Adapting SPaRCNet for Automated Seizure Detection in Pediatric ECMO Patients

by

Natalie Chen Barnouw

B.S. Computer Science and Molecular Biology, MIT, 2025.

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN COMPUTER SCIENCE AND MOLECULAR BIOLOGY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2026

Authored by: Natalie Chen Barnouw
Department of Electrical Engineering and Computer Science
May 15, 2026

Certified by: Mauro Caffarelli, MD
Assistant Professor of Neurology, UCSF, Thesis Supervisor

Certified by: Rohit Jain, PhD
Pediatric Critical Care EEG Lab Head of Data Science, UCSF, Thesis Supervisor

Certified by: Peter Szolovits
Professor, Massachusetts Institute of Technology, Thesis Supervisor

Accepted by: Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Adapting SPaRCNet for Automated Seizure Detection in Pediatric ECMO Patients

by

Natalie Chen Barnouw

Submitted to the Department of Electrical Engineering and Computer Science
on May 15, 2026 in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN COMPUTER SCIENCE AND MOLECULAR BIOLOGY

ABSTRACT

Electrographic seizures are common in critically ill children and are particularly frequent in high-risk pediatric intensive care unit (PICU) populations, where they may reflect both acute symptomatic seizures and underlying epilepsy[1]. Continuous EEG monitoring improves seizure detection but requires scarce expert interpreters, motivating automated tools such as SPaRCNet (Seizures, Periodic and Rhythmic Continuum patterns Deep Neural Network), a state-of-the-art deep neural network trained on clinical EEG recordings from patients with epilepsy or critical illness to detect seizures and related ictal-interictal-injury continuum patterns [2,3]. However, pediatric EEG exhibits age-dependent background and seizure patterns that may limit the direct applicability of seizure-detection models trained primarily on adult ICU cohorts in children on ECMO. In this project, we evaluated SPaRCNet for seizure detection in a cohort of 209 EEG runs from 44 pediatric patients on ECMO. Using a pipeline that applies SPaRCNet to long-term pediatric EEG, summarizes its segment-level probabilities over time, and links the outputs to clinically annotated seizure labels, I quantified both segment- and patient-level performance. In this pediatric ECMO dataset, a SPaRCNet-derived seizure burden metric defined as the area under the curve but above a given threshold (AUC/AT) showed limited discrimination and predictive value for seizures, yielding a maximum Youden's J value of 0.27 across the datasets tested. Logistic regression showed statistically insignificant predictive ability, further highlighting the need for pediatric-specific preprocessing and model adaptation. The future development of seizure-detection models trained specifically for pediatric patients will allow for real-time neurocritical-care decision support in the pediatric intensive care unit.

Thesis supervisor: Mauro Caffarelli, MD
Title: Assistant Professor of Neurology, UCSF

Thesis supervisor: Rohit Jain, PhD
Title: Pediatric Critical Care EEG Lab Head of Data Science, UCSF

Thesis supervisor: Peter Szolovits
Title: Professor, Massachusetts Institute of Technology

Acknowledgments

First and foremost, I would like to acknowledge my parents, Faye and Benjamin Barnouw, for supporting me throughout my undergraduate studies and masters degree. Thank you for always picking up the phone, providing endless emotional support, and sacrificing so much throughout the years for my education. Thank you also to my brother, Nick, for being so supportive with my academic career and always being a joy to catch up with. To my partner Andrew, thank you for providing endless emotional and moral support, answering countless technical questions, and for working alongside me during those many late nights.

Thank you to Dr. Mauro Caffarelli for taking me under his wing and teaching me about neurology and quantitative methods to detect stroke. I am also very grateful to Dr. Rohit Jain for meeting with me weekly to work through this project and pushing me to become a better researcher. Last but not least, thank you to Dr. Peter Szolovits for agreeing to co-advise my project and graciously supporting me throughout my MEng.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Seizures are among the most common serious neurologic conditions in childhood, affecting about 1-2% of the pediatric population in the United States[4]. In the pediatric intensive care unit (PICU), electrographic-only seizures are particularly common among high-risk children, with observational studies estimating that approximately 20-40% of comatose pediatric ICU patients with neurologic risk factors have seizures detectable only on EEG [2]. Extracorporeal membrane oxygenation (ECMO) is a life-saving intervention used in the PICU that provides temporary cardiopulmonary support in cases of severe cardiac or respiratory failure.

Figure 1.1: ECMO machine for a patient in the ICU. Reproduced from *Venoarterial ECMO Hemodynamics*, StatPearls Publishing LLC, distributed under the Creative Commons CC BY-NC-ND 4.0 license.

There are two main types of ECMO: veno-arterial (VA) and veno-venous (VV) support. VV ECMO drains and returns blood to the venous system, primarily supporting the lungs. Native cardiac output still provides forward flow in this case. In peripheral VA ECMO, oxygenated blood is pumped back into the arterial system, often the femoral artery, in a retrograde direct toward the aorta. Because this is against the normal physiologic flow, VA

ECMO increases left-ventricular afterload, which may create regions of abnormal shear or stasis. This is thought to contribute to thrombus formation and embolic stroke risk. Due to the critical status of patients who require ECMO, in addition to the non-physiologic flow pattern created by the machine, patients on ECMO are therefore placed at a higher risk of neurological injury, including both stroke (due to blood clotting) and seizure [5].

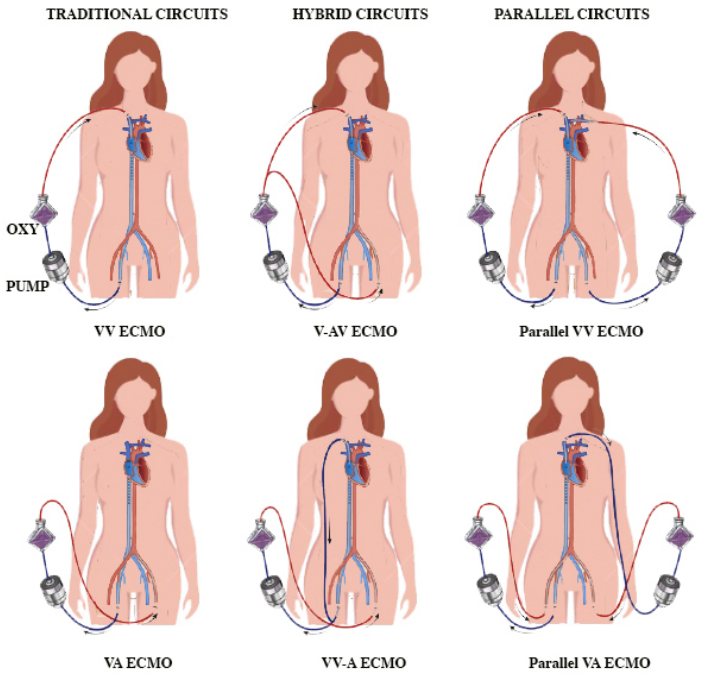


Figure 1.2: VA and VV ECMO circuit configurations, illustrating venous drainage, the extracorporeal pump and oxygenator, and arterial versus venous return. Reproduced from *Venoarterial ECMO Hemodynamics*, StatPearls Publishing LLC, distributed under the Creative Commons CC BY-NC-ND 4.0 license.

