the algorithm automatically identifies quality HFOs without any patient-specific tuning or operator intervention (Gliske et al., 2016). These artifact rejection steps were designed using the RMS detector for initial detection, making them appropriate for our analysis. In this study, the default parameters from Gliske et al., 2016 were used for all patients and all parameter sets.

The popDet criterion was designed to detect DC shifts and fast transients commonly found in EEG data (Figure 2A, B). When a DC shift or fast transient is filtered, it can have the appearance of an HFO in the 100-500 Hz frequency band (Bénar et al., 2010; Zelmann et al., 2010). However, these transients also contain power at very high frequencies, whereas a true HFO should have band-limited power. Therefore, the popDet identifies instances when the line length of a 0.1s window in the 850-990 Hz frequency range exceeds a threshold of 5 standard deviations above the mean line length calculated from baseline. Baseline is defined as a 5-second window preceding the window being evaluated.

The BkgStabaDet was designed to detect spatially diffuse HFOs, which are considered false positive detections because they contradict the idea that HFOs should be focal events (Bragin et al., 2002, 2011). If an HFO occurs in all channels of a depth electrode or grid, it will appear in the common average (Figure 2C). Therefore, detections in a single channel that occurred within 100ms of a detection in the common average were excluded from further analysis. As in the Gliske et al., 2016 implementation, we applied the RMS detector from Staba et al., 2002 to the common average within each depth electrode or grid using the default parameters..

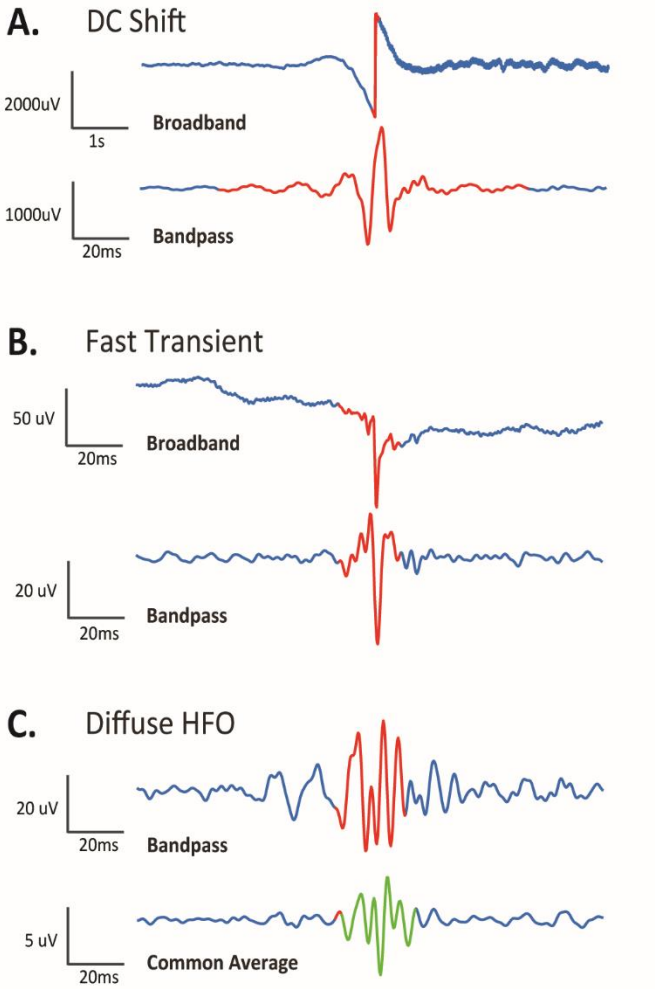

Figure 2. *Examples of rejected artifacts. (A) DC Shift in the broadband data (top) detected by popDet in the bandpass filtered data (bottom), and (B) fast transient in the broadband data (top) detected by popDet in bandpass filtered data (bottom), with the artifacts highlighted in red. (C) Diffuse HFO in the bandpass filtered data (top) detected by BkgStabaDet in the common average of the bandpass filtered data (bottom). The artifact (highlighted in green) represents the common average HFO that occurred at or near (±100ms) the same time as the detected HFO (highlighted in red) in the bandpass filtered signal.*

2.4 Data analysis

2.4.1 SOZ localization defined by rate thresholding

After HFO detection and artifact rejection were performed for a given parameter set, the HFO rate (number per minute) for each channel was averaged across epochs. A threshold was then

applied to the average HFO rates; channels with average HFO rates exceeding the rate threshold were classified as SOZ and the remaining channels were classified as non-SOZ (nSOZ).

2.4.2 Receiver operating characteristic curves and precision-recall curves

For each parameter set, we used a receiver operating characteristic (ROC) curve to characterize the ability of automatic HFO detection to localize the clinically-determined SOZ. For channels defined clinically as SOZ (based on visual interpretation of ictal iEEG), those with HFO rates that exceeded the threshold were marked as true positives (TP), and those that did not exceed the rate threshold were marked as false negatives (FN). For channels defined clinically as nSOZ, those that had HFO rates below the threshold were marked as true negatives (TN), and those with HFO rates above the threshold were marked as false positives (FP). The ROC curve was plotted by varying the HFO rate threshold and determining the true positive rate (TPR) and false positive rate (FPR) for each value:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \qquad (1)$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \qquad (2)$$

The area under the ROC curve (which we will abbreviate AUC) and the optimum cutoff (defined as the minimum distance between the ROC curve and the top left corner) were determined for each parameter set for all patients.

As a complementary measure, Precision-Recall (PR) scores were evaluated to examine the positive predictive value for each parameter set. Precision-Recall analyses are preferred when prediction power of imbalanced classes is being evaluated; in our case, we have imbalanced classes represented by the small subset of SOZ channels compared to the larger class of nSOZ channels for UCI patients, and the small subset of nSOZ channels compared to the larger class of SOZ channels for some ETH Zurich patients (specifically, patients 10 and 16). Precision (P), Recall (R), and the F1 score (F1) were computed as follows:

$$P = \frac{TP}{TP+FP} \quad (3)$$

$$R = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = 2\,\frac{P*R}{P+R} \quad (5)$$

Like the ROC curve, the PR curve was constructed by varying the rate threshold and calculating the precision, recall, and F1 score for each value. The optimal F1 score was defined as its maximum value, which represents the maximum harmonic mean of precision and recall. After calculating the ROC and PR results for each parameter set, the maximum AUC and optimal cutoff of the ROC curve, as well as the maximum AUC and optimal F1 score of the PR curve were determined for each patient.

2.4.3 Artifact counting

To assess the impact of the automated artifact rejection steps, we calculated the percentage of candidate HFOs rejected for each parameter set. The percentage of rejected candidates by popDet was calculated as the total number of candidates marked by popDet over the total number of candidates counted across all epochs. The same procedure was applied to candidates marked by BkgStabaDet. Because a candidate can be marked by both popDet and BkgStabaDet, the total percentage of rejected candidates by popDet and BkgStabaDet can exceed 100.

3. Results

3.1 Minimum event duration does not significantly impact SOZ localization

SOZ localization results were similar when HFO detection was done using minimum event duration (min_dur) values of 6ms or 12ms. We noted this pattern in all patients; one representative example is shown in Figure 3, and additional examples can be found in Supplementary Figure S1. Comparing the heatmaps in Figures 3A and 3B, the range of optimal F1 scores remained the same across the parameter space, with the highest optimal F1 score

achieved using a low threshold and short RMS window. A similar pattern of optimal parameters occurred in the AUC values. Because doubling the minimum event duration did not significantly alter localization results for any patient, the remaining results are reported using only a minimum event duration of 6ms, which was the value used in the original publication.

