



OPEN Machine learning predicts spinal cord stimulation surgery outcomes and reveals novel neural markers for chronic pain

Jay Gopal¹, Jonathan Bao², Tessa Harland³, Julie G. Pilitsis⁴, Steven Panicioli⁵, Rachael Grey⁵, Michael Briotte⁵, Kevin McCarthy⁵ & Ilknur Telkes^{4,6,7}✉

Spinal cord stimulation (SCS) is a well-accepted therapy for refractory chronic pain. However, predicting responders remain a challenge due to a lack of objective pain biomarkers. The present study applies machine learning to predict which patients will respond to SCS based on intraoperative electroencephalogram (EEG) data and recognized outcome measures. The study included 20 chronic pain patients who were undergoing SCS surgery. During intraoperative monitoring, EEG signals were recorded under SCS OFF (baseline) and ON conditions, including tonic and high density (HD) stimulation. Once spectral EEG features were extracted during offline analysis, principal component analysis (PCA) and a recursive feature elimination approach were used for feature selection. A subset of EEG features, clinical characteristics of the patients and preoperative patient reported outcome measures (PROMs) were used to build a predictive model. Responders and nonresponders were grouped based on 50% reduction in 3-month postoperative Numeric Rating Scale (NRS) scores. The two groups had no statistically significant differences with respect to demographics (including age, diagnosis, and pain location) or PROMs, except for the postoperative NRS (worst pain: $p = 0.028$; average pain: $p < 0.001$) and Oswestry Disability Index scores (ODI, $p = 0.030$). Alpha-theta peak power ratio differed significantly between CP3-CP4 and T3-T4 ($p = 0.019$), with the lowest activity in CP3-CP4 during tonic stimulation. The decision tree model performed best, achieving 88.2% accuracy, an F1 score of 0.857, and an area under the curve (AUC) of the receiver operating characteristic (ROC) of 0.879. Our findings suggest that combination of subjective self-reports, intraoperatively obtained EEGs, and well-designed machine learning algorithms might be potentially used to distinguish responders and nonresponders. Machine and deep learning hold enormous potential to predict patient responses to SCS therapy resulting in refined patient selection and improved patient outcomes.

Keywords Machine learning, EEG, Chronic pain, Spinal cord stimulation, Responders

Chronic pain affects 50 million Americans (20.4% of US adults) as per estimates by the centers for disease control and prevention (CDC), and it has been linked to high healthcare costs as well as decreased quality of life¹. Spinal cord stimulation (SCS)- an FDA-approved widely used treatment, has been associated with improved clinical outcomes when compared to opioids². However, a significant portion of pain patients receive suboptimal care due to the lack of clear criteria for parameter and device selection, and patient selection still remains a challenge³. Unfortunately, the absence of objective biomarkers further complicates pain management and treatment⁴.

To address this challenge, EEG offers several advantages for understanding chronic pain and its treatment with SCS, making it a valuable tool in both clinical and research settings⁵. One of the most important aspects might be its compatibility with all clinically available SCS devices and its integration with other assessment tools. Intraoperative EEG is a commonly used method in neurosurgery. It has been used for dorsal root ganglion stimulation lead placement⁶ and studies have observed nociceptive activation in intraoperative EEG while

¹The Warren Alpert Medical School of Brown University, Providence, RI, USA. ²Albany Medical College, Albany, NY, USA. ³Department of Neurosurgery, Albany Medical College, Albany, NY, USA. ⁴Department of Neurosurgery, University of Arizona College of Medicine - Tucson, Tucson, AZ, USA. ⁵Nuvasive Clinical Services, San Diego, CA, USA. ⁶Charles E. Schmidt College of Medicine, Florida Atlantic University, Boca Raton, FL, USA. ⁷College of Engineering and Computer Science, Florida Atlantic University, Boca Raton, FL, USA. ✉email: itelkes@arizona.edu

patients are under general anesthesia⁷. However, its application in chronic pain management remains novel, with limited studies utilizing intraoperative EEG for comprehensive analysis of neural pain patterns^{8,9}.

There are currently a variety of different self-reporting methods to quantify pain^{10–16}. However, due to the complex and multifactorial nature of pain, these methods often limited to capture its full scope, highlighting the need for machine learning (ML) to integrate diverse physiological, neural, and clinical data for more objective and comprehensive pain assessment. ML is a rapidly growing field in pain assessment. Studies have found that ML models can approximate self-reported pain scores, the current clinical standard, solely based on physiological features recorded with real-world context^{17,18}. For example, researchers demonstrated that ML models trained on baseline features, including demographics, diagnostic features, and self-reported pain scores - without incorporating any neural data - could predict responders to SCS with an accuracy of 50–73%^{19,20}. A recent study recorded intracranial EEG and showed that self-reported pain scores could be predicted with high sensitivity from neural features, but was limited by the small sample size of just four participants²¹.

The successful prediction of SCS outcomes for chronic pain patients would represent a significant stride forward for personalized medicine and science as a whole. In the short term, intraoperative prediction would accelerate precise device selection, avoiding the need for patients to have multiple operations to sample several devices. Predicting responders would also empower patients by giving them insight into their most likely outcomes before they go through the months-long process of having their SCS settings clinically optimized. In the long term, models trained to predict responders could be used to unlock the neurophysiological bases of chronic pain, helping us understand and design more personalized SCS. The current study aims to predict which chronic pain patients will respond positively to SCS based on intraoperative neural patterns. To date, no studies have published results for this task in the literature.

Methods

Participants

Twenty chronic pain patients who underwent spinal cord stimulation implant surgery for their chronic back and/or leg pain as standard of care were included in the present study. All patients provided written informed consent to participate in the study. The experimental protocol was approved by the Institutional Review Board of Albany Medical Center (IRB Number: 4973), and the research was conducted in accordance with the Helsinki Declaration.

Patient reported outcome measures

To monitor the patients' levels of pain over time, a number of self-reported scores were recorded as per routine protocol. These included Numeric Rating Scale (NRS) scores, which is used to assess pain intensity; Beck Depression Inventory (BDI), which measures the extent of depressive symptoms due to chronic pain¹³; Oswestry Disability Index (ODI), which quantifies the degree of disability associated with chronic pain¹⁴; McGill Pain Questionnaire (MPQ) assesses many aspects of pain, including its emotional impact and intensity¹⁵; and Pain Catastrophizing Scale (PCS) to measure how much patient's chronic pain induces thoughts about worst-case scenarios or extremely pessimistic futures¹⁶. These scores were collected preoperatively and 3 months (on average) after surgery. To add depth to the measurements, the patients were also asked about their absolute worst, best, and average levels of pain over the preceding week on the NRS at each follow-up. To obtain a more robust pain intensity score, NRS-now and NRS-average were averaged, and responders were defined as those with a decrease in this score (labeled as NRS- average) of at least 50% within the first three months. This threshold was used because it is a widely accepted standard and was included in the IMMPACT recommendations as the criteria for substantial improvement in pain²².

Surgical procedure and neuro monitoring

Surgeries were performed while patients were under general anesthesia at appropriate levels for intraoperative neurophysiological monitoring (IONM) per a widely accepted routine²³. Leads were placed covering T9–10 and T8–9 either via laminectomy or percutaneously using a standard of care surgical methods as previously described^{23,24}. Briefly, the incision site and lead placement were planned preoperatively with respect to the “sweet spot” from trial identifying the ideal stimulation location. An initial midline incision was made over the posterior spinous process, guided by fluoroscopic localization. Electrocautery and surgical dissection were used to expose the spinous process and lamina. Following intraoperative fluoroscopic confirmation of the correct level, a laminectomy was performed adequate for placement of the SCS paddle. Serial fluoroscopic imaging and IONM confirmed placement and intact neurological status. The routine IONM protocol included monitoring muscle activity in the abdominals, quadriceps, tibialis anterior, abductor hallucis, medial gastrocnemius, biceps femoris, and gluteus maximus under stimulation^{23,25}. To check laterality, each lead was temporarily connected to a manufacturer-specific testing device, a contact of interest is selected as the active contact based on the predicted ‘sweet spot’, and compound muscle action potentials (CMAPs) in response to tonic stimulation (60 Hz, 300 μ s) at gradually increased amplitudes (0-to-10 mA) were monitored. Once placement was deemed adequate for long-term therapy and implantable pulse generator (IPG) was placed subcutaneously, our EEG experiment was performed by connecting the commercial paddle to its pulse generator and delivering stimulation to the same active contacts. The surgery was completed in standard fashion^{23,26}.

EEG recording and analysis

Intraoperative EEG data was recorded at 10 channels (F3–4, FP1–2, C3–4, CP3–4, and T3–4) in accordance with the extended 10–20 system using the Cadwell Cascade Pro IONM systems (Cadwell Inc., Kennewick, Wash., USA). EEG signals were bandpass filtered at 1–58 Hz and sampled at 128 Hz. A 90-second baseline (stimulation OFF) was recorded at the beginning of each session, followed by randomized, 60-second-long alternating

stimulation ON and OFF conditions. The following types of stimulation were recorded twice each: (i) Tonic at 60 Hz/300 μ s and (ii) High-density (HD) Stimulation at 1 kHz/30 μ s. The stimulation OFF periods separated the stimulation ON periods to allow for signal recovery. The contacts and current amplitude were determined by SCS screening trials.