Indeed, cerebrovascular injury and electrographic seizures are well-recognized complications of pediatric ECMO. Meta-analyses and cohort studies have reported seizure incidences ranging from 12-18% in children on ECMO or other mechanical circulatory support, increasing to an estimated 18-23% when continuous EEG monitoring is used [6,7]. Seizures are associated with increased mortality and worse neuro-developmental outcomes, and they frequently co-occur with radiographic evidence of ischemic or hemorrhagic brain injury [8]. Timely, accurate seizure recognition in this setting is critical for guiding treatment and allocating limited neuro-critical care resources [9]. These findings have led expert groups to recommend routine or at least targeted EEG monitoring in pediatric ECMO patients, but there is no consensus on optimal monitoring duration or standardized response algorithms [6].

1.2 Deep Learning Methods and SPaRCNet

Despite these recommendations, implementation of care and standard procedure is variable. Continuous EEG is technically demanding, and interpretation requires specialized expertise which may not be available at all hours. Several studies highlight delays between EEG acquisition and formal review, during which seizures may go unrecognized. This context has motivated growing interest in automated EEG analysis pipelines that can support real-time detection of seizure activity and other high-risk patterns in the PICU [10]. Early automated seizure-detection approaches relied on hand-crafted features combined with classical machine learning classifiers such as support vector machines (SVMs) and random forests [11]. The features were hand-crafted from the time, frequency, and time-frequency domains, including band-limited power, Hjorth parameters, short-time fourier transform (STFT) [12,13]. These systems demonstrated reasonable performance in research settings but often required substantial feature engineering, patient-specific tuning, or artifact rejection steps that limited generalizability and real-time deployment [14].

Deep-learning methods have recently supplanted feature-based approaches. Convolutional and recurrent neural networks trained directly on raw or minimally pre-processed EEG can learn hierarchical temporal-spatial features and have achieved state-of-the-art performance on several benchmark datasets [15]. Recent work has explored architectures such as one-dimensional convolutional neural networks (CNNs), hybrid CNN-Long Short-Term Memory (CNN-LSTM) models to better capture short- and long-range temporal patterns, and autoencoder and encoder-decoder architectures for dimensionality reduction before classification [16]. Systematic reviews focused on pediatric epilepsy similarly report high overall accuracy and sensitivity for deep-learning-based seizure detection, although heterogeneity in datasets and evaluation protocols complicates direct comparison across studies [17].

1.2.1 SPaRCNet Model Architecture

Within the critical-care setting, Seizures, Periodic and Rhythmic Continuum patterns Deep Neural Network (SPaRCNet) represents a major step forward. Westover et al. trained SPaRCNet, a Dense-Net type of Convolutional Neural Network (CNN) on 50,697 labeled 10-second EEG segments from 6,095 long-term adult EEG recordings to detect seizures and other ictal-interictal-injury continuum (IIIC) patterns [3].

The model consists are seven dense blocks, each of which include four dense layers. Each dense layer contains two convolutional layers and two exponential linear unit (ELU) activation layers. Between each dense block is a transition block, consisting of an ELU, convolutional, and average pooling layer. The final two layers consist of a fully-connected layer followed by a SoftMax layer to classify the 10-second window as one of the five IIIC patterns or the "other" class.

Segments were labeled by board-certified epileptologists from 18 institutions, reducing single-center bias reflected in IIIC pattern interpretation. Across several IIIC classes, SPaRCNet matched or exceeded the aggregate performance of human experts in terms of Receiver Operating Characteristic (ROC) and precision-recall metrics, and it produced well-calibrated probability estimates. Subsequent commentary has highlighted SPaRCNet as the current standard for automated seizure detection in adult critical-care EEG [10]. Importantly, how-

ever, the training and validation cohorts were predominantly adult patients, and performance in purely pediatric populations - especially those on ECMO - has not been systematically characterized.

1.3 Motivation of Current Work

There is growing recognition that algorithms developed on adult EEG may not transfer directly to children. Pediatric EEG exhibits age-dependent changes in background frequency content, burst-suppression patterns, and transient waveforms, reflecting evolving cortical connectivity and myelination. Quantitative spectral studies show decreasing absolute lower power values in the δ (0.5 - 4 Hz) and θ (4-8 Hz) ranges, relative power increases in the higher α (8-12 Hz) and β (12-30 Hz) bands, and region-specific trajectories of maturation [18,19]. These maturational trends can alter both the baseline against which seizures are detected and the morphology of ictal discharges themselves.

A 2025 review of AI in pediatric epilepsy detection concluded that deep-learning models are promising but highlighted substantial gaps, including small sample sizes, limited external validation, and a lack of studies in critically ill pediatric populations [20]. To date, there are few, if any, published seizure-detection algorithms designed specifically for pediatric ECMO patients. Existing pediatric studies largely focus on epilepsy clinics or general PICU cohorts rather than the unique hemodynamic and artifact environment of ECMO (e.g., pump noise, frequent interventions, sedation regimens). This highlights a need to evaluate the state-of-the-art adult critical-care algorithm (SPaRCNet) on pediatric ECMO EEG, characterize its ability to identify seizure occurrences in this setting, and explore strategies to improve performance and calibration in this high-risk group. In this thesis, I address this gap by applying the SPaRCNet algorithm to EEG segments from a cohort of pediatric patients on ECMO. I quantify segment- and patient-level performance on seizure- and stroke-labeled data to assess SPaRCNet's ability to generalize to the pediatric critical care setting. I also characterize failure modes of SPaRCNet by age, seizure type, and stroke co-occurrence.