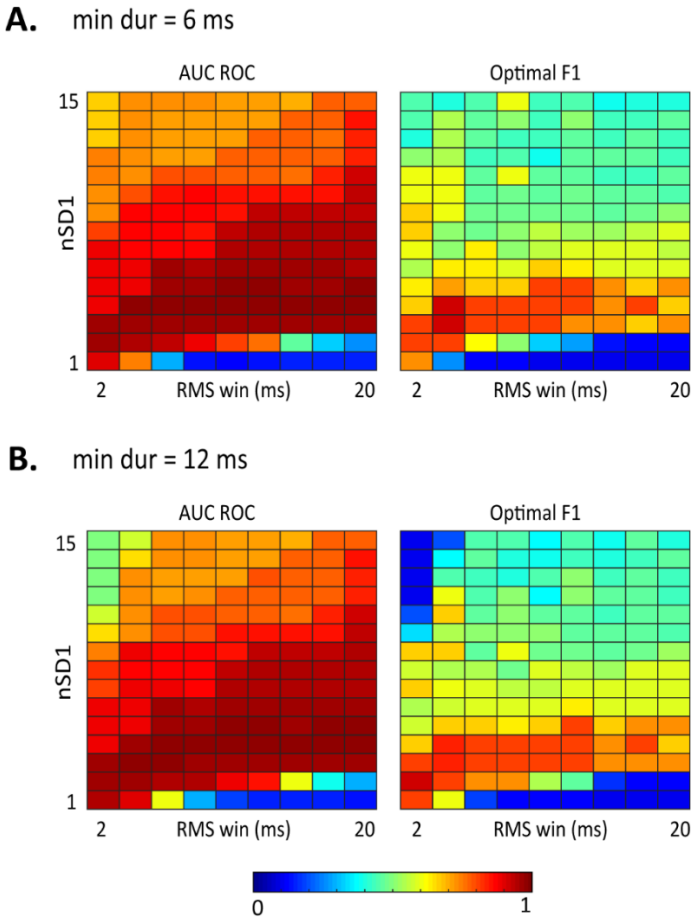

Figure 3. *Minimum event duration does not affect SOZ localization performance. Representative examples of SOZ localization performance across the parameter space, comparing results calculated using minimum event duration of (A) min_dur=6ms and (B) min_dur=12ms. Heatmaps of AUC of the ROC curve and optimal F1 score are shown across the parameter space for patient 5, with RMS window length (RMS_win) varying on the horizontal axis and RMS threshold (nSD1) varying on the vertical axis. All values range from 0 to 1, with results closer to 1 indicating good classification of SOZ and nSOZ channels.*

3.2 Artifact rejection does not significantly impact SOZ localization

The addition of artifact rejection steps did not change the detection parameters that produced the best classification of SOZ and nSOZ channels. A representative example of localization performance before artifact rejection is shown in Figure 4A. For this subject, the best performance, denoted by the highest optimal F1 score, occurred when the threshold was low (around nSD1 = 3-5) and across most RMS window sizes. Similar results were achieved after artifact rejection (Figure 4B). The optimal parameter set was the same in both cases (RMS_win = 3ms, nSD1 = 4), but we note that artifact rejection improved performance for higher values of nSD1 for this patient. Across all patients, the optimal parameters were preserved after artifact rejection, with some improvement for other parameter sets.

The percent of rejected candidates varied across the parameter space (Figure 4C, D). Near the optimal parameter set, the percent of rejected candidates by popDet ranged from 5.2 – 8% and by BkgStabaDet ranged from 6.6 – 9% for this representative example. Generally, higher percentages of detected events were rejected as artifacts when the RMS window was short and the threshold was high (top left corner of Figure 4C, D). In contrast, the smallest percentages of rejected candidates occurred when the RMS window was long and the threshold was low (bottom right). This pattern was similar across all patients and for both artifact rejection methods. The total percentage of rejected candidates was also generally consistent across patients.

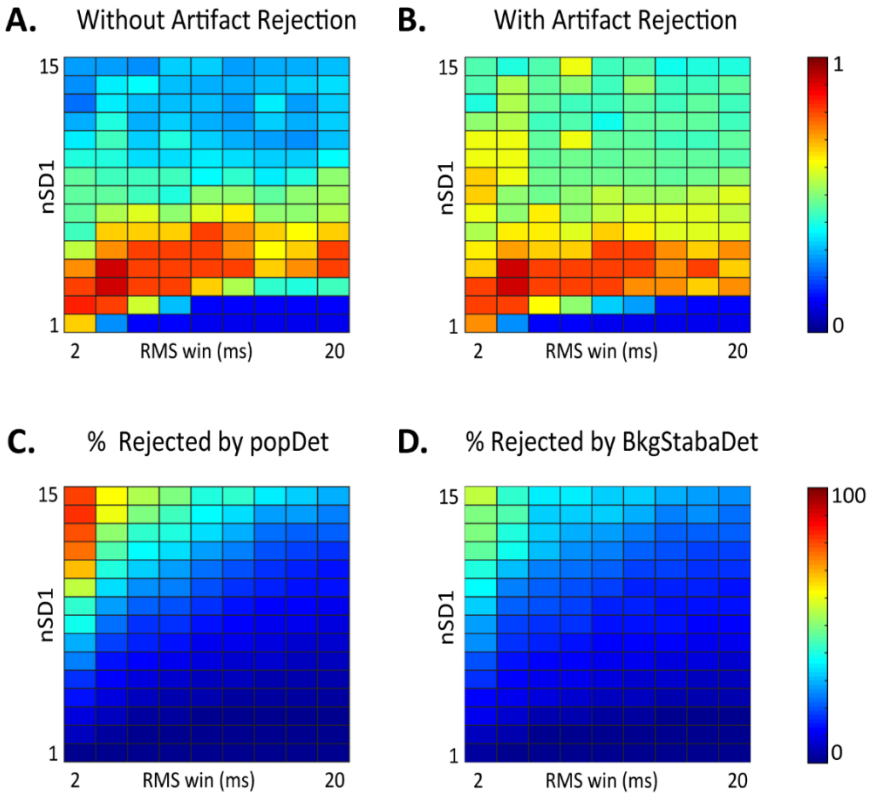

Figure 4. *Artifact rejection does not impact the optimal detection parameters. Comparison of optimal F1 scores (A) without and (B) with artifact rejection across the parameter space. Percentage of candidate events rejected by (C) popDet and (D) BkgStabaDet across the parameter space. Results from patient 5 are shown as a representative example.*

3.3 Optimal parameters are patient-specific and vary across a wide range

The optimal HFO detection parameters, which provided the most accurate classification of SOZ and nSOZ channels, varied across patients. Figure 5 displays the SOZ localization results for all twenty patients reported as optimal F1 score of the PR curve (left) and AUC of the ROC curve (middle). Similar results were obtained for the maximum AUC of the PR curve and optimal cutoff of the ROC curve (Supplementary Figure S2). A consistent range of generally good performance was found for all patients as both the RMS window and threshold increased (indicated by the red bands stretching from the lower left corner to the upper right corner). However, within this range, patients had their own subrange of optimal parameters. Four of the

twenty patients (patients 4, 5, 7, and 17) displayed a pattern with optimal values at or near the default detection parameters. These same patients reported high optimal F1 scores and AUC values, indicating a clear distinction between channels classified as SOZ and nSOZ based on HFO rate. Based on the optimal F1 score results, four other patients (patients 2, 3, 8, and 19) had optimal performance using a higher threshold and larger RMS window than the default parameters. However, patient 8 obtained similar localization accuracy at the default parameters and the optimal parameters. In six patients (patients 4, 7, 15, 16, 17, 18), the optimal parameters were spread over a wide range that included various RMS window lengths, indicating little influence of the RMS window parameter on detection. Five patients (patients 1, 6, 14, 19, and 20) had low optimal F1 scores across the parameter space, suggesting a substantial presence of false positives. However, three of these patients (patients 6, 14, and 20) displayed favorable results based on the ROC curve, indicating that the number of true positives was also high.