Data was exported to MATLAB (Mathworks, Natick, MA) for offline analysis using custom scripts in MATLAB R2021a (Mathworks, Natick, MA). All EEG recordings were visually examined to identify corrupted channels/trials, as well as mechanical artifacts. After corrupted data was rejected, signals were filtered via a second order Butterworth high pass filter (2 Hz). Independent component analysis (ICA) was performed to remove heartbeat artifacts²⁷. Subsequently, a common average reference (CAR) filter was applied to each trial to remove common activity across all channels²⁸.

To estimate power spectral density of EEG signals in each condition, a fast Fourier transform (FFT) was computed in 5-second epochs with a 1-second sliding Hanning window²⁹. Next, the average magnitude of Fourier coefficients across trials was computed for the theta (4–7 Hz), alpha (8–12 Hz), beta (13–33 Hz), and gamma (35–50 Hz) subbands, as well as the overall spectrum (1–50 Hz). Subband power values were normalized for each stimulation condition relative to baseline on a logarithmic decibel (dB) scale. Additionally, band specific peak frequency and peak power ratio between alpha and theta bands were calculated. Each of these EEG features were extracted for all contacts, allowing for an investigation of the topological distributions. This feature extraction approach yielded total of 100 relative power features (2 SCS conditions x 10 channels x 5 bands), 30 peak power ratio features (3 SCS conditions x 10 channels x 1 band), and 150 peak frequency features (3 SCS conditions x 10 channels x 5 bands) per subject.

Statistics

Participant data was summarized by mean \pm standard error of mean. A Mann-Whitney *U* test was used to compare age, sex, and all PROMs between responders and nonresponders. The Fisher's exact test was used to compare diagnosis (4 categories) and pain location (5 categories) between responders and nonresponders and identify the significant variables. To compare the preoperative and postoperative PROMs *within* groups, a two-tailed Wilcoxon signed-rank test was conducted.

The Shapiro-Wilk test and Levene's test were used to assess the normality of the feature distributions and the homogeneity of variances, respectively. Spectral features passed the tests were analyzed using an independent samples t-test and a 3-way analysis of variance (ANOVA) with the independent factors of group (responders and nonresponders), SCS condition (OFF, tonic, HD), and cortical region (prefrontal, frontal, motor, somatosensory, and temporal). Features with non-normal distributions were analyzed using the Mann-Whitney *U*. *Post hoc* analyses were done by Bonferroni correction. Descriptive statistics were reported as mean \pm standard deviation. Analysis was conducted using IBM SPSS version 29.0.1.1 (SPSS Inc., Chicago, IL, United States), with a significance threshold of $p=0.05$. Correlations between the neural features in each region and preoperative PROMs were analyzed via Spearman's rank order correlation using MATLAB statistical toolbox (Mathworks, Natick, Massachusetts).

Machine learning

To determine EEG features that account for the most variance, we applied principal component analysis (PCA) to rank-order the features, enabling objective selection of inputs for the ML classifiers³⁰. We then used leave-one-out cross-validation (LOOCV) across all patients, training on all but one patient in each iteration and testing on the excluded patient. To promote model generalizability, feature selection was conducted only once prior to LOOCV, ensuring that the same features were used to predict across all folds.

For hyperparameter tuning and random seed selection, we conducted a grid search before beginning LOOCV, which ensured consistent hyperparameters across all iterations and avoided tailoring model settings or initializations to specific data subsets. We trained various classifiers, including logistic regression, decision trees, support vector machines (SVM), XGBoost, and random forests^{31–34}. Models were evaluated using accuracy, F1 score, and area under the curve (AUC) of receiver operating characteristic (ROC) aggregated across folds. For the best architecture from the initial set of experiments, we performed a second grid search to optimize the number of EEG features that were to be used as inputs from the rank-ordered list from PCA, continually adding the remaining features with the highest rank.

Results

Patient cohort and outcome measures

A total of 20 patients with chronic pain (10 female, aged 54.4 ± 2.95 years on the date of surgery) were recruited for this study. Three patients were excluded from the study due to insufficient postoperative NRS data for classification. Of the 17 patients included, 7 subjects were grouped as responders with a mean age of 54.14 ± 5.82 years and 10 subjects were grouped as nonresponders with a mean age of 53.1 ± 3.41 years (Table 1). All patients were diagnosed with chronic neuropathic pain (CNP), including specific types such as chronic low back pain (CLBP), persistent spinal pain syndrome (PSPS), and complex regional pain syndrome (CRPS). Particularly in responders, 3 patients were treated for CLBP, 2 were treated for CNP, 1 was treated for PSPS, and 1 was treated for CRPS while, in nonresponders, 3 were treated for CLBP, 6 were treated for PSPS, and 1 were treated for CNP. Most of the participants suffered from lower back and leg pain (responder/nonresponder: 5/4) while 2 responders suffered from lower back pain only, 2 nonresponders suffered from bilateral leg pain only, and 4 nonresponders suffered from pain in another body area(s). Statistical analyses did not show a significance difference in age ($U=31$, $p=0.696$), sex ($U=32.5$, $p=0.778$), diagnosis ($p=0.199$), or pain location ($p=0.111$) between responders and nonresponders.

	Responders	Nonresponders	P-value
Sample size	7	10	
Sex (n)			
male: female	4:3	5:5	0.778
Age (years)	54.14 ± 5.82	53.1 ± 3.41	0.696
Diagnosis (n)			0.199
Persistent spinal pain syndrome	1	6	
Chronic neuropathic pain	2	1	
Chronic low back pain	3	3	
Complex regional pain syndrome	1	0	
Pain location			0.111
Lower back pain	2	0	
Bilateral leg pain	0	2	
Lower back pain and leg pain	5	4	
Other, single area	0	2	
Other, multiple areas	0	2	
Baseline scores			
NRS-worst	7.14 ± 1.06	9 ± 0.37	0.118
NRS-average	5.93 ± 0.73	7.05 ± 0.38	0.169
BDI	10.71 ± 2.67	15.30 ± 4.22	0.733
MPQ	5.14 ± 1.10	6.90 ± 1.19	0.404
PCS	26.60 ± 4.45	18.33 ± 4.53	0.317
ODI	49.14 ± 3.02	48.80 ± 4.67	0.768
Postoperative scores (3-month)			
NRS-worst	4.14 ± 1.01	7.4 ± 0.70	0.028
NRS-average	1.93 ± 0.34	6.60 ± 0.55	<0.001
BDI	8.14 ± 1.67	12.80 ± 3.21	0.524
MPQ	3.00 ± 0.49	5.80 ± 1.31	0.128
PCS	8.71 ± 2.01	18.22 ± 5.15	0.396
ODI	27.67 ± 5.48	46 ± 4.69	0.030

Table 1. Demographic and clinical data for responders and nonresponders. Age and all pain scores are reported as mean ± standard error of mean. NRS: Numeric rating scale; BDI: Beck's depression inventory; MPQ: McGill pain questionnaire; PCS: Pain catastrophizing scale; ODI: Oswestry disability index. *P* values are generated by Mann-Whitney *U* test or Fisher's exact test. The statistically significant differences are highlighted in bold.

PROMs were collected from each patient preoperatively and at 3 months. None of the preoperative outcome measures, including NRS-worst ($U = 19.5$, $p = 0.118$), NRS-average ($U = 21$, $p = 0.169$), BDI ($U = 31.5$, $p = 0.733$), MPQ ($U = 26.5$, $p = 0.404$), PCS ($U = 15$, $p = 0.317$), and ODI ($U = 32$, $p = 0.768$), demonstrated statistically significant difference between responders and nonresponders. Statistical analysis indicated significantly lower postoperative scores of NRS-worst ($U = 13$, $p = 0.028$), NRS-average ($U = 1$, $p < 0.001$), and ODI ($U = 10$, $p = 0.030$) in responders while BDI, MPQ, and PCS did not show a significant difference ($U = 28.5$, $p = 0.524$; $U = 19.5$, $p = 0.128$; and $U = 23.5$, $p = 0.396$, respectively). Within group comparisons in responders demonstrated significant improvement in NRS-average ($p = 0.018$), PCS ($p = 0.043$), and ODI ($p = 0.046$) from preoperative to 3-month postoperative period while no significant improvement was noted in NRS-worst ($p = 0.088$), BDI ($p = 0.343$), and MPQ ($p = 0.089$). In nonresponders, only significant reduction was noted in NRS-worst ($p = 0.041$) following surgery (NRS-average: $p = 0.192$; BDI: $p = 0.125$; MPQ: $p = 0.089$; ODI: $p = 0.672$; PCS: $p = 0.469$). Additional patient characteristics can be found in Table 1.