Beyond SPaRCNet, newer AI systems such as MORGOTH are being developed for broader EEG anomaly and seizure detection in critical care, and are likely to be even more powerful than existing models[21]. Whereas SPaRCNet focuses on segment-level classification of ictal–interictal-injury continuum patterns, MORGOTH appears to extend toward event-level and EEG-level abnormality detection, making it potentially more aligned with real-world clinical interpretation workflows. However, because these systems are also trained primarily on large adult ICU datasets, their performance and calibration in pediatric EEG remain largely unknown. Assessing the applicability of SPaRCNet in a pediatric critical-care context therefore provides an early indication of how well newer models like MORGOTH may generalize to this population.

Importantly, stroke is a major neurologic complication of pediatric ECMO, with intracranial hemorrhage or cerebral infarction occurring in approximately 7-10% of ECMO patients. Recent work developing the Correlate of Injury to the Nervous System (COIN) index has leveraged frequency-domain characteristics and asymmetrical power differences to identify pediatric stroke, successfully distinguishing children with acute ischemic stroke from controls with $>$90% accuracy at an optimal cutoff [22]. We therefore leverage both

seizure and stroke labels derived from clinical annotations, allowing us to assess SPaRCNet's seizure-detection performance in the context of co-occurring ischemic injury and to explore how pediatric-adapted EEG features might ultimately support broader neurological injury risk stratification at the bedside. Consideration of neurological injury in the context of both stroke and seizure will allow our results to illustrate the potential of EEG-derived biomarkers for anomaly detection in collaboration with deep learning models.

Chapter 2

Methods

2.1 Methods

2.1.1 Data Collection

EEG data was collected from 46 patients at the University of California, San Francisco Benioff Children's Hospital pediatric intensive care unit (PICU). 218 EEG runs were collected in total. This dataset has IRB approval to work with, and patients have been de-identified.

Patient Exclusion

On a patient level, we excluded two neonatal patients (<1 month of age) to focus the analysis on non-neonatal pediatric EEG, because neonatal EEG has distinct background and seizure patterns that require different analytic and interpretive frameworks. On an EEG level, EEG runs were excluded if their duration was less than 2 hours.

2.1.2 Data Processing Pipeline

A full processing pipeline is detailed. All analyses were performed in Python 3.11; the full analysis pipeline and scripts are provided in the accompanying code repository available in the appendix.

SPaRCNet is a temporal deep convolutional neural network for EEG event classification, based on a DenseNet-style architecture. The model produces probabilities for six ictal-interictal-injury continuum (IIIC) classes: seizure, lateralized periodic discharges (LPD), generalized periodic discharges (GPD), lateralized rhythmic delta activities (LRDA), generalized rhythmic delta activities (GRDA), and "other". It was originally trained on 50,697 expert-annotated EEG segments from 6,095 recordings in 2,711 patients with epilepsy or critical illness undergoing EEG monitoring. In this study, we use the published SPaRCNet weights as a fixed classifier and do not modify the core model architecture.

SPaRCNet receives raw EEG data sampled at 512 Hz and windowed into 10-second segments from pediatric ECMO patients. SPaRCNet utilizes a standard clinical montage based on the international 10-20 system. The raw EEG files were run through the SPaRCNet

algorithm to obtain per-segment probabilities for six injury classes across the 10-second segments.

The SPaRCNet output was a n x 6 Pandas DataFrame, with n representing the number of 10-second windowed segments. For each injury class, the area under the probability curve (AUC) that lies above a chosen threshold was computed and subsequently normalized by the recording length (AUC/AT).

These summary features were merged with clinical variables including seizure presence, stroke, and age. From the clinical annotations, the patient was considered to have a seizure if the variable "focalsz-se" was one, while it was labeled as non-seizure if the value was zero. Similarly, the clinical variable "img-fii" indicated if the patient had a stroke, with one as a positive case and zero as a negative case.

To determine the optimal AUC/AT threshold that would allow for the greatest separation between seizure and non-seizure patients, the AUC value was calculated for threshold values ranging from 0.0 to 1.0 in increments of 0.1. Plots were labeled by stroke and seizure presence. This allowed for an initial visual inspection of an optimal threshold value.

Youden's J statistic was used to quantitatively determine the optimal threshold value. Youden's J summarizes the performance of a binary classifier, accounting for both sensitivity and specificity metrics. $J = sensitivity + specificity - 1$, where a score of 1 indicates a perfect test with no false positive or false negative classifications.

For each internal SPaRCNet probability threshold $thr \in \{0, 0.1, \ldots, 1.0\}$, we computed a patient-level AUC/AT summary feature. Treating this feature as a continuous biomarker, we used Youden's J statistic to identify the optimal cutpoint on the AUC/AT scale for discriminating seizure versus non-seizure patients, and recorded the resulting maximum J for each thr. The internal probability threshold thr^* with the highest Youden's J was selected as the optimal integration threshold. All downstream logistic regression analyses used the AUC/AT feature computed at this optimal threshold thr^* as a continuous predictor.

While the AUC/AT value was the primary independent variable used, stroke presence was also used as an independent input variable. For the AUC/AT value, the area under the seizure, LPD, and seizure + LPD curves were all considered in separate logistic regression models. The probability of a clinically-labeled seizure was the dependent variable. For each logistic regression that used an interaction variable, the logistic regression equation produced additive and multiplicative terms:

$$p(seizure) = \frac{1}{1 + e^{-(b_0 + b_1 * x_1)}} \tag{2.1}$$

where x_1 is AUC/AT.

$$p(seizure) = \frac{1}{1 + e^{-(b_0 + b_1 * x_1 + b_2 * x_2 + b_3 * x_1 * x_2)}} \tag{2.2}$$

where x_1 is AUC/AT and x_2 is stroke presence.