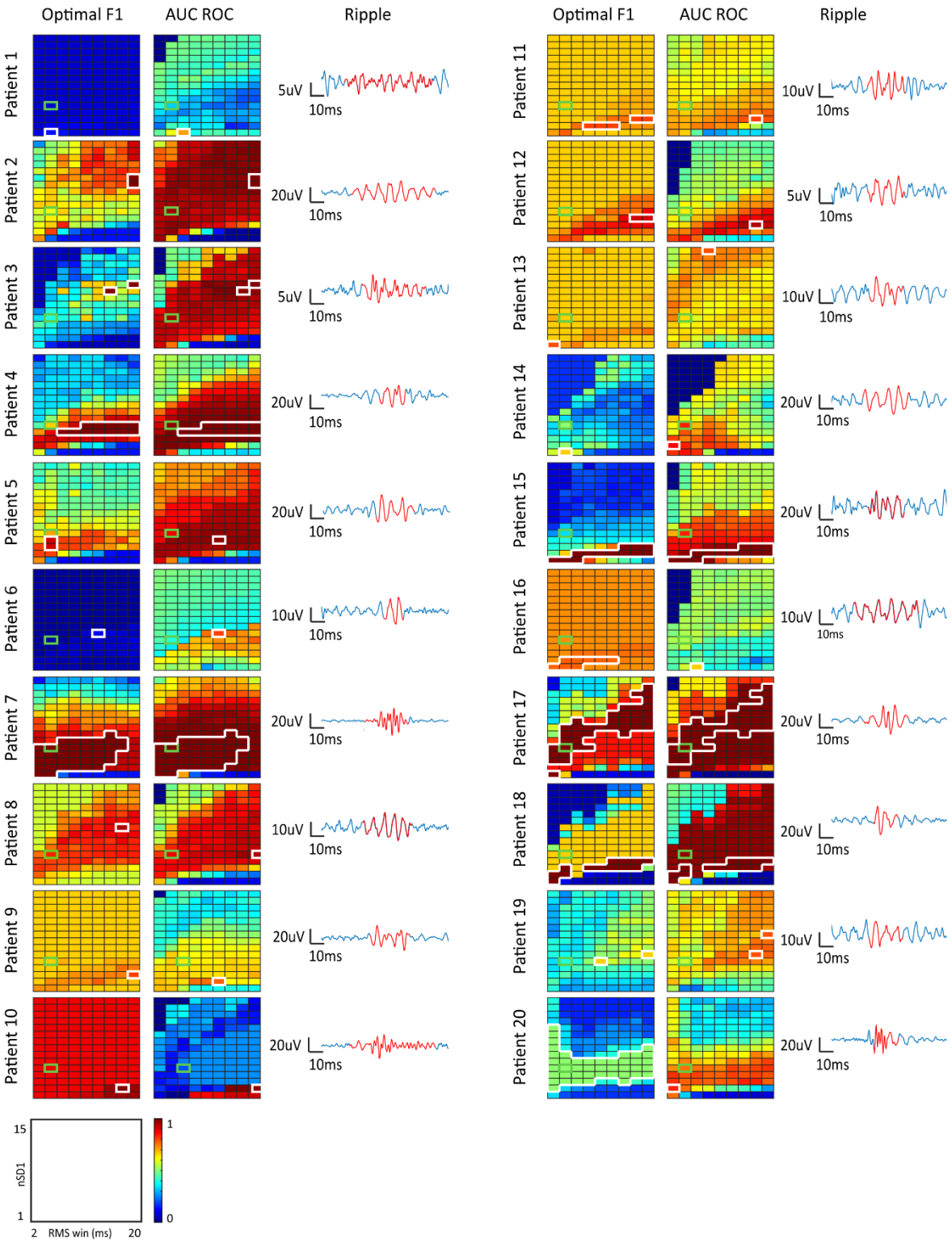

Figure 5. *Optimal HFO detection parameters are patient-specific. Optimal F1 scores (left column) and AUC values of the ROC curve (middle column) across the parameter space are shown for all patients. Results based on the default detection parameters are outlined in green.*

Parameter sets achieving optimal SOZ/nSOZ classification are outlined in white. Representative ripples detected using optimal parameters are shown in the right column.

3.4 Class imbalance affects interpretation of PR and ROC results

For all seven patients from the UCI dataset (patients 1-7), the ROC results provided a greater range of optimal detection parameters than the PR results. This can be explained by the definitions of precision and false positive rate — two distinct components of the PR and ROC curve. Precision is interpreted as a positive predictive value determined by the number of true positives relative to all predicted positives, as defined in equation 3. In patients from the UCI dataset, the number of nSOZ channels considerably outnumbers the SOZ channels; this increases the probability of marking a false positive, which increases the denominator of equation 3. Furthermore, since there are fewer SOZ channels (and therefore fewer true positives), the probability of obtaining a significant positive predictive value is low. In contrast, the false positive rate in ROC analysis is interpreted as the proportion of nSOZ channels with HFO rates above the threshold. Due to the large number of nSOZ channels (and thereby the high likelihood of having true negative channels based on HFO rate), the false positive rate can generally be kept low while the true positive rate increases, resulting in more favorable ROC results within the UCI dataset.

For example, in patient 6, good classification performance was only visible in the AUC of the ROC curve (Figure 5). In this case, we observe that the HFO rate distribution contains a considerable number of nSOZ channels with high HFO rates (Figure 6A). Twenty-four nSOZ channels have higher HFO rates than the SOZ channel, thereby giving a poor PR curve (Figure 6B). However, we see a better ROC result due to the large proportion of nSOZ channels with lower rates compared the SOZ channel (Figure 6C). Therefore, nSOZ channels with high HFO rates significantly affect the positive predictive value in PR, but they have a modest effect on ROC analysis when the proportion of nSOZ channels outweighs the SOZ channels.

In contrast to the UCI dataset, most patients from the ETH Zurich dataset (patients 8-20) had wider ranges of optimal parameters based on the PR results (Figure 5, left column) compared to the ROC results (Figure 5, right column). This relationship is inverted compared to the UCI dataset due to the smaller sampling of analyzed electrodes and broader interpretation of the SOZ, which was defined by the resected volume. Therefore, the proportion of SOZ channels outweighs the nSOZ channels, improving the likelihood of obtaining favorable positive predictive performance. Furthermore, due to the decreased number of electrodes, we see less variation in results across the parameter space for some ETH Zurich patients (patients 9, 10, 11, 12, 13, and 16).

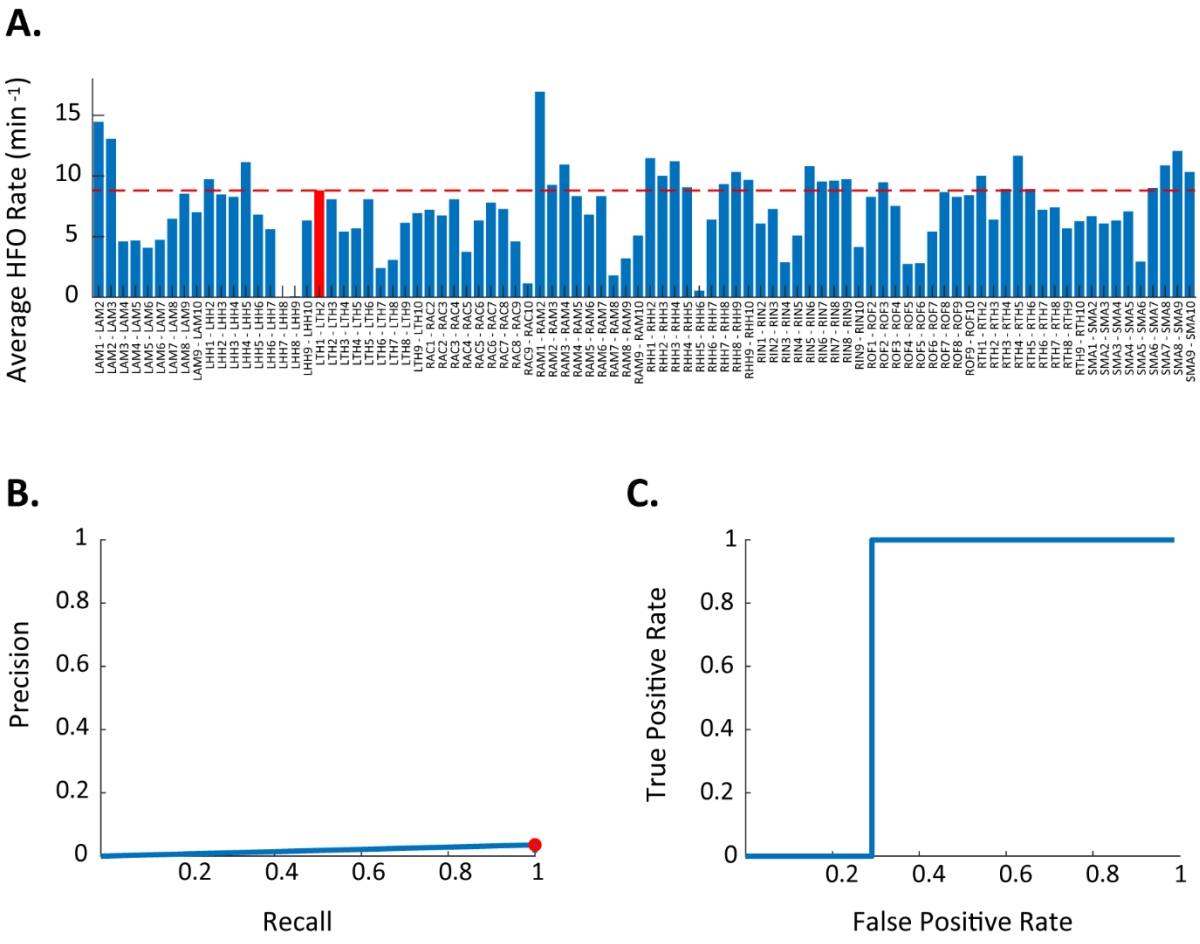

Figure 6. *Imbalanced proportion of SOZ channels and nSOZ channels alters interpretations of precision-recall and receiver-operator-characteristic results. (A) Average HFO rate for each channel in Patient 6 using an RMS threshold of three standard deviations (nSD1=3), RMS window size of 20ms (rms_win=20ms), and minimum event duration of 6ms (min_dur=6ms). (B)*

The resulting PR curve with the optimal F1 point circled in red, and (C) the ROC curve. The low number of SOZ channels (relative to the large number of nSOZ channels) negatively impacts the PR curve when the number of false positives is high, but it minimally impacts the ROC curve.