Spectral features with correlations

Global activity

When examining global activity changes across different conditions (baseline, tonic, HD), we observed that average spectrum (1–50 Hz, Fig. 1A, B) in responders showed stronger and faster alpha trends with greater variance ($5.39 \pm 4.43 \mu\text{V}^2/\text{Hz}$; $5.36 \pm 4.84 \mu\text{V}^2/\text{Hz}$; $5.59 \pm 5.05 \mu\text{V}^2/\text{Hz}$, respectively) compared to nonresponders ($4.43 \pm 4.67 \mu\text{V}^2/\text{Hz}$; $4.65 \pm 2.19 \mu\text{V}^2/\text{Hz}$; $4.75 \pm 2.07 \mu\text{V}^2/\text{Hz}$, respectively). However, these differences were not statistically significant in any condition ($p > 0.05$). While none of the differences in global activity between responders and nonresponders reached statistical significance, higher trends in variance in HD may suggest varying stimulation-induced cortical responses. The global theta trends (Supplementary Fig. 1) were higher in nonresponders ($5.09 \pm 1.86 \mu\text{V}^2/\text{Hz}$; $5.15 \pm 2.17 \mu\text{V}^2/\text{Hz}$; $5.24 \pm 2.39 \mu\text{V}^2/\text{Hz}$, respectively) compared to responders ($4.50 \pm 2.76 \mu\text{V}^2/\text{Hz}$; $4.08 \pm 2.76 \mu\text{V}^2/\text{Hz}$; $4.02 \pm 2.69 \mu\text{V}^2/\text{Hz}$, respectively) in all conditions; however,

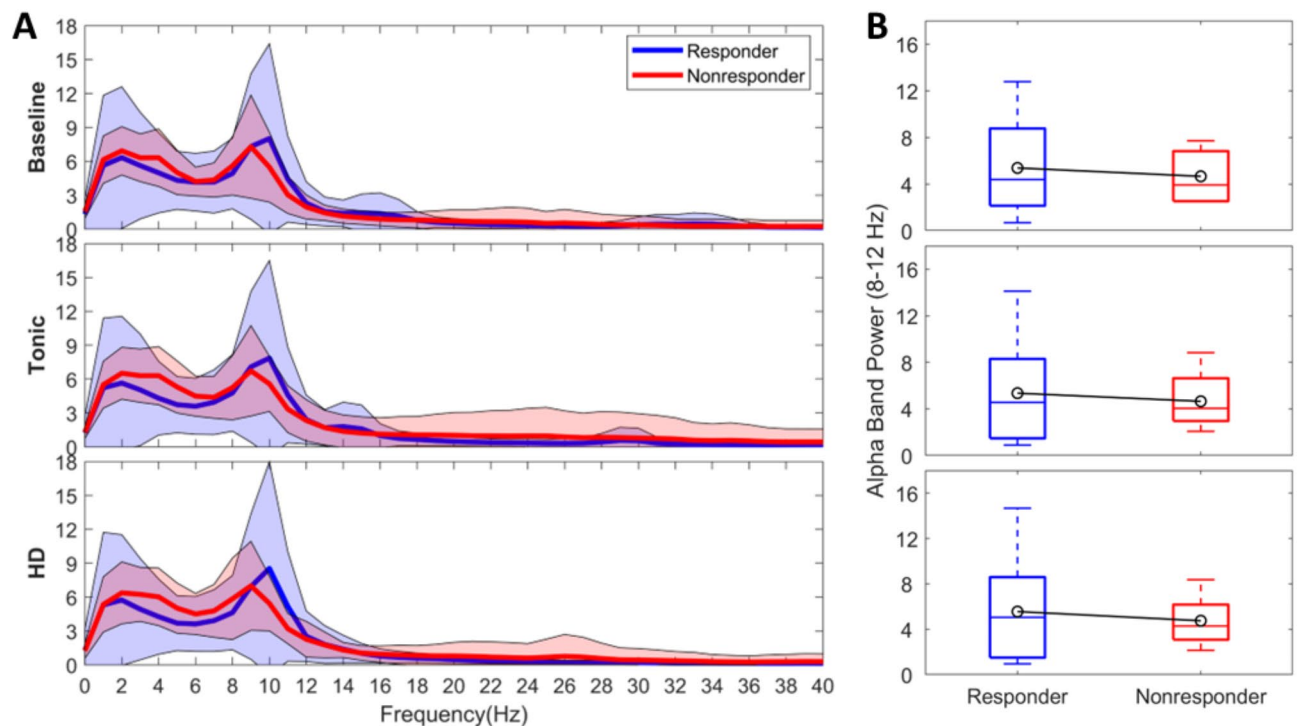


Fig. 1. Global activity. **(A)** Power spectrum density (PSD) estimates. PSDs were averaged across EEG regions and patients per group during stimulation OFF, tonic and HD. Blue: Responders. Red: Nonresponders. Shaded area indicating \pm standard deviation. HD: high-density. **(B)** Global alpha band activity. Boxplots indicating global alpha activity (8–12 Hz) during baseline (top row), tonic stimulation (middle row), and HD stimulation (bottom row). Blue: Responders. Red: Nonresponders. Black lines indicating the groups' mean.

these changes were not statistically different ($p > 0.05$). Global beta band activity (13–33 Hz) decreased with both stimulation conditions in responders, more uniformly with HD resulting in a smaller variance ($0.50 \pm 0.41 \mu\text{V}^2/\text{Hz}$). The highest variance in gamma band (35–50 Hz) was localized to nonresponders and noted in tonic stimulation ($0.39 \pm 1.03 \mu\text{V}^2/\text{Hz}$).

Alpha-theta peak power ratio

A three-way ANOVA showed a marginal difference for group factor indicating that responder-nonresponder factor might affect the alpha-theta peak power ratio ($F(1,225) = 3.44$, $p = 0.055$). We noted that responders showed higher alpha-theta ratio compared to nonresponders (1.47 ± 0.26 dB vs. 0.84 ± 0.22 dB). Additionally, pairwise comparisons demonstrated a significant difference in alpha-theta peak power ratio between CP3-CP4 and T3-T4 (*adjusted* $p = 0.019$), with the strongest activity in T3-T4 during HD stimulation (Fig. 2A, bottom-row), and the lowest in CP3-CP4 during tonic stimulation in nonresponders (Fig. 2A, middle-row). Alpha-theta peak power ratio was notably higher in frontal (F3-F4) and prefrontal (FP1-FP2) regions in responders; however, it did not survive multi-comparison correction. No significant group \times region \times SCS interaction was found ($p > 0.05$).

As shown in Fig. 2B, activity in FP1-FP2 during baseline was significantly correlated with preoperative NRS-worst and NRS-average scores in responders ($r = -0.829$, $p = 0.030$; $r = -0.800$, $p = 0.041$, respectively). This indicated that patients with severe pain scores tended to have smaller prefrontal region alpha-theta peak power ratio. While trends were similar in the nonresponders, correlation was not significant. Figure 2C presents the significant correlation between alpha-theta peak power ratio in C3-C4 during baseline and preoperative BDI scores in nonresponders ($r = -0.709$, $p = 0.028$). Similarly, a significant correlation was found between baseline CP3-CP4 and preoperative BDI scores in nonresponders ($r = -0.660$, $p = 0.044$). These indicated that patients with severe depression scores tended to have smaller central-centroparietal alpha-theta peak power, indicating stronger theta activity. While trends were similar in the responders, correlation was not significant.

Global relative power

The relative global power (1–50 Hz) also demonstrated a significant main effect of group factor, indicating that responder-nonresponder factor might affect the global activity - which was computed with respect to baseline ($F(1,150) = 5.99$, $p = 0.016$). The relative global activity was significantly smaller in responders (*adjusted* $p = 0.016$) and the smallest activity was localized to CP3-CP4 under HD (-0.72 ± 1.25 dB) followed by tonic stimulation (-0.62 ± 1.25 dB). While not statistically significant ($p > 0.05$) following the multi-comparison correction, descriptive statistics of *within*-group distribution indicated that responders demonstrated lower activity in central and frontal-prefrontal regions under tonic stimulation (Fig. 3A). While trends in central area