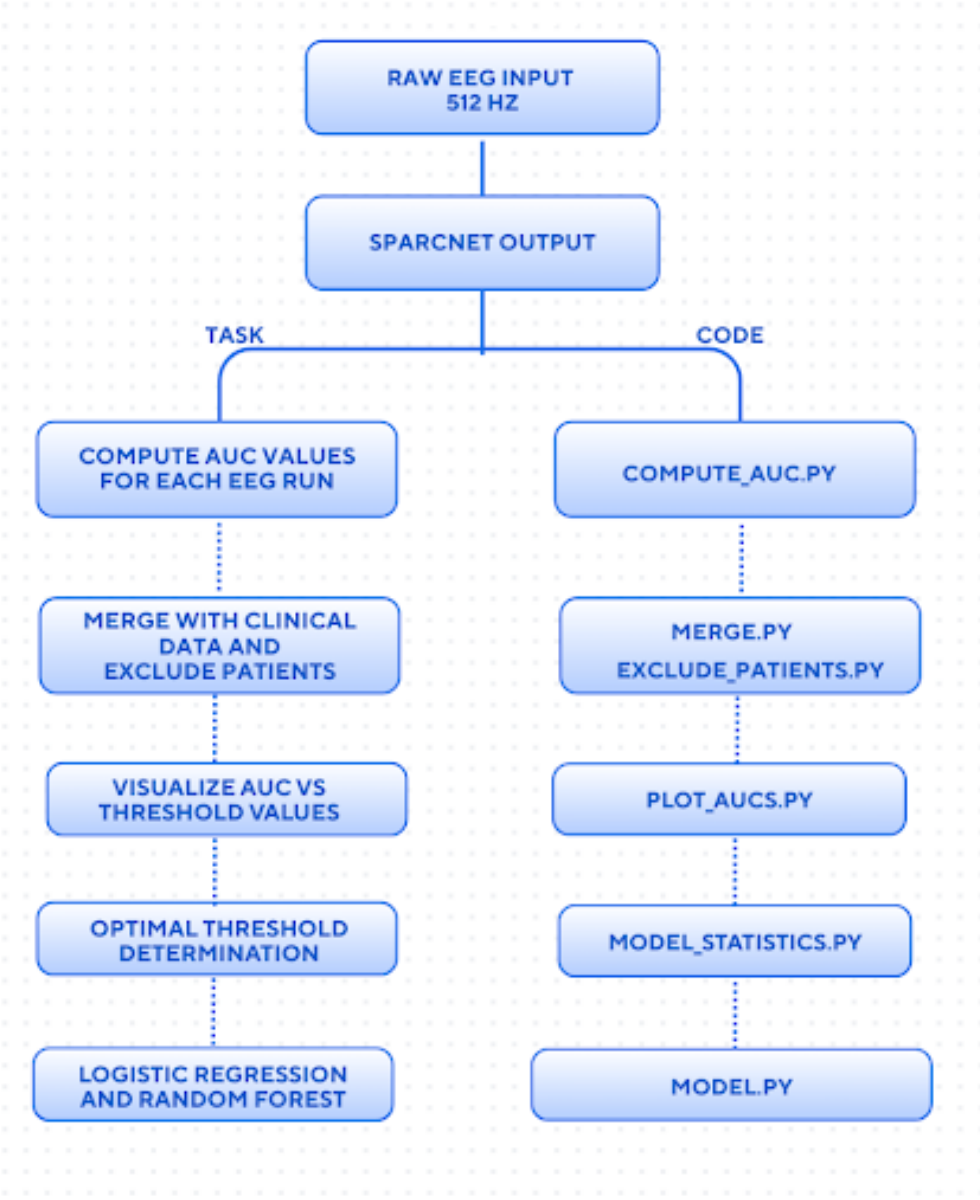


Figure 2.1: Pre-processing pipeline overview with task and accompanying code filename. AUC values were calculated for each EEG run at various thresholds. Results were merged with clinical data. Optimal threshold was determined, followed by logistic regression to evaluate AUC/AT predictive ability.

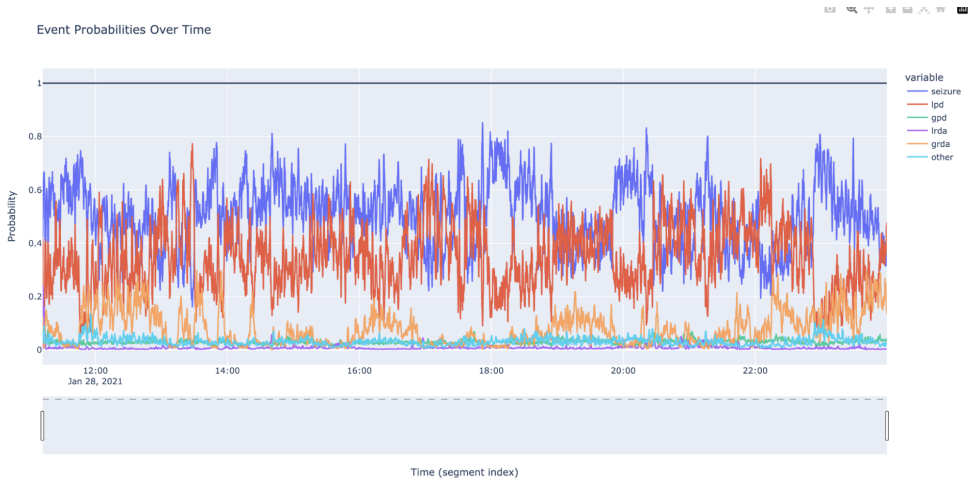


Figure 2.2: Example SPaRCNet output. Probability of each of the five IIIC patterns and "other" (each represented by a different color line) are plotted against time.

Chapter 3

Results

After excluding 2 patients due to neonatal age status from an initial pool of 46 patients, 209 EEG runs from 44 patients were kept for further analysis. Of the 44 patients, 17 patients had a seizure, while 27 patients did not have a seizure during their EEG recording. 19 patients had a stroke, while 25 patients did not have a stroke during their recording.

Table 3.1: Patient characteristics for the pediatric ECMO cohort.

Characteristic	Value
Number of patients	44
Age, mean (years)	3.05
Age, median (years)	0.53
Age, minimum (years)	0.09
Age, maximum (years)	15.16
EEG duration, mean (hours)	17.44
EEG duration, median (hours)	18.68
EEG duration, minimum (hours)	2.27
EEG duration, maximum (hours)	28.79
Patients with seizures, n	17
Patients with seizures, %	38.64
Patients with stroke, n	19
Patients with stroke, %	43.18

To avoid conflating neonatal and pediatric EEG patterns, we restricted the cohort to patients at least 1 month of age; age was therefore not included as a primary predictor in the logistic regression models, but was summarized descriptively and examined in sensitivity analyses. The distribution of seizure AUC/AT values as a function of age is shown in Figure B.2. Across the non-neonatal pediatric range, AUC/AT values did not exhibit a clear monotonic relationship with age. Pearson correlation between age and seizure AUC/AT value was low (Pearson's r = -0.016180, p-value: 0.916975, indicating no strong monotonic association between age and this summary feature.