3.5 Optimizing HFO detection parameters improves SOZ localization

For each patient, we determined the optimal parameter sets based on the highest optimal F1 score and AUC of the ROC curve, then compared the results obtained from these parameters to the default detection parameters (Figure 7). In total, seventeen of the twenty patients (85%) showed improved localization results using optimized detection parameters, while the remaining three maintained the same results. We found substantial increases in optimal F1 scores (>0.2) for eight of the twenty patients (patients 2, 3, 4, 5, 12, 15, 18, 19; Figure 7A). Two patients (patients 7 and 17) showed no change in optimal F1 score or AUC because the optimal parameters matched the default parameters with good localization accuracy. Patient 20 also showed no change in the optimal F1 score due to consistent optimal and default parameters, however showed improvement in AUC. Six patients (patients 8, 9, 10, 11, 13, and 16) had reasonable localization results (optimal F1 score > 0.6) using default parameters, however still showed improvement in optimal F1 using optimal parameters. The remaining three patients (patients 1, 6 and 14) showed marginal growth in optimal F1 scores with optimization, but the value did not exceed 0.65. Note that improvements in ROC results carry more weight than improvements in optimal F1 scores because the optimal F1 score represents the best localization possible (at a single rate threshold), while the AUC represents the overall localization accuracy across a range of rate thresholds.

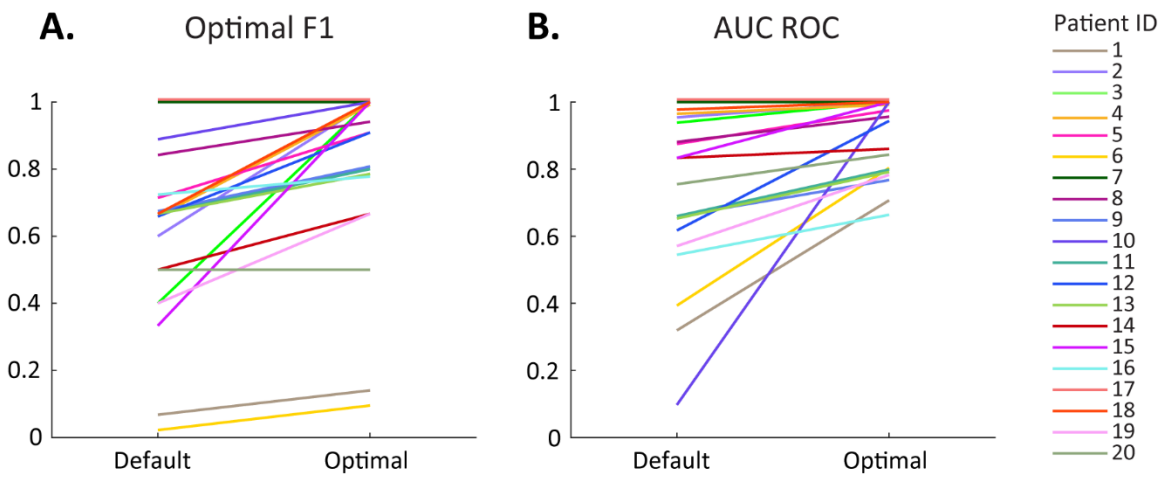

Figure 7. *The use of optimal HFO detection parameters improves precision, sensitivity, and specificity. Results for default values vs. optimal values are shown for (A) Optimal F1 scores and (B) AUC of the ROC curve for all patients.*

We assessed the number of observed true positive and false positive channels in each patient to evaluate the clinical relevance of the difference in classification accuracy using optimal versus default detection parameters (Figure 8). The optimal parameter set and optimal rate thresholds were determined based on the highest optimal F1 score. In four patients (patients 2, 4, 9, and 19), the number of true positives increased using optimal detection parameters compared to default (Figure 8A). For three of the four patients (patients 2, 4, and 19), the difference in marked true positives with optimal detection parameters was at least 20% of all SOZ channels, indicating a substantial improvement in SOZ localization. Only two patients (patients 12 and 13), both from the ETH Zurich dataset, had a greater number of true positives with default parameters, however the difference was only one channel. In sixteen of the twenty patients, the number of true negatives increased with optimal detection parameters, with differences ranging from +1 to +70 (Figure 8B). This suggests that optimization of HFO detection can substantially decrease the likelihood of misidentifying nSOZ channels as SOZ channels based on HFO rate. Overall, seventeen of the twenty patients showed improved SOZ

localization with optimal parameters, while the remaining three showed results consistent with default parameters.

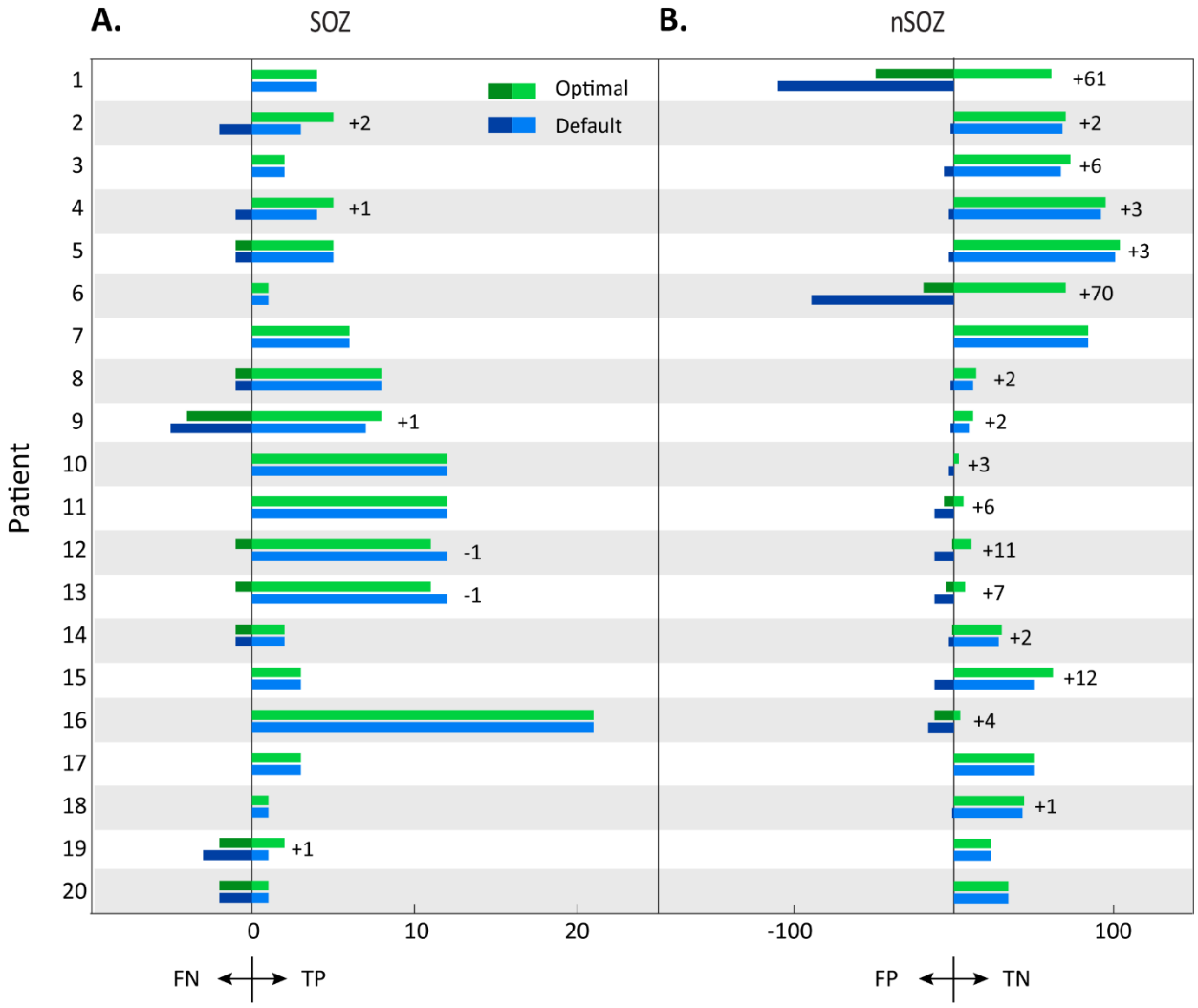

***Figure 8**. The use of optimal HFO detection parameters increases the number of observed true positives in the SOZ and true negatives in the nSOZ. The number of channels for default values (blue) and optimal values (green) are shown for all patients. (A) The number of SOZ channels correctly marked as true positives (TP) are displayed as positive values, while the number of SOZ channels marked as false negatives (FN) are depicted as negative values. (B) The number of nSOZ channels correctly marked as true negatives (TN) are shown as positive values, while the nSOZ channels marked as false positives (FP) are negative. The difference in correctly classified channels using default and optimal values are listed to the right of the bar graph for each patient.*

3.6 Optimal parameters vary widely across the parameter space

For many patients, the highest classification accuracy was obtained with more than one parameter set. For each patient, we plotted all optimal parameter sets across the ranges of tested values (Figure 9). The optimal parameters derived from the optimal F1 score (Figure 8A) and AUC of the ROC curve (Figure 8B) spanned the parameter space. Most patients had the same or similar optimal parameters determined from both measures; deviants from this pattern included patients 5, 8, 13, and 20. In some patients, the parameter sets spanned the entire range of RMS window sizes, but with different values of the amplitude threshold (nSD1).