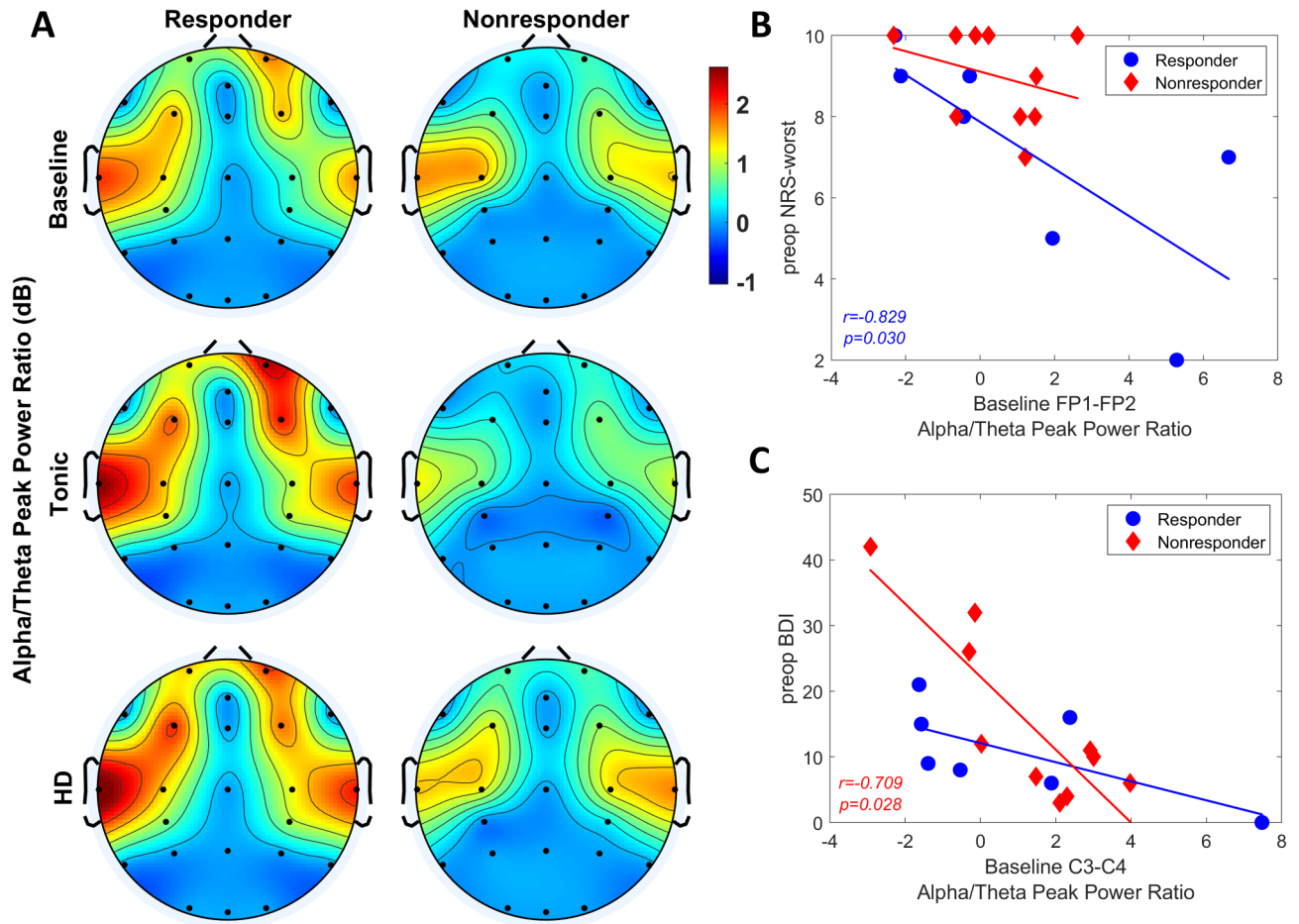


Fig. 2. Comparison of alpha/theta peak power ratio between responders and nonresponders. **(A)** Topographical maps of alpha-theta peak power ratio in responders and nonresponders during stimulation OFF (top-row), tonic (middle-row) and HD (bottom-row). Color bar indicating alpha-theta peak power ratio in decibel (dB) scale. **(B)** Correlation analysis between alpha/theta peak power ratio in FP1-FP2 during baseline and preoperative Numeric Rating Scale (NRS) worst pain scores. Blue circle: Responders. Red diamond: Nonresponders. **(C)** Correlation analysis between alpha/theta peak power ratio in C3-C4 during baseline and preoperative Beck Depression Inventory (BDI) scores.

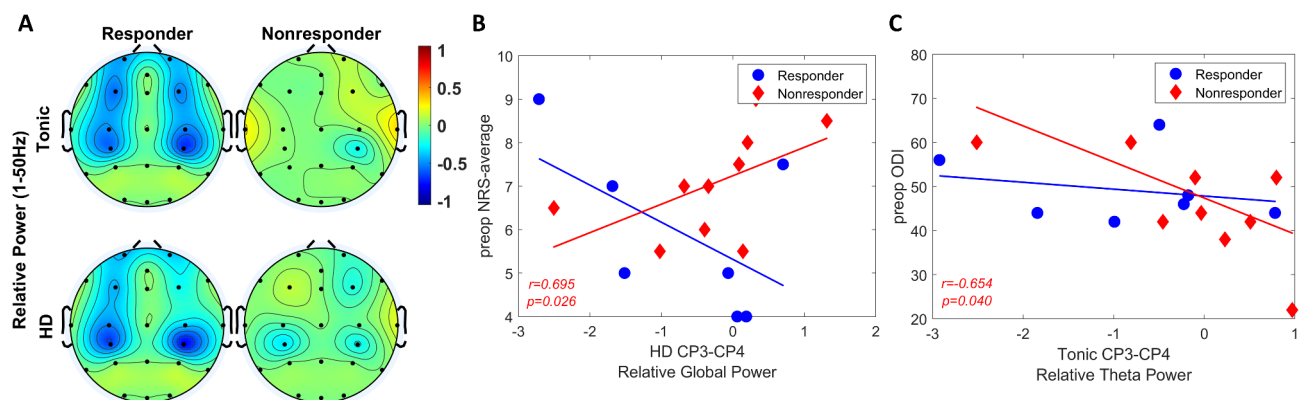


Fig. 3. Comparison of relative power between responders and nonresponders. **(A)** Topographical distribution of relative global (1–50 Hz) power in tonic (top-row) and HD (bottom-row). Color bar indicating relative power in decibel (dB) scale. **(B)** Correlation analysis between relative global power in CP3-CP4 during HD stimulation and preoperative Numeric Rating Scale (NRS) average pain scores. Blue circle: Responders. Red diamond: Nonresponders. **(C)** Correlation analysis between relative theta (4–7 Hz) power in CP3-CP4 during tonic stimulation and preoperative Oswestry Disability Index (ODI) scores.

remained low under HD similar to tonic, HD induced stronger activity in frontal-prefrontal regions. Trends in nonresponders did not show a change in response to tonic or HD stimulation. Analysis did not show a significant interaction across group, region, and SCS condition.

Correlation analysis between global relative power and the subjective scores demonstrated opposite trends between the groups under HD and it was limited to centroparietal region. Particularly, preoperative NRS-average scores in nonresponders indicated positive correlation with CP3-CP4 relative power ($r=0.695$, $p=0.026$; Fig. 3B).

Subband relative power

Analyzing relative theta power (4–7 Hz), a significant main effect of group ($F(1,150)=13.54$, $p<0.001$) was observed. The relative theta activity was significantly smaller in responders (*adjusted* $p<0.001$) and the smallest activity was localized to CP3-CP4 under HD (-0.90 ± 0.99 dB and -0.24 ± 1.13 dB) followed by tonic stimulation (-0.84 ± 1.22 dB and -0.17 ± 0.99 dB) in both groups (Supplementary Fig. 2A). No significant interaction was found among group, region, and SCS conditions. Relative theta power showed significant correlation only in CP3-CP4 region under tonic stimulation with preoperative ODI scores in nonresponders ($r=-0.654$, $p=0.040$; Fig. 3C).

Relative alpha power (8–12 Hz) displayed no significant main effect of group or interaction. While not significant, descriptive statistics showed higher alpha activity in HD compared to tonic condition (0.0 ± 0.12 dB vs. -0.13 ± 0.12 dB). Interestingly, the smallest relative alpha activity was induced by HD in responders in CP3-CP4 region while it was tonic stimulation in nonresponders in the same cortical area (Supplementary Fig. 2B). In beta band (13–33 Hz), a significant group effect ($F(1,150)=8.43$, $p=0.004$) indicated reduced relative beta power in responders (*adjusted* $p=0.004$), with no significant group-region-SCS interaction. The weakest beta activity was observed in CP3-CP4 with HD stimulation for both groups (Supplementary Fig. 2C). No significant correlation was found between the relative alpha or beta power and the PROMs. In contrast, the lowest gamma activity (35–50 Hz) was localized to FP1-FP2 region with HD stimulation in responders (-0.91 ± 1.99 dB), and to CP3-CP4 region (-0.91 ± 2.40 dB) in nonresponders (Supplementary Fig. 2D). Tonic stimulation produced the strongest gamma activity in C3-C4 for both groups. Relative gamma power showed no significant group main effect or 3-factor interaction. A significant correlation was found between the tonic-induced relative gamma power in C3-C4 and the preoperative NRS-average scores in responders ($r=-0.855$, $p=0.021$).

Subband peak frequencies

Three-way ANOVA revealed a significant main effect of group in theta frequency ($F(1,225)=34.29$, $p<0.001$); however, there was no significant interaction noted between group, region, and SCS condition ($p=0.965$). Subband peak frequencies computed per region were compared between groups across SCS conditions (Fig. 4A). While not statistically significant ($p>0.05$) following the multi-comparison correction, descriptive statistics indicated trend where responders exhibited faster theta peak frequency (5.3 ± 0.09 Hz) compared to nonresponders (4.6 ± 0.07 Hz). Further, the fastest theta rhythm was noted in the C3-C4 area for both groups (Responder: 5.4 ± 0.2 Hz; Nonresponder: 4.7 ± 0.2 Hz) while slowest theta rhythm was found in F3-F4 for responders (5.1 ± 0.2 Hz) and CP3-CP4 for nonresponders (4.5 ± 0.2 Hz). HD stimulation showed faster theta frequency in CP3-CP4 (Responder: 5.5 ± 0.3 Hz; Nonresponder: 5.1 ± 0.3 Hz) compared to baseline ($p=0.046$).