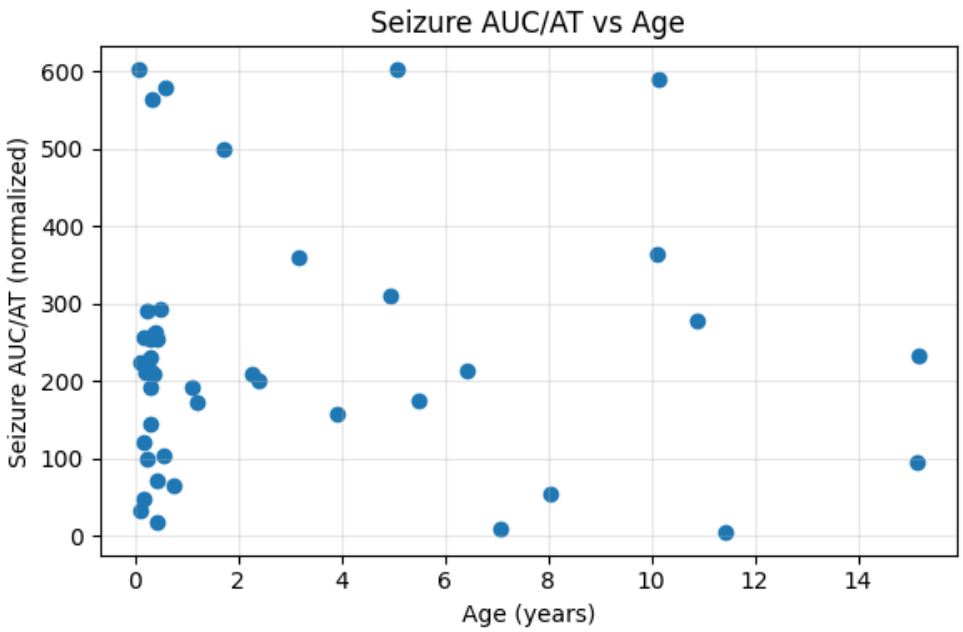


Figure 3.1: Scatter plot of patient-level AUC/AT values plotted against age for the pediatric ECMO cohort. Each point represents a single patient.

3.1 Optimal Threshold Value Determination

Figure 3.2 showed no obvious separation between seizure and non-seizure patients at any threshold for the seizure dataset. There was also not a clear separation based on stroke presence that was indicative of a seizure or non-seizure label. The AUC vs. threshold plots for the LPD and Seizure + LPD datasets are included in the appendix.

For each SPaRCNet-derived feature set (seizure, LPD, and seizure + LPD), the AUC/AT threshold used for downstream logistic regression was selected using Youden's J statistic.

Because the maximum Youden's J values were small, these thresholds were interpreted as exploratory rather than definitive clinical cutoffs.

Table 3.2 shows the maximum J values for each dataset, with a maximum J value of 0.0741 achieved for the seizure dataset. These results indicated that the thresholds are unstable and not strongly meaningful.

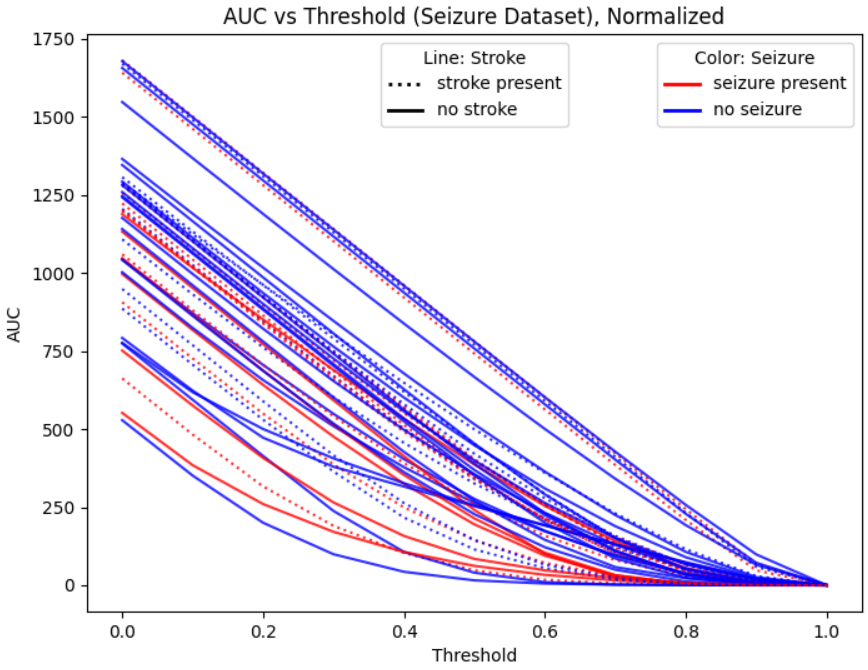


Figure 3.2: Patient-level AUC values plotted against threshold values for seizure dataset. Color indicates seizure presence, while line style (dashed or solid) indicates stroke presence.

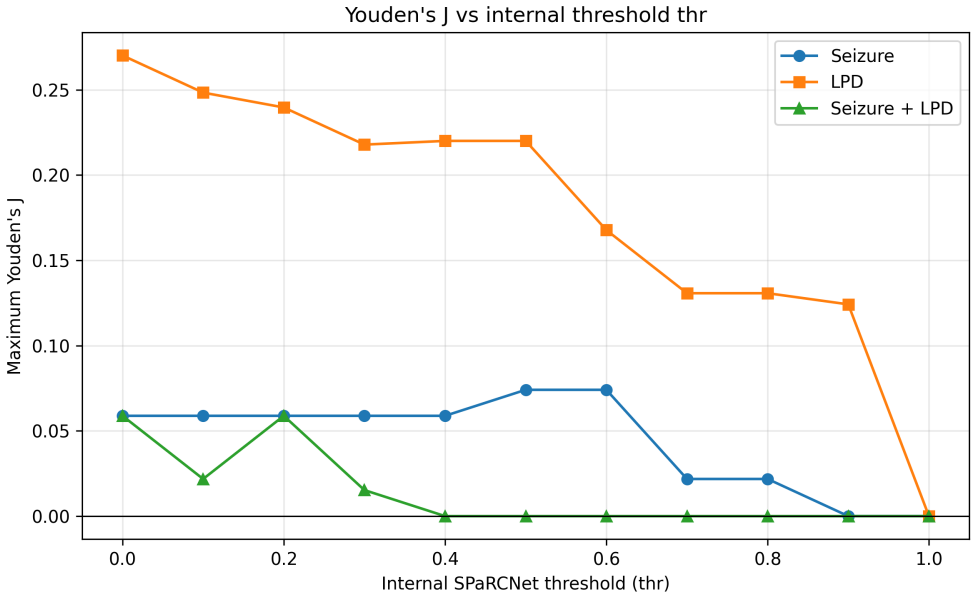


Figure 3.3: Youden's J statistic (sensitivity + specificity - 1) for the seizure dataset across AUC/AT thresholds from 0 to 1. The maximum J value achieved is 0.27, consistent with limited discriminative performance.