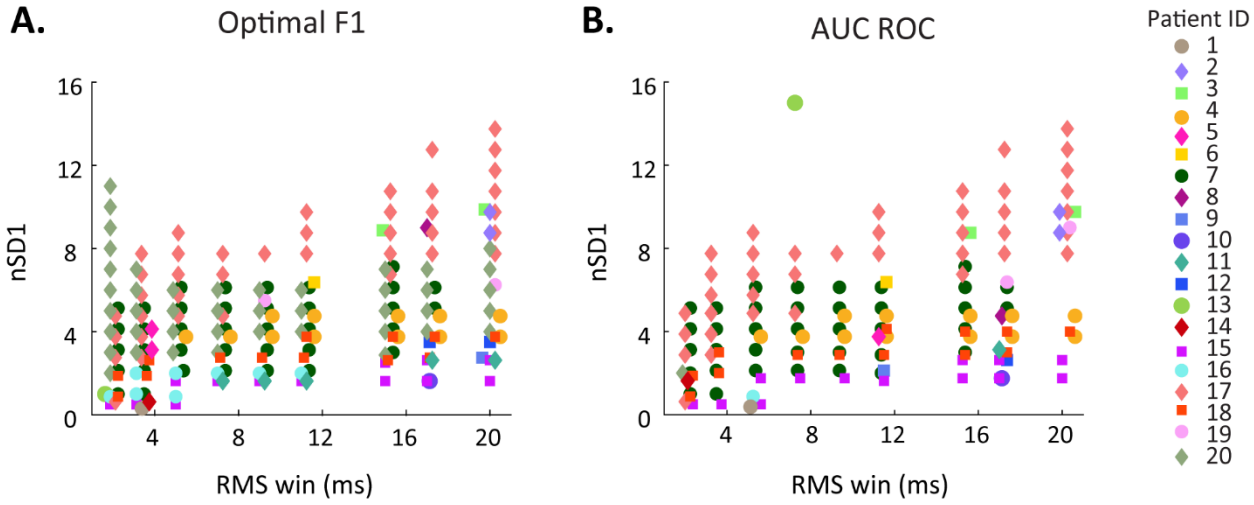

Figure 9. *Optimal HFO detection parameters based on (A) optimal F1 score and (B) AUC of the ROC curve for all patients.*

4. Discussion

4.1 Summary of findings

This study assessed SOZ localization accuracy over a wide range of HFO detection parameters to measure the impact of parameter optimization in two patient cohorts. In both patient populations, we found that the range of parameters for optimal classification of SOZ/nSOZ electrodes was patient-specific and spanned a wider range of values than those used in most

prior studies. The parameters with the biggest effect on SOZ localization accuracy were the RMS window length and the amplitude thresholds; in contrast, artifact rejection had minimal impact. Additionally, detection using optimal parameters led to improved localization results compared to the standard parameters established in the original publication. This suggests that proper selection of automatic detection parameters can increase the power of HFOs as a biomarker of the SOZ, and future efforts should be directed to development of patient-specific optimization procedures.

4.2 Influence of class imbalance on interpretation of results

In this study, the class imbalance (number of SOZ channels relative to nSOZ channels) affected the interpretation of the results for each patient. The number, type, and placement of electrodes varied per patient based on clinical indication, resulting in inconsistent sample sizes. Furthermore, the localized SOZ was specific to each patient, leading to variable numbers of SOZ channels and nSOZ channels. Because of this, we characterized the classification performance using both the ROC and PR curves, as they are complementary measures. The ROC curve provides valuable information on the classification sensitivity for each parameter set, and it has been widely used to characterize classification performance (K. Charupanit & Lopour, 2017; Roehri et al., 2018; Zelmann et al., 2012). In the UCI dataset, characterized by larger proportions of nSOZ channels for all patients, we saw smaller variations in AUC ROC results across the parameter space. This was because the small proportion of SOZ channels led to true positive rates that varied only marginally as the rate threshold was varied. In contrast, the Precision-Recall curves provided valuable information on the positive predictive value (PPV) of the classification, and this method is considered to be appropriate for analysis in imbalanced datasets (Saito & Rehmsmeier, 2015). Because the PPV is indicative of the proportion of true positives relative to the predicted positives, this metric varies more significantly than the true positive rate (which is used in ROC analysis). Furthermore, the PPV provides key information on the presence of false positives, which occur more frequently in the UCI dataset than false negatives due to the small proportion of SOZ channels. As a demonstration of this, we observed

more localized optimal parameter ranges for the optimal F1 scores than the AUC for all UCI patients in Figure 5.

The patterns described here for both ROC and Precision-Recall curves were also observed in the ETH Zurich dataset, albeit reversed because the proportion of SOZ channels was greater. Because we cannot compare classification accuracy between the UCI dataset and ETH Zurich dataset directly due to the variables described, we only conclude that the patient's optimal parameters differ, and we cannot make comparisons on the performance of HFO rate as a biomarker across the patient populations. In summary, the dissimilarities in electrode sampling led to different ROC and PR outcomes, and for some patients, localization was deemed poor based on these results.

4.3 Characteristics of resulting HFO detections

Detecting HFOs over a wide range of parameters led to substantial differences in the number of HFOs detected. For parameter sets with low thresholds and small RMS windows, we detected more than 1,000 HFOs per channel over 15 minutes of data on average for all patients. These detections were likely a mix of true HFOs and false positive detections; however, in some cases, the resulting distribution of average detection rates across channels yielded favorable localization results when a rate threshold was applied. In the opposite corner of the parameter space, parameter sets with high thresholds and large RMS windows contained very few detections. These events needed to have inordinately high amplitude to meet the threshold and minimum event duration criteria, especially because the amplitude was damped by the use of a large RMS window. We detected very few of these events, as low as 0-3 detections across all channels, which likely led to the unreliable localization results when using these extreme parameters. However, the best localization results for patient 2 occurred for such parameter sets (nSD1 = 9-10, RMS_win = 20 ms). This is consistent with previous findings that demonstrate amplitude as an important metric in HFO analysis (Charupanit et al., 2020; Malinowska et al., 2015; Matsumoto et al., 2013).

4.4 Influence of underlying pathologies

A subset of patients did not achieve accurate SOZ localizations for any parameter set, and the presence of secondary pathologies or specific epilepsy disorders is one possible explanation for this. In the UCI dataset, MRI results of patients 1, 2, 3, and 6 revealed lesional or tissue abnormalities consisting of hippocampal and temporal sclerosis, cortical dysplasia of the prefrontal cortex, and imaging abnormalities in the frontal cortex. The presence of these secondary pathologies may have influenced the HFO rates detected in these regions, due to various types of neuronal derangements in abnormal brain tissue (Ferrari-Marinho et al., 2015). Regarding the influence of epilepsy disorder on HFO rate in the SOZ, previous studies have found distinctions between patients with temporal and extratemporal epilepsy. For example, Guragain et al., 2018 and Haegelen et al., 2013 examined the spatial mapping of HFOs and found that patients with extratemporal epilepsy disorders did not contain significantly elevated HFOs rates in the SOZ compared to the nSOZ. In contrast, patients with mesial temporal lobe epilepsy showed significantly elevated HFO rates in the SOZ, suggesting that HFOs may be a specific marker for patients with mesial temporal lobe epilepsies (Bragin et al., 1999; Staba et al., 2004; Worrell et al., 2008). Similar distinctions were found between patients with neocortical epilepsies and mesial temporal lobe epilepsies, where patients with neocortical epilepsies did not exhibit HFOs in epileptic structures or healthy regions (Crépon et al., 2010). However, HFOs were recorded in the SOZ of all patients with mesial temporal lobe epilepsy (Crépon et al., 2010). Of the UCI patients included in this study, patient 1 was diagnosed with frontal lobe epilepsy, which may explain the suboptimal SOZ localization results across all parameter sets. Similarly, patients 14-20 from the ETH Zurich dataset were diagnosed with extratemporal epilepsy and underwent lesionectomy, which may have influenced localization because different lesion types have various levels of intrinsic epileptogenicity and may contribute dissimilarly to epileptogenic networks (Jacobs et al., 2009). In summary, underlying pathologies may play a role in the ability to accurately localize the SOZ based on findings within our cohort; however, additional patients should be studied to make robust conclusions.