Analysis of peak frequency showed no significant interaction between group, region, and SCS for any subband or the entire spectrum ($p>0.05$). Descriptive statistics demonstrated that the fastest alpha rhythm was in C3-C4 under HD stimulation (9.3 ± 0.2 Hz), and in T3-T4 under tonic stimulation for both responder (9.4 ± 0.3 Hz) and nonresponders (9.4 ± 0.2 Hz). Beta peak frequency averaged 13.5 ± 0.2 Hz in responders and 13.8 ± 0.1 Hz in nonresponders, with the highest beta activity was in F3-F4 during baseline for responders (14.1 ± 0.6 Hz) and in FP1-FP2 with tonic for nonresponders (14.5 ± 0.5 Hz). The slowest beta frequency occurred in HD T3-T4 for responders (13.2 ± 0.6 Hz) and baseline C3-C4 for nonresponders (13.2 ± 0.5 Hz). Under HD stimulation, trends in gamma rhythms were slower in responders (36.7 ± 0.3 Hz) than in nonresponders (37.2 ± 0.3 Hz). For responders, trends in global rhythms averaged 5.7 ± 0.7 Hz at baseline and increased to 6.1 ± 0.7 Hz with tonic stimulation, while nonresponders showed no change in trends (6.3 ± 0.6 Hz).

Further analysis demonstrated significant correlations in responders localized to alpha band and nonresponders to the theta band. More specifically, theta peak frequency under HD stimulation in nonresponders was negatively correlated with preoperative MPQ and ODI scores in both C3-C4 ($r=-0.878$, $p<0.001$; $r=-0.819$, $p=0.004$, respectively) and CP3-CP4 ($r=-0.744$, $p=0.014$; $r=-0.673$, $p=0.033$, respectively) regions (Fig. 4B). In alpha band, significant correlations were found to be between HD-induced alpha peak frequency and preoperative NRS-average scores in responders, localized to CP3-CP4 ($r=-0.982$, $p=0.002$) and FP1-FP2 ($r=-0.844$, $p=0.029$) regions (Fig. 4C).

Machine learning

PCA was performed with the EEG data to rank top features of hundreds of neural features by explained variance. Our feature selection approach resulted in 12 features that can be considered an indicator of response to SCS (Fig. 5A). It was indicated that the five most important EEG features were baseline theta peak frequency in C4, HD-induced alpha peak frequency in C4, baseline alpha peak frequency in C4, HD-induced global relative power in F4, and tonic-induced gamma relative power in F3. A list of the top EEG features included as inputs to the classifier is provided in Table 2. The top EEG features, along with demographics, clinical features, and self-reported pain scores, were used as inputs to an ML model that predicted responders.

The decision tree classifier performed best, with an accuracy of 88.2%, F1 score of 0.857, and AUC-ROC of 0.879 (Table 3). Importantly, the optimal number of EEG features were relatively consistent across the majority of the architectures compared (see Supplementary Fig. 3 for further information). For the decision tree, 12 EEG

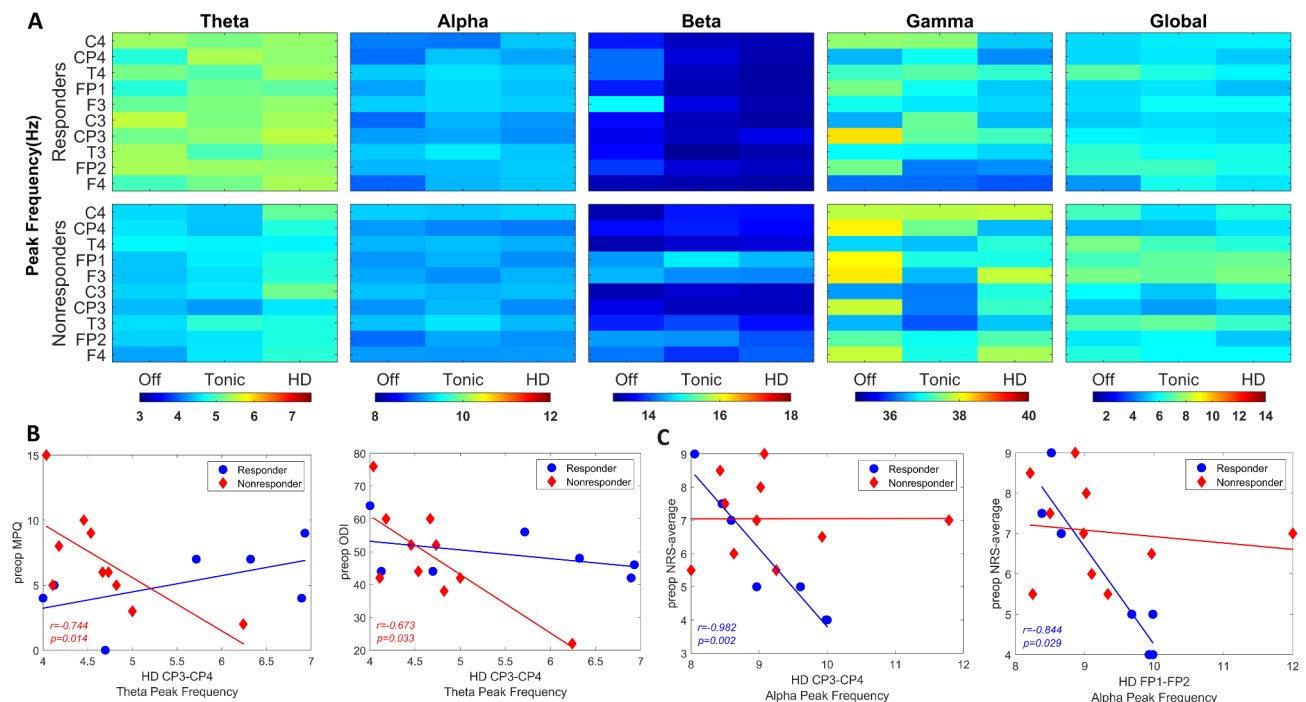


Fig. 4. Subband peak frequencies between responders and nonresponders. **(A)** Heatmaps showing the subband peak frequency per SCS condition in each EEG region in responders (top-row) and nonresponders (bottom-row). Color bar indicating the frequency in Hz. **(B)** Correlation analysis between theta peak frequency in CP3-CP4 during HD stimulation and preoperative McGill Pain Questionnaire (MPQ) and Oswestry Disability Index (ODI) scores. **(C)** Correlation analysis between alpha peak frequency in CP3-CP4 and FP1-FP2 regions during HD stimulation and preoperative Numeric Rating Scale (NRS) scores. Blue circle: Responders. Red diamond: Nonresponders.

features as inputs led to the best prediction as measured by AUC-ROC, providing enough information without diluting the signal with noise (Fig. 5B). The decision tree also had the highest accuracy and F1 score at this input setting. When each model was given its optimal number of EEG features, the decision tree had the maximal AUC-ROC (Fig. 5C). The confusion matrix of the optimized decision tree is shown in (Fig. 5D). The model only had one false positive and one false negative across the 17 patients included for the ML experiments.

Discussion

The present study develops predictive models for SCS outcomes using intraoperative EEG features as inputs for a machine learning classifier. Selecting candidates for SCS surgery is challenging due to a lack of clear biomarkers for chronic pain, but our work brings the field one step closer towards objective patient and device selection.

Numerous works have highlighted variable neural responses, particularly in areas like the prefrontal and somatosensory cortices, associated with pain processing in patients suffering from conditions like CRPS, PSPS, and NP^{35,36}. Our study has revealed distinct neural patterns associated with responders to SCS. For example, our findings suggest that the responder status may influence the alpha-theta peak power ratio, with responders showing higher values compared to nonresponders. The observed differences in the alpha-theta peak power ratio between regions, particularly the stronger difference between somatosensory and temporal area, highlight the region-specific effects of SCS. A similar group effect was observed in features such as relative global power, relative theta power, and theta peak frequency, where the relative activity was significantly lower in responders, but peak frequency was significantly higher. Furthermore, the smallest activity in these spectral features was consistently localized to the somatosensory region under HD. While no significant interaction between group, region, and SCS was found, these findings indicate that SCS may modulate neural activity differently in responders, potentially reflecting distinct neurophysiological responses to treatment.