Table 3.2: Youden's J across internal thresholds for the seizure dataset.

thr	Max J	Optimal AUC/AT cutpoint	Sensitivity	Specificity
0.0	0.058824	1680.610674	0.058824	1.000000
0.1	0.058824	1500.610674	0.058824	1.000000
0.2	0.058824	1320.610674	0.058824	1.000000
0.3	0.058824	109.678221	0.058824	1.000000
0.4	0.058824	52.326063	0.058824	1.000000
0.5	0.074074	46.480306	1.000000	0.074074
0.6	0.074074	11.466354	1.000000	0.074074
0.7	0.021786	23.218677	0.058824	0.962963
0.8	0.021786	2.998843	0.058824	0.962963
0.9	0.000000	0.032743	1.000000	0.000000
1.0	0.000000	0.000000	1.000000	0.000000

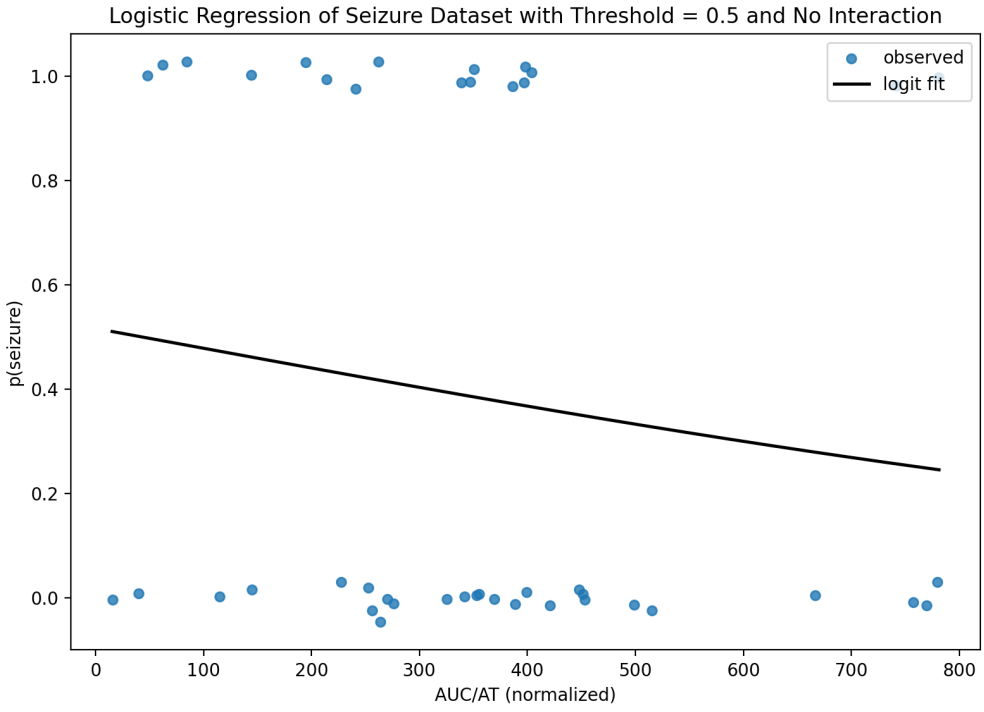
Table 3.3: Maximum Youden's J across internal thresholds for each dataset.

dataset	Seizure	LPD	Seizure + LPD
Threshold value	0.5	0.0	0.0
Max J	0.074074	0.270152	0.058824
Optimal AUC/AT cutpoint	46.213064	663.498109	1613.448717
Sensitivity	1.0	0.529412	0.058823
Specificity	0.074074	0.740741	1.000000

3.2 Logistic regression curves at optimal threshold values

Logistic regression was then performed separately on each dataset using the AUC threshold maximized by the Youden's J statistic. Because the maximum Youden's J values were small, these thresholds were interpreted as exploratory rather than definitive clinical cutoffs. Additionally, with 17 seizure events, the study is underpowered to reliably estimate logistic regression models with multiple predictors or interaction terms. Using the 10 events-per-parameter rule of thumb, only 1-2 parameters can be stably estimated.

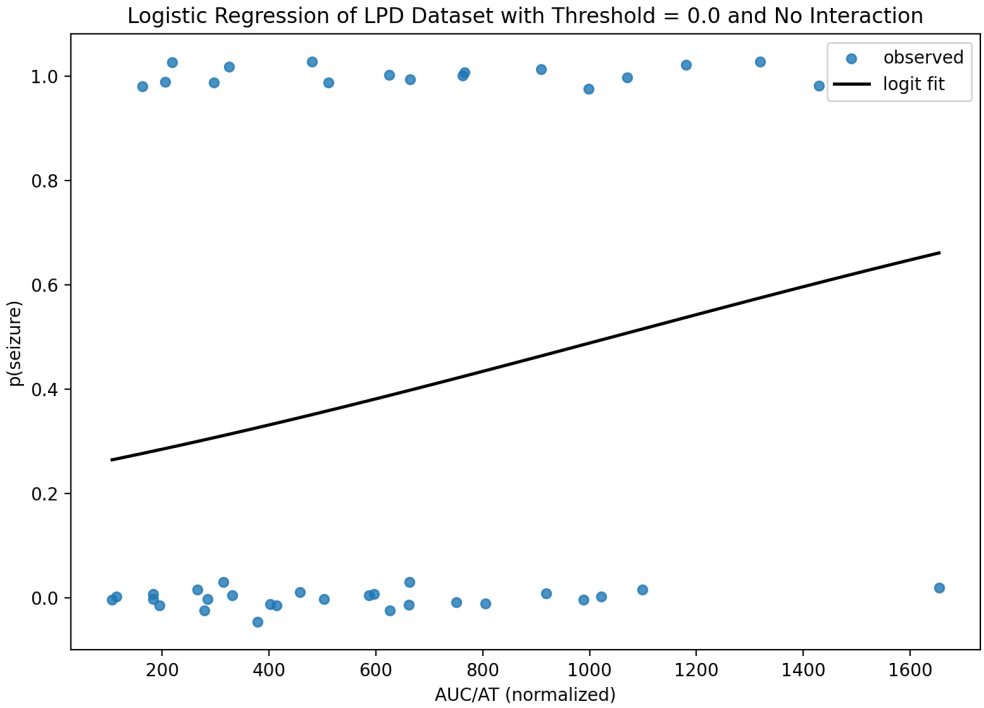
For the seizure dataset, logistic regression was run with a threshold of 0.5. Figure 3.4 shows the model run on the seizure-calculated AUC/AT values without any interaction terms or additional predictive values. The p-value of the constant and AUC/AT terms did not reveal any statistically significant predictive value for seizure presence ($p = 0.92$ and $p = 0.35$ respectively).



term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	0.0636899	-1.17442	1.3018	1.06576	0.308998	3.67591	0.919691
auc_normalized	-0.00152218	-0.00470653	0.00166217	0.998479	0.995305	1.00166	0.34881

Figure 3.4: Predicted probability of clinical seizures from a logistic regression model using seizure AUC/AT as the sole predictor on seizure only dataset. Points show observed patient outcomes, and the curve shows the fitted probability; the wide overlap between seizure and non-seizure patients illustrates the lack of a strong association.