4.5 Limitations and future work

There are several limitations to our work. Our study demonstrates that patient-specific optimization can significantly improve SOZ localization using HFOs; however, two important questions remain unanswered. First, how could such optimization be implemented in clinical practice? Our analysis was retrospective, and we reported the best possible localization results assuming that the optimal detection parameters were chosen; however, in clinical practice, optimization would need to be performed on a reserved subset of patient data and without prior knowledge of the SOZ. Without knowledge of the SOZ, the detector could be optimized to match visually detected HFOs, but this approach may not reliably correlate to optimal surgical outcome. Thus, future work should address methods for implementation of per-patient optimization in clinical settings. Second, is the improvement in SOZ localization due to HFO detection parameter optimization clinically meaningful? For example, if the AUC increases from 0.80 to 0.85 due to parameter optimization, would this change the patient's treatment or outcome? We have estimated the improvement due to optimization by quantifying the differences in observed true positive and false positive channels; however, the long-term impact on surgical outcome for each patient cannot be predicted. Future work should aim to address the question of what constitutes clinically meaningful improvement.

Another limitation was the relatively small cohort size. However, our results were consistent across two independent datasets, with each dataset demonstrating patient-specific variability in optimal parameters. This was the main goal of our study. Also, the ETH Zurich dataset delineated the SOZ based on resected volume, while the UCI dataset used ictal EEG analysis, and both datasets contained varying class imbalances that affected our interpretation of the SOZ localization results. These differences restricted our ability to make more definitive conclusions regarding SOZ localization. The resected volume is typically a larger region of the brain that encompasses the EEG-defined SOZ. Therefore, the SOZ was more loosely defined for the ETH Zurich dataset, which also had varying numbers of recording channels, and this likely affected ROC and PR results.

Regarding automatic HFO detection, we aimed to analyze artifact-free intracranial EEG data during sleep. However, concurrent scalp EEG was not available for the UCI dataset, and the data were therefore not sleep staged. Previous studies have shown that HFO activity varies across wakefulness and sleep and for specific sleep stages (Bagshaw et al., 2009; Clemens et al., 2003; Staba et al., 2004). Therefore, ensuring that all data were recorded during the same sleep stage may provide more accurate results. Another limitation in our study was the lack of optimization of our artifact rejection methods. Although Gliske et al., 2016 performed an optimization procedure to determine the rejection threshold of the popDet, one set of parameters does not necessarily work equally well for all patients, as we have shown here. Furthermore, Gliske et al., 2016 tested the threshold ranging from 5 to 15 standard deviations; however, in some of our patients, we found that the popDet performed better using a lower threshold. Ideally, to improve localization results, visual validation on a reserved subset of data should be used to optimize parameters for artifact rejection.

5. Conclusion

This study examined the impact of optimization of automatic HFO detection parameters on SOZ localization in medically refractory epilepsy patients. All patients achieved seizure freedom following resective surgery, and the majority of patients exhibited elevated HFO rates in the SOZ, which led to favorable localization results. However, only 15% of these patients achieved their maximum localization accuracy using conventional detection parameters, highlighting the significance of patient variability and the importance of tuning automatic HFO detectors. Future studies should propose patient-specific optimization techniques for the clinical setting and examine the characteristics of pathological HFOs that provide optimal classification of SOZ and nSOZ channels.

Conflict of Interest Statement

None of the authors have potential conflicts of interest to be disclosed.

Acknowledgements

The authors would like to acknowledge Dr. Krit Charupanit for technical assistance and the staff and epilepsy patients at the UCI Medical Center for their contributions to the study.

6. References

1. Akiyama, T., McCoy, B., Go, C. Y., Ochi, A., Elliott, I. M., Akiyama, M., Donner, E. J., Weiss, S. K., Snead, O. C., Rutka, J. T., Drake, J. M., & Otsubo, H. (2011). Focal resection of fast ripples on extraoperative intracranial EEG improves seizure outcome in pediatric epilepsy. *Epilepsia*, *52*(10), 1802–1811. https://doi.org/10.1111/j.1528-1167.2011.03199.x

2. Bagshaw, A. P., Jacobs, J., Levan, P., Dubeau, F., & Gotman, J. (2009). Effect of sleep stage on interictal high-frequency oscillations recorded from depth macroelectrodes in patients with focal epilepsy. *Epilepsia*, *50*(4), 617–628. https://doi.org/10.1111/j.1528-1167.2008.01784.x

3. Bautista, R. E. D., Cobbs, M. A., Spencer, D. D., Spencer, S. S., Spencer, D. D., & Spencer, S. S. (1999). *Prediction of Surgical Outcome by Interictal Epileptiform Abnormalities During Intracranial EEG Monitoring in Patients with Extrahippocampal Seizures* (Vol. 40, Issue 7).

4. Bénar, C. G., Chauvière, L., Bartolomei, F., & Wendling, F. (2010). Pitfalls of high-pass filtering for detecting epileptic oscillations: A technical note on “false” ripples. *Clinical Neurophysiology*, *121*(3), 301–310. https://doi.org/10.1016/j.clinph.2009.10.019

5. Blanco, J. A., Stead, M., Krieger, A., Viventi, J., Marsh, W. R., Lee, K. H., Worrell, G. A., & Litt, B. (2010). Unsupervised classification of high-frequency oscillations in human neocortical epilepsy and control patients. *Journal of Neurophysiology*, *104*(5), 2900–2912. https://doi.org/10.1152/jn.01082.2009

6. Bragin, A., Benassi, S. K., Kheiri, F., & Engel, J. (2011). Further evidence that pathologic high-frequency oscillations are bursts of population spikes derived from recordings of identified cells in dentate gyrus. *Epilepsia*, *52*(1), 45–52. https://doi.org/10.1111/j.1528-1167.2010.02896.x

7. Bragin, A., Engel, J., Wilson, C. L., Fried, I., & Buzsáki, G. (1999). High-Frequency

Oscillations in Human Brain. In *Hippocampus* (Vol. 9).

8. Bragin, A., Mody, I., Wilson, C. L., & Engel, J. (2002). Local generation of fast ripples in epileptic brain. *Journal of Neuroscience*, *22*(5), 2012–2021. https://doi.org/10.1523/jneurosci.22-05-02012.2002

9. Burnos, S., Hilfiker, P., Sürücü, O., Scholkmann, F., Krayenbühl, N., Grunwald, T., & Sarnthein, J. (2014). Human intracranial high frequency oscillations (HFOs) detected by automatic time-frequency analysis. *PLoS ONE*, *9*(4). https://doi.org/10.1371/journal.pone.0094381

10. Chaibi, S., Lajnef, T., Sakka, Z., Samet, M., & Kachouri, A. (2013). A Comparaison of Methods for Detection of High Frequency Oscillations (HFOs) in Human Intacerberal EEG Recordings. *American Journal of Signal Processing*, *2013*(2), 25–34. https://doi.org/10.5923/j.ajsp.20130302.02

11. Charupanit, K., & Lopour, B. A. (2017). A Simple Statistical Method for the Automatic Detection of Ripples in Human Intracranial EEG. *Brain Topography*, *30*(6), 724–738. https://doi.org/10.1007/s10548-017-0579-6

12. Charupanit, Krit, Sen-Gupta, I., Lin, J. J., & Lopour, B. A. (2019). Detection of anomalous high frequency events in human intracranial EEG. *BioRxiv*, 782912. https://doi.org/10.1101/782912

13. Charupanit, Krit, Sen-Gupta, I., Lin, J. J., & Lopour, B. A. (2020). Detection of anomalous high frequency events in human intracranial EEG. *Epilepsia Open*, epi4.12397. https://doi.org/10.1002/epi4.12397

14. Clemens, Z., Janszky, J., Szucs, A., Békésy, M., Clemens, B., & Halász, P. (2003). Interictal epileptic spiking during sleep and wakefulness in mesial temporal lobe epilepsy: A comparative study of scalp and foramen ovale electrodes. *Epilepsia*, *44*(2), 186–192. https://doi.org/10.1046/j.1528-1157.2003.27302.x

15. Crépon, B., Navarro, V., Hasboun, D., Clemenceau, S., Martinerie, J., Baulac, M., Adam, C., & Le Van Quyen, M. (2010). Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*, *133*(1), 33–45. https://doi.org/10.1093/brain/awp277

16. Dümpelmann, M., Jacobs, J., Kerber, K., & Schulze-Bonhage, A. (2012). Automatic 80-250Hz "ripple" high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network. *Clinical Neurophysiology*, *123*(9), 1721–1731. https://doi.org/10.1016/j.clinph.2012.02.072

17. Edelvik, A., Rydenhag, B., Olsson, I., Flink, R., Kumlien, E., Källén, K., & Malmgren, K. (2013). Long-term outcomes of epilepsy surgery in Sweden: A national prospective and longitudinal study. *Neurology*, *81*(14), 1244–1251. https://doi.org/10.1212/WNL.0b013e3182a6ca7b

18. Fedele, T., Burnos, S., Boran, E., Krayenbühl, N., Hilfiker, P., Grunwald, T., & Sarnthein, J. (2017). Resection of high frequency oscillations predicts seizure outcome in the individual patient. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-13064-1