Alpha-theta peak power ratio in prefrontal cortex and central regions during baseline was inversely correlated with pain intensity, indicating that responders experiencing severe pain displayed slower rhythms before stimulation. Chronic pain and depression are closely related³⁷ and high levels of psychological distress are associated with altered pain processing and perception³⁸. Although no significant differences in BDI scores were observed between groups, nonresponders with elevated depression scores showed a tendency toward theta activity in central-centroparietal regions, potentially indicating altered attentional and emotional processing of pain stimuli. Correlations between objective and subjective measures showed parallel trends in most features between the groups, except for somatosensory global power and theta peak frequency in HD. Particularly, relative global power tended to be higher in nonresponders with higher pain intensity; and slower theta rhythms with stronger sensory experience of pain. Overall, these findings provide insights into the neurophysiological

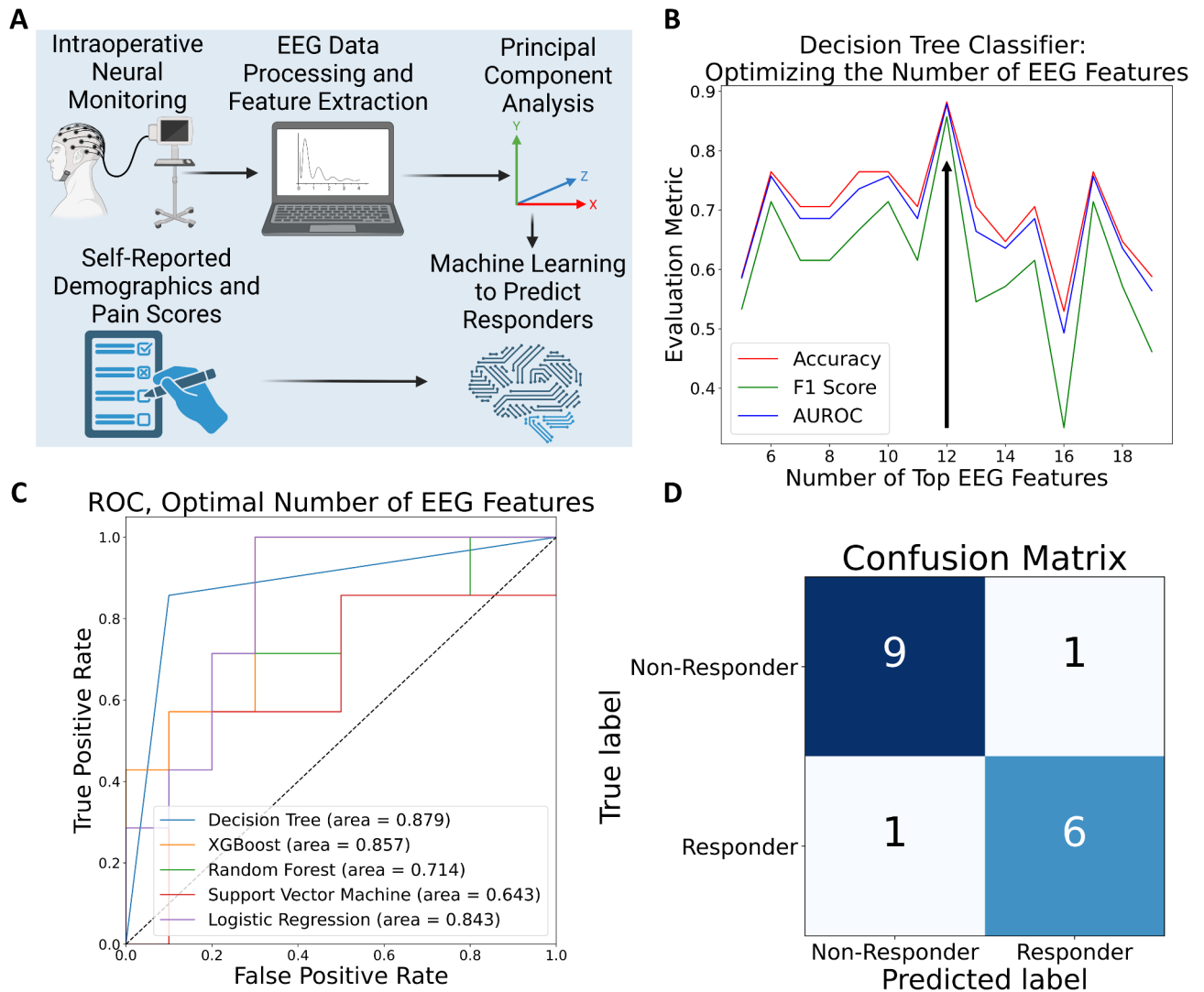


Fig. 5. Machine learning pipeline. **(A)** Diagram of full pipeline. EEG data was collected during SCS implant surgery and was subsequently exported to MATLAB for offline analysis. The data was preprocessed and the relevant neural features were extracted. Principle component analysis (PCA) was employed to rank these features by explained variance. The top features, along with demographics, clinical features, and pain scores, were used as inputs to a machine learning model that predicted responders. **(B)** Comparison of different numbers of EEG features as input to the decision tree classifier. Accuracy, F1 score, and area under the receiver operating characteristic curve (AUROC) are shown. The optimal number of EEG features was defined as that with the highest AUROC, indicated by a vertical arrow. **(C)** ROC for the optimal version of each of the architectures compared. AUROC is noted in the figure legend. **(D)** Confusion matrix for the optimal decision tree classifier.

mechanisms underlying the response to SCS in individuals with chronic pain, highlighting potential EEG measures for predicting treatment response and identifying differences in neural processing between responders and nonresponders.

The use of ML in neurosurgery is notable for its ability to handle large, complex datasets and generate predictive outputs crucial for patient-specific, AI-driven decision-making. In spine surgery, ML applications include predicting post-surgical outcomes, diagnosing complex conditions from CT or MRI, and even assisting in surgeries such as pedicle screw placement, fostering personalized medicine^{39–41}. More specifically, ML has shown incredible promise in the chronic pain literature. For instance, ML has shown promise in pain assessment by approximating self-reported pain scores based solely on physiological features, which could lead to more holistic, multidimensional approaches to pain assessment compared to conventional self-reporting methods^{17,18}. Researchers have predicted individual chronic pain severity scores via intracranial orbitofrontal cortex signals²¹. Our findings confirm and extend the work of Gram et al., who showed ML based on EEG signals could distinguish between responders and nonresponders to morphine, even when conventional statistics reveal no significant differences between the two groups⁴². Levitt et al. trained an SVM to distinguish between healthy controls, patients with chronic lumbar radiculopathy, and those with treatment-resistant chronic back pain who

EEG feature	Rank (PCA)
Baseline theta peak frequency C4	1
HD alpha peak frequency C4	2
Baseline alpha peak frequency C4	3
HD relative global power F4	4
Tonic relative gamma power F3	5
Tonic relative alpha power F3	6
Tonic relative global power T4	7
Tonic relative global power F3	8
Tonic relative gamma power C3	9
Tonic relative global power C3	10
Tonic alpha peak frequency C4	11
Tonic relative global power FP1	12

Table 2. Top-ranked EEG features based on PCA. *PCA* principal component analysis, *HD* high-density stimulation.

Model	EEG features (<i>n</i>)	Accuracy (%)	F1 Score	AUC-ROC
Decision Tree	12	88.2	0.857	0.879
Random Forest	15	35.3	0.154	0.700
XGBoost	5	17.6	0	0.857
Logistic Regression	19	35.3	0	0.843
SVM	5	58.8	0	0.643

Table 3. Performance metrics of predictive models: responders vs. Nonresponders. *AUC-ROC* area under the curve of receiver operating characteristic, *SVM* support vector machine, *n* optimal number of EEG features.

were considered candidates for SCS by a healthcare committee. Their model predicted clinical decisions for SCS implantation, thus highlighting the relevance of EEG features for assessing patient suitability for SCS⁴³.

Interpretability is essential when applying ML to clinical decision-making. Beyond showcasing how responders can be predicted with great accuracy, this work highlights which EEG features out of a relatively large set of options are the most significant, and we show this is consistent across a large cohort of patients in the intraoperative setting. Very few studies have investigated EEG differences across baseline, tonic, and HD stimulation^{8,9}, and none have applied ML to successfully predict which patients will be responders to SCS. The twelve most discriminative features identified in this study, selected through PCA, hold significant promise for future application in distinguishing responders from nonresponders.

The decision tree’s highly accurate performance for this task is reasonable given the lack of statistically significant differences in individual features between the two groups. Decision trees, like XGBoost and Random forest classifiers, excel at handling non-linear relationships. These three tree-based models had better performance than the other architectures, which may have struggled with the complex relationships between the EEG features we extracted. Without extensive kernel tuning, the SVM may have struggled with the non-linear separability. Furthermore, while the dataset of 20 patients is immense in the context of intraoperative EEG studies⁹, it is relatively small for ML problems. This likely gave the decision tree an edge over XGBoost, which is a more complex model that could be prone to overfitting.

At present, SCS device settings are manually tuned for several months after surgery in a trial-and-error based manner, and patients cannot know whether they will be a responder until after this process is complete. Intraoperative EEG data provides a dynamic and real-time insight into the patient’s neurophysiological state, free from potential confounds related to cognitive conditions. A recent study used intraoperative frontal EEG signatures to predict postoperative delirium⁴⁴. This underlines the potential of this modality to predict surgical outcomes.

In the surgical context, both percutaneous and paddle SCS leads offer distinct advantages and disadvantages. The existing SCS literature includes numerous clinical trials reporting successful outcomes with both implantation techniques (laminectomy vs. percutaneous)^{45,46}. However, studies directly comparing these approaches in terms of patient outcomes remain limited⁴⁷. One study reported that patients who underwent laminectomy for paddle lead placement had a higher rate of successful outcomes at a 1.9-year average follow-up, but this advantage was not observed at the 2.9-year follow-up, where success rates were similar between the two groups⁴⁸. A retrospective review demonstrated significant pain relief for both groups based on visual analog scale (VAS) scores within an 8.6- to 10.3-month postoperative period while long-term follow-up (up to 66 months) revealed significantly greater pain relief in patients who underwent laminectomy compared to those who had percutaneous leads⁴⁹. In our dataset, the distribution of percutaneous versus paddle leads was comparable between responders and nonresponders: 3 vs. 4 in responders and 5 vs. 5 in nonresponders, respectively. Using a 50% reduction in NRS

scores as the threshold for response, the implantation technique (percutaneous vs. paddle) did not significantly affect patient outcomes at the 3-month follow-up. This suggests that while certain studies indicate potential advantages for paddle leads in long-term pain relief, the short-term outcomes in our cohort were independent of the surgical method.