Similarly, logistic regression was performed on the LPD-only dataset with a threshold of 0.0. Figure 3.5 shows that constant term and the AUC/AT predictor variable were insignificant in predicting seizure presence, with p-values of 0.07 and 0.20 respectively.



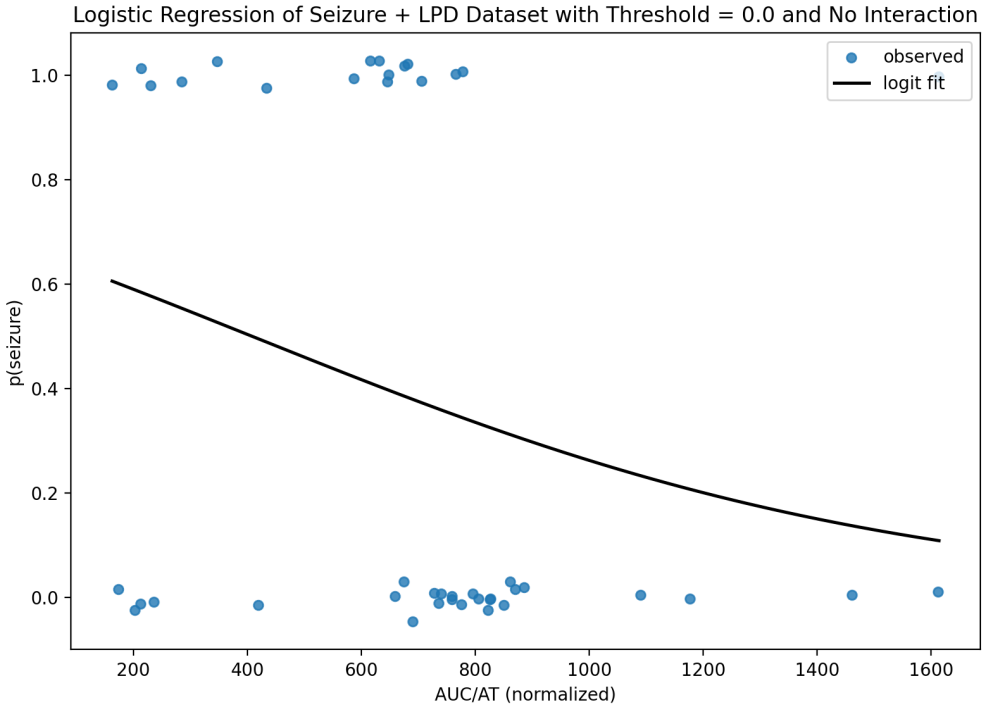
term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	-1.1408	-2.35323	0.0716389	0.319564	0.0950614	1.07427	0.0651604
auc_normalized	0.00109251	-0.00056077	0.00274578	1.00109	0.999439	1.00275	0.195262

Figure 3.5: Predicted probability of clinical seizures from a logistic regression model using seizure AUC/AT as the sole predictor on LPD only dataset. Points show observed patient outcomes, and the curve shows the fitted probability.

Lastly, logistic regression was performed on the seizure + LPD dataset with a threshold of 0.0. Figure 3.6 shows that constant term and the AUC/AT predictor variable were insignificant in predicting seizure presence, with p-values of 0.35 and 0.11 respectively.

We also wanted to explore whether knowledge of stroke presence contributed to the predictive abilities of the logistic regression model. For each logistic regression model, we included binary stroke presence as an independent variable and interaction variable. Figure 3.7 depicts the model run on the seizure-calculated AUC/AT values with stroke as an interaction variable. The p-value of the constant, AUC/AT, stroke, and interaction terms were all insignificant ($p = 0.50$, 0.07, 0.80, and 0.13 respectively).

Figure 3.8 depicts the model run on the LPD-calculated AUC/AT values with stroke as

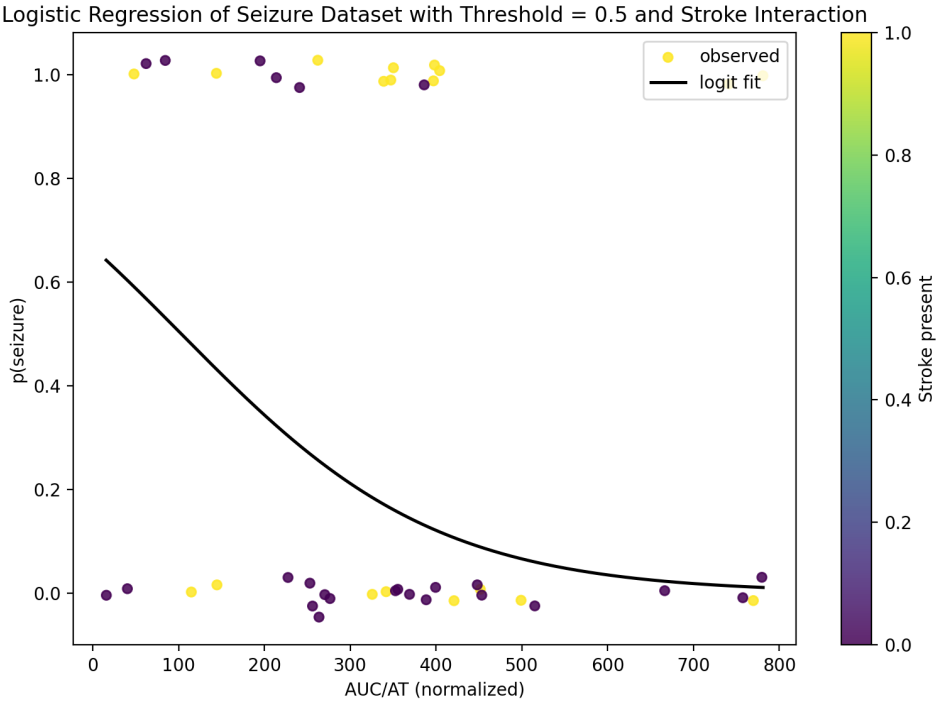


term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	0.712223	-0.803716	2.22816	2.03852	0.447662	9.28279	0.357136
auc_normalized	-0.00174809	-0.00387535	0.000379171	0.998253	0.996132	1.00038	0.107264

Figure 3.6: Predicted probability of clinical seizures from a logistic regression model using seizure AUC/AT as the sole predictor on seizure + LPD dataset. Points show observed patient outcomes, and the curve shows the fitted probability.