19. Ferrari-Marinho, T., Perucca, P., Mok, K., Olivier, A., Hall, J., Dubeau, F., & Gotman, J. (2015). Pathologic substrates of focal epilepsy influence the generation of high-frequency oscillations. *Epilepsia*, *56*(4), 592–598. https://doi.org/10.1111/epi.12940

20. Gardner, A. B., Worrell, G. A., Marsh, E., Dlugos, D., & Litt, B. (2007). Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. *Clinical Neurophysiology*, *118*(5), 1134–1143. https://doi.org/10.1016/j.clinph.2006.12.019

21. Gliske, S. V., Irwin, Z. T., Davis, K. A., Sahaya, K., Chestek, C., & Stacey, W. C. (2016). Universal automated high frequency oscillation detector for real-time, long term EEG. *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2015.07.016

22. Guragain, H., Cimbalnik, J., Stead, M., Groppe, D. M., Berry, B. M., Kremen, V., Kenney-Jung, D., Britton, J., Worrell, G. A., & Brinkmann, B. H. (2018). Spatial variation in high-frequency oscillation rates and amplitudes in intracranial EEG. *Neurology*, *90*(8), E639–E646. https://doi.org/10.1212/WNL.0000000000004998

23. Haegelen, C., Perucca, P., Châtillon, C. E., Andrade-Valença, L., Zelmann, R., Jacobs, J., Collins, D. L., Dubeau, F., Olivier, A., & Gotman, J. (2013). High-frequency oscillations, extent of surgical resection, and surgical outcome in drug-resistant focal epilepsy.

Epilepsia, *54*(5), 848–857. https://doi.org/10.1111/epi.12075

24. Jacobs, J., Asano, E., Otsubo, H., Wu, J. Y., Zijlmans, M., Mohamed, I., Kahane, P., Dubeau, F., Navarro, V., & Gotman, J. (2012). High-frequency oscillations (HFOs) in clinical epilepsy. *Progress in Neurobiology*, *98*(3), 302–315. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

25. Jacobs, J., Golla, T., Mader, M., Schelter, B., Dümpelmann, M., Korinthenberg, R., & Schulze-Bonhage, A. (2014). Electrical stimulation for cortical mapping reduces the density of high frequency oscillations. *Epilepsy Research*, *108*(10), 1758–1769. https://doi.org/10.1016/j.eplepsyres.2014.09.022

26. Jacobs, J., LeVan, P., Chander, R., Hall, J., Dubeau, F., & Gotman, J. (2008). Interictal high-frequency oscillations (80-500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. *Epilepsia*, *49*(11), 1893–1907. https://doi.org/10.1111/j.1528-1167.2008.01656.x

27. Jacobs, J., Levan, P., Chtillon, C. D., Olivier, A., Dubeau, F., & Gotman, J. (2009). High frequency oscillations in intracranial EEGs mark epileptogenicity rather than lesion type. *Brain*, *132*(4), 1022–1037. https://doi.org/10.1093/brain/awn351

28. Jacobs, J., Wu, J. Y., Perucca, P., Zelmann, R., Mader, M., Dubeau, F., Mathern, G. W., Schulze-Bonhage, A., & Gotman, J. (2018). Removing high-frequency oscillations: A prospective multicenter study on seizure outcome. *Neurology*, *91*(11), e1040–e1052. https://doi.org/10.1212/WNL.0000000000006158

29. Jacobs, J., Zijlmans, M., Zelmann, R., Chatillon, C. É., Hall, J., Olivier, A., Dubeau, F., & Gotman, J. (2010). High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery. *Annals of Neurology*, *67*(2), 209–220. https://doi.org/10.1002/ana.21847

30. Kovac, S., Vakharia, V. N., Scott, C., & Diehl, B. (2017). Invasive epilepsy surgery evaluation. *Seizure*, *44*, 125–136. https://doi.org/10.1016/j.seizure.2016.10.016

31. Malinowska, U., Bergey, G. K., Harezlak, J., & Jouny, C. C. (2015). Identification of seizure onset zone and preictal state based on characteristics of high frequency oscillations. *Clinical Neurophysiology*, *126*(8), 1505–1513.

https://doi.org/10.1016/j.clinph.2014.11.007

32. Matsumoto, A., Brinkmann, B. H., Stead, S. M., Matsumoto, J., Kucewicz, M. T., Marsh, W. R., Meyer, F., & Worrell, G. (2013). Pathological and physiological high-frequency oscillations in focal human epilepsy. *Journal of Neurophysiology*. https://doi.org/10.1152/jn.00341.2013

33. Migliorelli, C., Alonso, J. F., Romero, S., Nowak, R., Russi, A., & Mañanas, M. A. (2017). Automated detection of epileptic ripples in MEG using beamformer-based virtual sensors. *Journal of Neural Engineering*, *14*(4). https://doi.org/10.1088/1741-2552/aa684c

34. Noachtar, S., & Borggraefe, I. (2009). Epilepsy surgery: A critical review. *Epilepsy and Behavior*, *15*(1), 66–72. https://doi.org/10.1016/j.yebeh.2009.02.028

35. Pearce, A., Wulsin, D., Blanco, J. A., Krieger, A., Litt, B., & Stacey, W. C. (2013). Temporal changes of neocortical high-frequency oscillations in epilepsy. *Journal of Neurophysiology*, *110*(5), 1167–1179. https://doi.org/10.1152/jn.01009.2012

36. Roehri, N., Pizzo, F., Lagarde, S., Lambert, I., Nica, A., McGonigal, A., Giusiano, B., Bartolomei, F., & Bénar, C. G. (2018). High-frequency oscillations are not better biomarkers of epileptogenic tissues than spikes. *Annals of Neurology*, *83*(1), 84–97. https://doi.org/10.1002/ana.25124

37. Ryvlin, P., Cross, J. H., & Rheims, S. (2014). Epilepsy surgery in children and adults. *The Lancet Neurology*, *13*(11), 1114–1126. https://doi.org/10.1016/S1474-4422(14)70156-5

38. Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, *10*(3), 1–21. https://doi.org/10.1371/journal.pone.0118432

39. Schuele, S. U., & Lüders, H. O. (2008). Intractable epilepsy: management and therapeutic alternatives. *The Lancet Neurology*, *7*(6), 514–524. https://doi.org/10.1016/S1474-4422(08)70108-X

40. Sindhu, K. R., Staba, R. J., & Lopour, B. A. (2020). Trends in the use of automated algorithms for the detection of high-frequency oscillations associated with human epilepsy. *Epilepsia*, *00*, 1–17. https://doi.org/10.1111/epi.16622

41. Spring, A. M., Pittman, D. J., Aghakhani, Y., Jirsch, J., Pillay, N., Bello-Espinosa, L. E., Josephson, C., & Federico, P. (2017). Interrater reliability of visually evaluated high frequency oscillations. *Clinical Neurophysiology*, *128*(3), 433–441. https://doi.org/10.1016/j.clinph.2016.12.017

42. Spring, A. M., Pittman, D. J., Aghakhani, Y., Jirsch, J., Pillay, N., Bello-Espinosa, L. E., Josephson, C., & Federico, P. (2018). Generalizability of High Frequency Oscillation Evaluations in the Ripple Band. *Frontiers in Neurology*, *9*(June), 1–11. https://doi.org/10.3389/fneur.2018.00510

43. Staba, R. J., & Bragin, A. (2011). High-frequency oscillations and other electrophysiological biomarkers of epilepsy: Underlying mechanisms. In *Biomarkers in Medicine* (Vol. 5, Issue 5, pp. 545–556). https://doi.org/10.2217/bmm.11.72

44. Staba, R. J., Wilson, C. L., Bragin, A., Fried, I., & Engel, J. (2002). *Quantitative Analysis of High-Frequency Oscillations (80-500 Hz) Recorded in Human Epileptic Hippocampus and Entorhinal Cortex*. https://doi.org/10.1152/jn.00322.2002

45. Staba, R. J., Wilson, C. L., Bragin, A., Jhung, D., Fried, I., & Engel, J. (2004). High-frequency oscillations recorded in human medial temporal lobe during sleep. *Annals of Neurology*, *56*(1), 108–115. https://doi.org/10.1002/ana.20164

46. Westmoreland, B. F. (1996). *Epileptiform Electroencephalographic Patterns*.

47. Worrell, G. A., Jerbi, K., Kobayashi, K., Lina, J. M., Zelmann, R., & Le Van Quyen, M. (2012). Recording and analysis techniques for high-frequency oscillations. In *Progress in Neurobiology* (Vol. 98, Issue 3, pp. 265–278). https://doi.org/10.1016/j.pneurobio.2012.02.006

48. Worrell, Greg A., Gardner, A. B., Stead, S. M., Hu, S., Goerss, S., Cascino, G. J., Meyer, F. B., Marsh, R., & Litt, B. (2008). High-frequency oscillations in human temporal lobe: Simultaneous microwire and clinical macroelectrode recordings. *Brain*, *131*(4), 928–937. https://doi.org/10.1093/brain/awn006