Limitations and future directions

Some limitations should be considered when interpreting our results. To be reliably integrated into clinical practice, ML algorithms need to be tested across multiple institutions and in larger cohorts. Although our model's high performance across a variety of metrics adds credibility, external validation remains a critical step for broader application.

In our study, we focused on a 3-month follow-up as this timeframe is clinically critical for optimizing SCS settings. Early response to SCS often guides the subsequent parameter adjustments, which are crucial for long-term management of chronic pain. Several pivotal trials also use a primary evaluation at 3 months, underscoring the importance of this period in determining initial efficacy and patient-specific adjustments. For example, EVOKE study conducted their primary analysis at 3 months to assess the effectiveness of closed-loop SCS⁵⁰. Deer et al. similarly assessed the efficacy of dorsal root ganglion stimulation at this interval⁵¹. However, future studies should follow patients for longer periods of time after surgery and attempt to predict 12- or even 24-month outcomes.

Although our sample size is relatively high compared to other recent intraoperative SCS studies^{52–55}, we acknowledge that a multi-center study with a larger sample size would further strengthen our findings. Specifically, a study with more patients across different etiologies could help test the generalizability of our algorithms and explore how varying etiologies influence responses to different SCS waveforms. Such a follow-up study would provide a broader base for understanding how these treatments impact diverse patient populations and could ultimately enhance the clinical utility of our model. Similarly, in our study, the average age and female-male ratios were comparable between groups, with no significant differences. While age and sex might influence SCS-related EEG changes, their comparable distribution likely minimizes such effects. Thus, strict age- and sex-matching may not be essential; however, future studies with larger sample sizes could consider strict matching to ensure fully unbiased analyses.

Finally, intraoperative mapping (IOM) poses unique challenges and opportunities for EEG recordings, impacting their accuracy and interpretation. In our study, IOM was performed after positioning the SCS lead according to routine surgical procedures. EEG testing began after lead and IPG implantation, following surgical closure, approximately 15 min before surgery concluded. To ensure consistency, stimulation was delivered through the “sweet spot” contacts identified during the trial period, which provided >50% pain relief and were also used for laterality testing. Motor mapping involved evoked responses from 18 muscles (9 per side), alongside somatosensory evoked potentials (SSEPs). While the specific stimulation site along the dorsal column may affect neural responses⁵⁴, our findings reflect stimulation consistently delivered at a clinically validated location, ensuring reliability and relevance.

Future research should focus on exploring preoperative neural signatures to predict surgical outcomes, aiding in better patient selection and enhancing the efficacy of neuromodulation therapy. There is potential to apply the present study's predictive model in the preoperative setting, allowing for a more informed selection of candidates for SCS before patients enter the operating room. Leveraging preoperative data could reduce unnecessary procedures and help identify patients who are most likely to benefit from SCS. This approach aligns with the growing trend of using AI for predictive modeling in medicine to avoid unnecessary invasive procedures or surgeries. The effective application of machine learning algorithms to the clinical settings has the potential to bring medicine into an era of more personalized, precise, and accessible care.

Conclusion

Neural signatures in response to various SCS waveforms were extracted in chronic pain patients who were undergoing SCS implant surgery for their lower back/leg pain. Using clinically recognized outcome measures and the objective EEG markers, ML algorithms were developed to predict responders to SCS therapy. A decision tree model with 12 selected spatio-spectral features performed best and predicted the responders with 88% accuracy. Our research offers insights into the basic science behind chronic pain, identifying specific EEG patterns correlating with pain relief and bridges the gap by applying ML to EEG signals from chronic pain patients undergoing SCS surgery. Our novel approach provides a potentially powerful tool for better patient selection, with the potential to improve chronic pain management.

Data availability

The data analyzed during the current study will be available from the corresponding author upon reasonable request.

Received: 22 March 2024; Accepted: 25 February 2025

Published online: 18 March 2025

References

1. Dahlhamer, J. et al. Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. *Mmwr-Morbidity and Mortality* **67**, 1001–1006. <https://doi.org/10.15585/mmwr.mm6736a2> (2018).
2. Gee, L. et al. Spinal cord stimulation for the treatment of chronic pain reduces opioid use and results in superior clinical outcomes when used without opioids. *Neurosurgery* **84**, 217–226. <https://doi.org/10.1093/neuros/nyy065> (2019).