49. Wu, J. Y., Sankar, R., Lerner, J. T., Matsumoto, J. H., Vinters, H. V., & Mathern, G. W. (2010). Removing interictal fast ripples on electrocorticography linked with seizure freedom in children. *Neurology*, *75*(19), 1686–1694.

https://doi.org/10.1212/WNL.0b013e3181fc27d0

50. Wu, M., Wan, T., Ding, M., Wan, X., Du, Y., & She, J. (2018). A New Unsupervised Detector of High-Frequency Oscillations in Accurate Localization of Epileptic Seizure Onset Zones. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(12), 2280–2289. https://doi.org/10.1109/TNSRE.2018.2877820

51. Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Chander, R., & Gotman, J. (2010). Automatic detector of High Frequency Oscillations for human recordings with macroelectrodes. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10*. https://doi.org/10.1109/IEMBS.2010.5627464

52. Zelmann, R., Mari, F., Jacobs, J., Zijlmans, M., Dubeau, F., & Gotman, J. (2012). A comparison between detectors of high frequency oscillations. *Clinical Neurophysiology*, *123*(1), 106–116. https://doi.org/10.1016/j.clinph.2011.06.006

53. Zijlmans, M., Jiruska, P., Zelmann, R., Leijten, F. S. S., Jefferys, J. G. R., & Gotman, J. (2012). High-frequency oscillations as a new biomarker in epilepsy. *Annals of Neurology*, *71*(2), 169–178. https://doi.org/10.1002/ana.22548

Supplementary Figure 1.

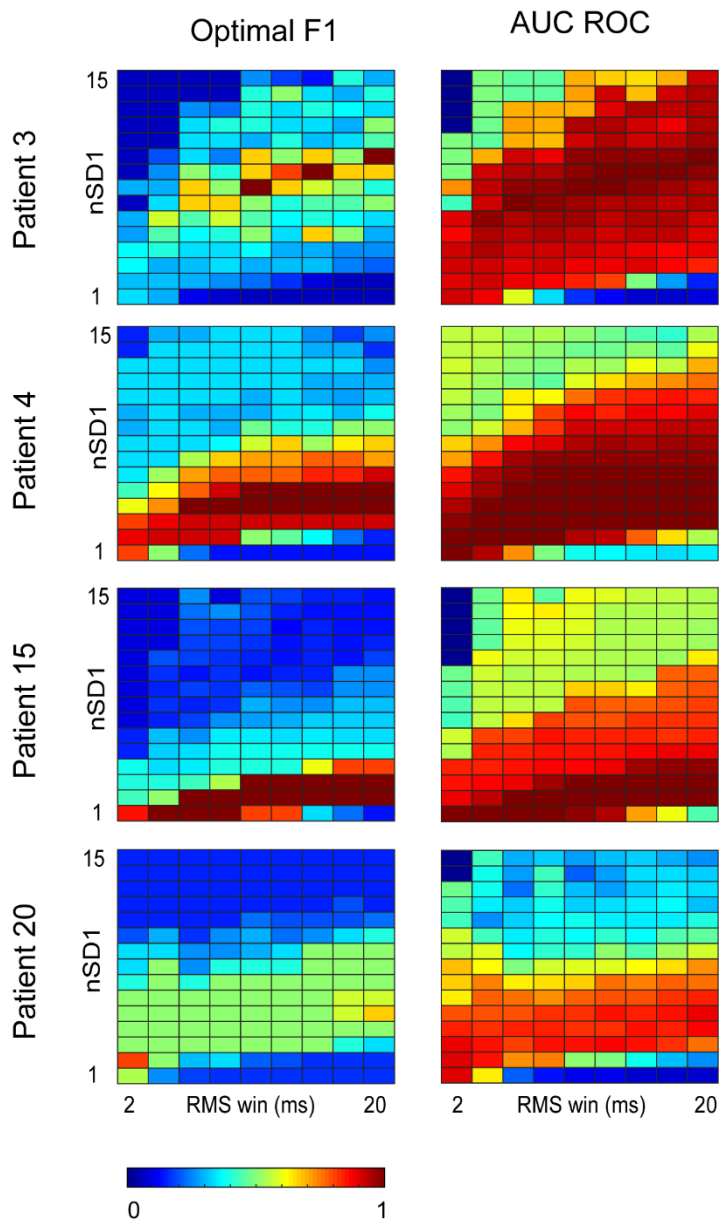

Figure S1*. Representative heatmaps of optimal F1 and AUC of the ROC curve using minimum event duration of 12ms (min_dur=12ms) are shown across the parameter space for patients 3 and 4 from the UCI dataset and patients 15 and 20 from the ETH Zurich dataset. This can be compared to the results in Figure 5 which used min_dur=6ms.*

Supplementary Figure 2

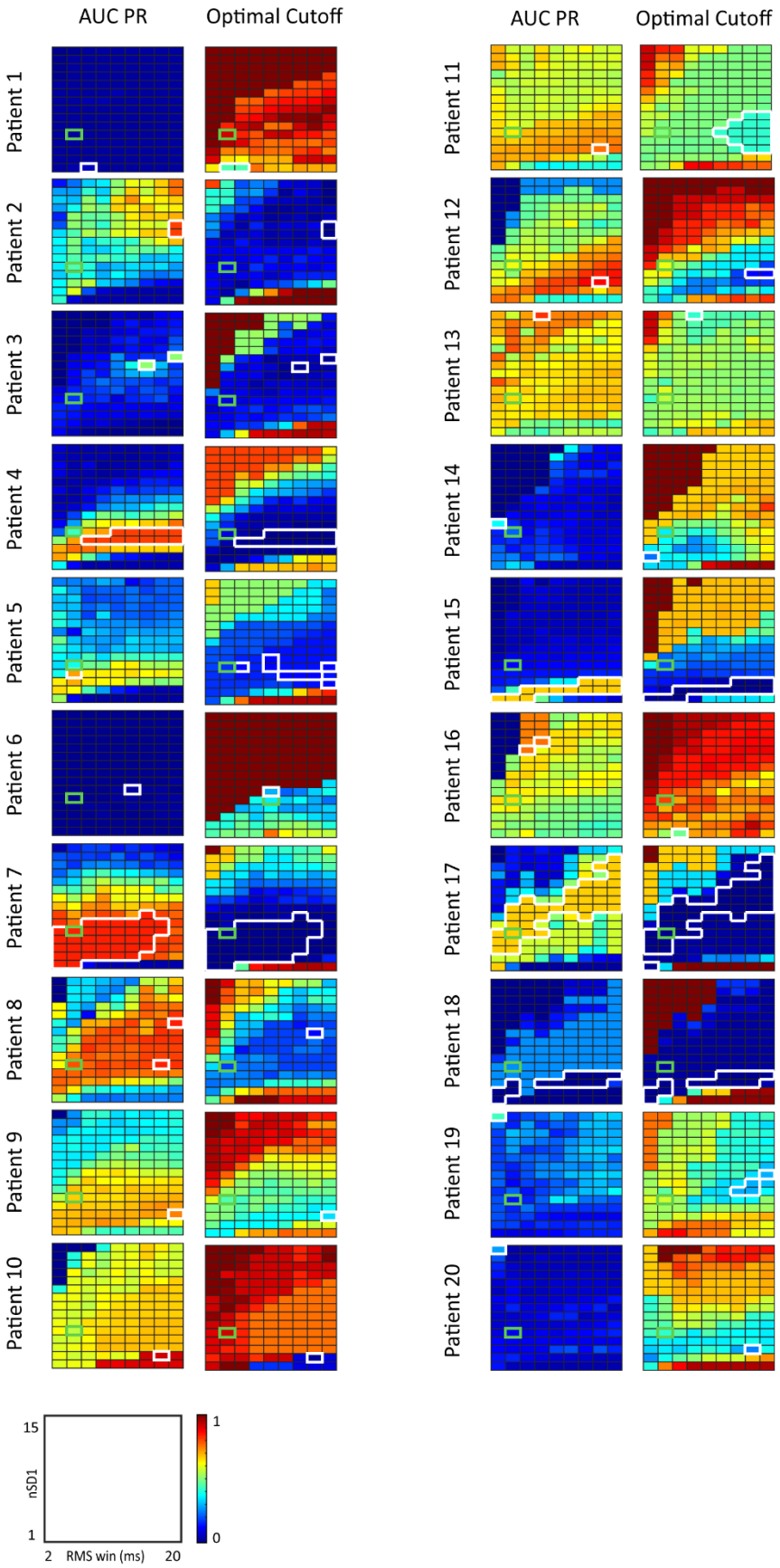

Figure S2. *AUC of the PR curve (left column) and optimal cutoff distance for the ROC curve (right column) across the parameter space are shown for all patients. Results based on the default parameters are outlined in green. Parameter sets achieving optimal SOZ/nSOZ classification are outlined in white.*