3. Goudman, L. et al. Patient selection for spinal cord stimulation in treatment of pain: sequential decision-making model—a narrative review. *J. Pain Res.* **15**, 1163–1171. <https://doi.org/10.2147/JPR.S250455> (2022).
4. Tracey, I., Woolf, C. J. & Andrews, N. A. Composite pain biomarker signatures for objective assessment and effective treatment. *Neuron* **101**, 783–800. <https://doi.org/10.1016/j.neuron.2019.02.019> (2019).
5. Nunez, P. L. & Srinivasan, R. *Electric Fields of the Brain: the Neurophysics of EEG* (Oxford University Press, 2006).
6. Lu, L. et al. Intraoperative neurophysiological monitoring during lead placement for dorsal root ganglion stimulation: A literature review and case series. *Neuromodulation* **27**, 160–171. <https://doi.org/10.1016/j.neurom.2023.04.468> (2024).
7. Lichtner, G. et al. Nociceptive activation in spinal cord and brain persists during deep general anaesthesia. *Br. J. Anaesth.* **121**, 291–302. <https://doi.org/10.1016/j.bja.2018.03.031> (2018).
8. Telkes, I. et al. Corrigendum to Differences in EEG patterns between tonic and high frequency spinal cord stimulation in chronic pain patients. *Clin. Neurophysiol.* **131** (8), 1731–1740 <https://doi.org/10.1016/j.clinph.2020.10.001> (2020).
9. Witjes, B., Ottenheim, L. A., Huygen, F. & de Vos, C. C. A review of effects of spinal cord stimulation on spectral features in resting-state electroencephalography. *Neuromodulation* **26**, 35–42. <https://doi.org/10.1016/j.neurom.2022.04.036> (2023).
10. Jensen, M. P., Karoly, P. & Braver, S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* **27**, 117–126. [https://doi.org/10.1016/0304-3959\(86\)90228-9](https://doi.org/10.1016/0304-3959(86)90228-9) (1986).
11. Kremer, E., Atkinson, H. J. & Ignelzi, R. J. Measurement of pain: patient preference does not confound pain measurement. *Pain* **10**, 241–248. [https://doi.org/10.1016/0304-3959\(81\)90199-8](https://doi.org/10.1016/0304-3959(81)90199-8) (1981).
12. Pilitsis, J. G., Fahey, M., Custozzo, A., Chakravarthy, K. & Capobianco, R. Composite score is a better reflection of patient response to chronic pain therapy compared with pain intensity alone. *Neuromodulation* **24**, 68–75. <https://doi.org/10.1111/ner.13212> (2021).
13. Richardson, P. H. & de, C. W. A. C. & What does the BDI measure in chronic pain? *Pain* **55**, 259–266. [https://doi.org/10.1016/0304-3959\(93\)90155-1](https://doi.org/10.1016/0304-3959(93)90155-1) (1993).
14. Vianin, M. Psychometric properties and clinical usefulness of the Oswestry disability index. *J. Chiropr. Med.* **7**, 161–163. <https://doi.org/10.1016/j.jcm.2008.07.001> (2008).
15. Main, C. J. Pain assessment in context: a state of the science review of the McGill pain questionnaire 40 years on. *Pain* **157**, 1387–1399. <https://doi.org/10.1097/j.pain.0000000000000457> (2016).
16. Sullivan, M. J. L., Bishop, S. R. & Pivik, J. The pain catastrophizing scale: development and validation. *Psychol. Assess.* **7**, 524–532. <https://doi.org/10.1037/1040-3590.7.4.524> (1995).
17. Moscato, S., Orlandi, S., Giannelli, A., Ostan, R. & Chiari, L. Automatic pain assessment on cancer patients using physiological signals recorded in real-world contexts. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2022**, 1931–1934. <https://doi.org/10.1109/EMBC48229.2022.9871990> (2022).
18. Pouroumran, F., Radhakrishnan, S. & Kamarthi, S. Exploration of physiological sensors, features, and machine learning models for pain intensity Estimation. *PLoS One* **16**, e0254108. <https://doi.org/10.1371/journal.pone.0254108> (2021).
19. Hadanny, A. et al. Development of machine learning-based models to predict treatment response to spinal cord stimulation. *Neurosurgery* **90**, 523–532. <https://doi.org/10.1227/neu.0000000000001855> (2022).
20. Goudman, L. et al. Predicting the response of high frequency spinal cord stimulation in patients with failed back surgery syndrome: A retrospective study with machine learning techniques. *J. Clin. Med.* **9** <https://doi.org/10.3390/jcm9124131> (2020).
21. Shirvalkar, P. et al. First-in-human prediction of chronic pain state using intracranial neural biomarkers. *Nat. Neurosci.* **26**, 1090–1099. <https://doi.org/10.1038/s41593-023-01338-z> (2023).
22. Dworkin, R. H. et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J. Pain* **9**, 105–121. <https://doi.org/10.1016/j.jpain.2007.09.005> (2008).
23. Roth, S. G. et al. A prospective study of the intra- and postoperative efficacy of intraoperative neuromonitoring in spinal cord stimulation. *Stereotact. Funct. Neurosurg.* **93**, 348–354. <https://doi.org/10.1159/000437388> (2015).
24. Sheldon, B. et al. Spinal cord stimulation programming: a crash course. *Neurosurg. Rev.* **44**, 709–772. <https://doi.org/10.1007/s10143-020-01299-y> (2021).
25. Telkes, I. et al. High-resolution spinal motor mapping using thoracic spinal cord stimulation in patients with chronic pain. *Neurosurg* **91** (3), 459–469 <https://doi.org/10.1227/neu.0000000000002054> (2022).
26. Air, E. L., Toczyl, G. R. & Mandybur, G. T. Electrophysiologic monitoring for placement of laminectomy leads for spinal cord stimulation under general anesthesia. *Neuromodulation* **15** (6), 573–579. <https://doi.org/10.1111/j.1525-1403.2012.00475> 579–580 (2012).
27. Cohen, M. X. *Issues in Clinical and Cognitive Neuropsychology 1 Online Resource* 578 (The MIT Press, 2014).
28. Hayes, M. H. *Statistical Digital Signal Processing and Modeling* (Wiley, 1996).
29. Telkes, I., Jimenez-Shahed, J., Viswanathan, A., Abosch, A. & Ince, N. F. Prediction of STN-DBS electrode implantation track in Parkinson's disease by using local field potentials. *Front. Neurosci.* **10**, 198. <https://doi.org/10.3389/fnins.2016.00198> (2016).
30. Jolliffe, I. T. In *Principal Component Analysis* 338–372 (Springer, 2022).
31. Madan Somvanshi, P. C. & Tambade, S. Swati V. Shinde. in *2016 International Conference on Computing Communication Control and automation (ICCCUBEA)*. (IEEE, 2017).
32. Vapnik, C. C. V. *Machine Learning*. **20**, 273–297 (Kluwer Academic Publishers, 1995).
33. Chen, T. C. G. *22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794.
34. Breiman, L. *Machine Learning* Vol. 45 (ed Robert E. Schapire) 5–32 (Kluwer Academic Publishers, 2001).
35. Petersen, E. A., Schatman, M. E., Sayed, D. & Deer, T. Persistent spinal pain syndrome: new terminology for a new era. *J. Pain Res.* **14**, 1627–1630. <https://doi.org/10.2147/JPR.S320923> (2021).
36. Wasan, A. D. et al. Neural correlates of chronic low back pain measured by arterial spin labeling. *Anesthesiology* **115**, 364–374. <https://doi.org/10.1097/ALN.0b013e318220e880> (2011).
37. Sheng, J., Liu, S., Wang, Y., Cui, R. & Zhang, X. The link between depression and chronic pain: neural mechanisms in the brain. *Neural Plast.* **2017** (9724371). <https://doi.org/10.1155/2017/9724371> (2017).
38. King, C. D. & Sibille, A. K. K. T. *Stress: Concepts, Cognition, Emotion, and Behavior* 1 (ed George Fink) 413–421 (Academic Press, 2016).
39. Halicka, M., Wilby, M., Duarte, R. & Brown, C. Predicting patient-reported outcomes following lumbar spine surgery: development and external validation of multivariable prediction models. *BMC Musculoskelet. Disord.* **24**, 333. <https://doi.org/10.1186/s12891-023-06446-2> (2023).
40. Zhou, S. et al. The application of artificial intelligence in spine surgery. *Front. Surg.* **9**, 885599. <https://doi.org/10.3389/fsurg.2022.885599> (2022).
41. Saravi, B. et al. Artificial intelligence-driven prediction modeling and decision making in spine surgery using hybrid machine learning models. *J. Pers. Med.* **12** <https://doi.org/10.3390/jpm12040509> (2022).
42. Gram, M., Graversen, C., Olesen, A. E. & Drewes, A. M. Machine learning on encephalographic activity May predict opioid analgesia. *Eur. J. Pain* **19**, 1552–1561. <https://doi.org/10.1002/ejp.734> (2015).
43. Levitt, J. et al. Pain phenotypes classified by machine learning using electroencephalography features. *Neuroimage* **223**, 117256. <https://doi.org/10.1016/j.neuroimage.2020.117256> (2020).
44. Rohr, V. et al. Machine-learning model predicting postoperative delirium in older patients using intraoperative frontal electroencephalographic signatures. *Front. Aging Neurosci.* **14**, 911088. <https://doi.org/10.3389/fnagi.2022.911088> (2022).

45. Rock, A. K., Truong, H., Park, Y. L. & Pilitsis, J. G. Spinal cord stimulation. *Neurosurg. Clin. N Am.* **30**, 169–194. <https://doi.org/10.1016/j.nec.2018.12.003> (2019).
46. Fontaine, D. Spinal cord stimulation for neuropathic pain. *Rev. Neurol. (Paris)* **177**, 838–842. <https://doi.org/10.1016/j.neurol.2021.07.014> (2021).
47. Van Buyten, J. P. et al. High-frequency spinal cord stimulation for the treatment of chronic back pain patients: results of a prospective multicenter European clinical study. *Neuromodulation* **16**, 59–65. <https://doi.org/10.1111/ner.12006> (2013).
48. North, R. B., Kidd, D. H., Petrucci, L. & Dorsi, M. J. Spinal cord stimulation electrode design: a prospective, randomized, controlled trial comparing percutaneous with laminectomy electrodes: part II—clinical outcomes. *Neurosurgery* **57**, 990–996 (2005).
49. Villavicencio, A. T., Leveque, J. C., Rubin, L., Bulsara, K. & Gorecki, J. P. Laminectomy versus percutaneous electrode placement for spinal cord stimulation. *Neurosurgery* **46**, 399–405 (2000).
50. Mekhail, N. et al. Long-term safety and efficacy of closed-loop spinal cord stimulation to treat chronic back and leg pain (Evoke): a double-blind, randomised, controlled trial. *Lancet Neurol.* **19** (2), 123–134. [https://doi.org/10.1016/S1474-4422\(19\)30414-4](https://doi.org/10.1016/S1474-4422(19)30414-4) (2020).
51. Deer, T. R. et al. Dorsal root ganglion stimulation yielded higher treatment success rate for complex regional pain syndrome and causalgia at 3 and 12 months: a randomized comparative trial. *Pain* **158** (4), 669–681. <https://doi.org/10.1097/j.pain.0000000000000814> (2017).
52. Falowski, S. M., Benison, A. M. & Nanivadekar, A. C. Regional coverage differences with single-and multi-area burst spinal cord stimulation for treatment of chronic pain. *Neuromodulation* **26** (7), 1471–7 (2023).
53. Hamm-Faber, T. E., Gültuna, I., van Gorp, E. J. & Aukes, H. High-dose spinal cord stimulation for treatment of chronic low back pain and leg pain in patients with FBSS, 12-month results: a prospective pilot study. *Neuromodulation* **23** (1), 118–125 (2020).
54. Berwal, D. et al. Investigation of the intraoperative cortical responses to spinal motor mapping in a patient with chronic pain. *J. Neurophysiol.* **130** (3), 768–774. <https://doi.org/10.1152/jn.00221.2023> (2023).
55. Taghlabi, K. M. et al. Spinal cord stimulation for chronic pain treatment following sacral Chordoma resection: illustrative case. *J. Neurosurg. Case Lessons* **25** (26), CASE23540. <https://doi.org/10.3171/CASE23540> (2023).

Author contributions

I.T. and J.G. conceptualized the study, conducted the analyses, and wrote the manuscript. J.G.P. and T.H. performed the surgeries, contributed to data collection, and interpretation of the results. J.B. contributed to clinical data collection. S.P., R.G., M.B., and K.M. performed the intraoperative neuromonitoring during the surgeries and contributed to data collection. All authors reviewed the manuscript and approved the final version of the manuscript.

Funding

This study was supported by the National Institutes of Health (NIH) award number R00NS119672. Dr. Telkes has grant support from NIH R00NS119672, FAU COECS/I-SENSE, and University of Arizona TRIF Funding. Dr. Pilitsis receives grant support from Medtronic, Boston Scientific, Abbott, Focused Ultrasound Foundation, NIH 2R01CA166379, NIH R01EB030324, and NIH-NeuroBlueprint MedTech 5U54EB033650.

Declarations

Competing interests

Dr. Pilitsis is the medical advisor for Aim Medical Robotics and has stock equity. Steven Panicioli, Rachael Grey, Michael Briotte, and Kevin McCarthy are employees of Nuvasive Clinical Services. The remaining authors declare no competing interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-92111-8>.

Correspondence and requests for materials should be addressed to I.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025