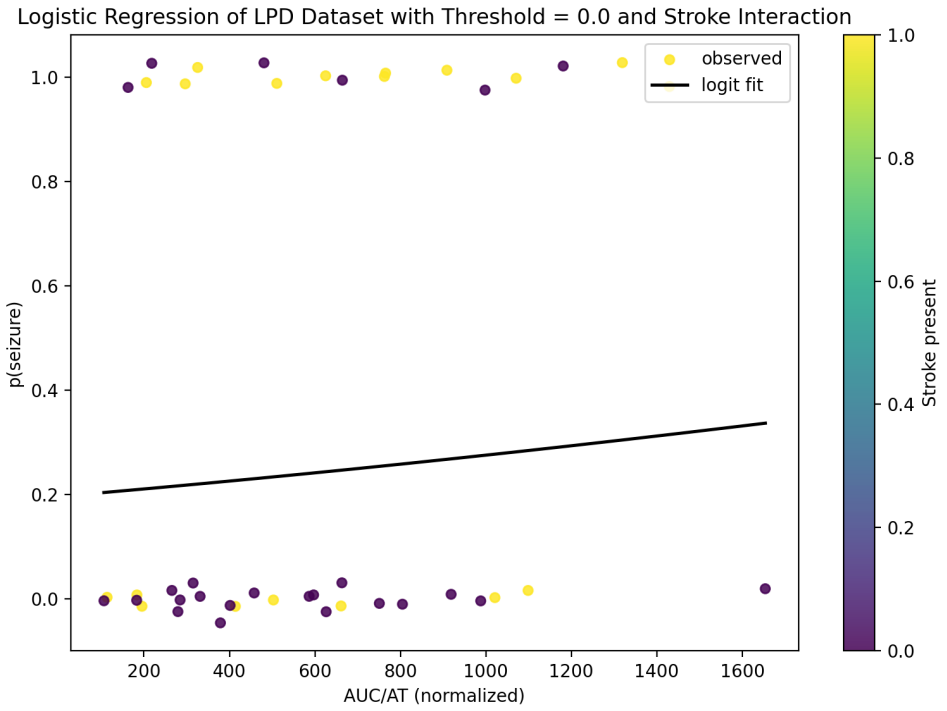
an interaction variable. The p-value of every term was insignificant (p = 0.11, 0.72, 0.57, and 0.53).

Figure 3.9 displays results of the model run on the seizure + LPD calculated AUC/AT values with stroke as an interaction variable. The p-value of each term was insignificant (p = 0.56, 0.10, 0.82, and 0.43).



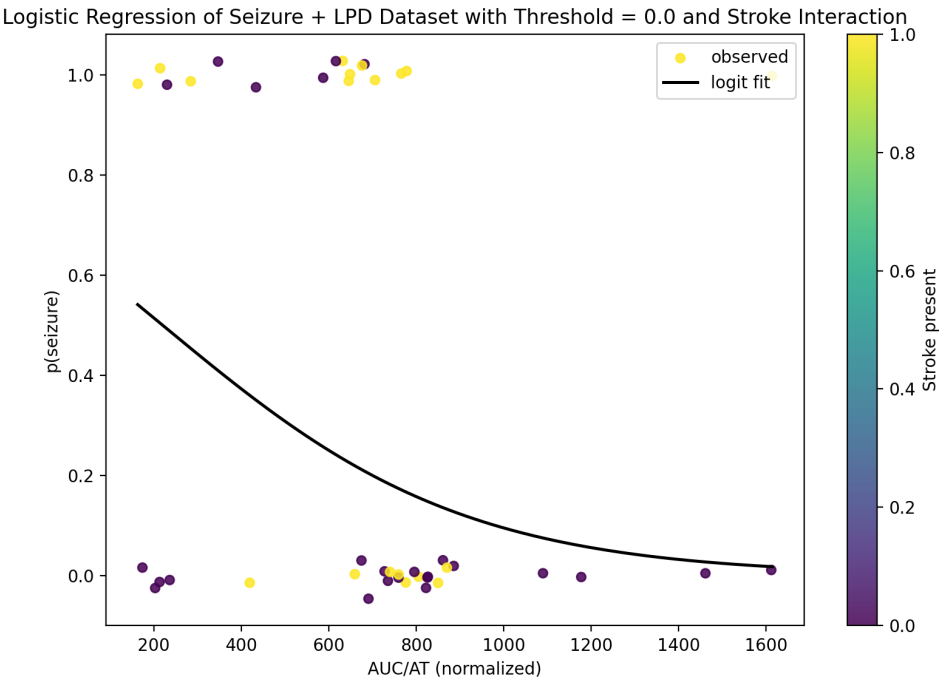
term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	0.685929	-1.32894	2.7008	1.98562	0.264758	14.8916	0.504621
auc_normalized	-0.00668025	-0.0140285	0.000667981	0.993342	0.986069	1.00067	0.0747825
stroke_present	-0.361903	-3.16525	2.44145	0.69635	0.0422035	11.4896	0.800249
interaction	0.00666571	-0.00194951	0.0152809	1.00669	0.998052	1.0154	0.129406

Figure 3.7: Logistic regression model of clinical seizure status in the seizure dataset including AUC/AT (normalized), stroke status, and their interaction as predictors. Points show observed seizure outcomes for individual patients, colored by presence of stroke, and the solid line shows the fitted probability of seizure across AUC/AT values for the reference group. The accompanying table reports coefficient estimates, odds ratios, confidence intervals, and p-values, illustrating that neither AUC/AT nor the stroke interaction term is significantly associated with seizure risk in this cohort.



term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	-1.41251	-3.15108	0.326051	0.24353	0.0428058	1.38549	0.111297
auc_normalized	0.0004429	-0.00201147	0.00289727	1.00044	0.997991	1.0029	0.723577
stroke_present	0.732518	-1.79758	3.26261	2.08031	0.1657	26.1176	0.570407
interaction	0.00114049	-0.00243418	0.00471515	1.00114	0.997569	1.00473	0.53176

Figure 3.8: Logistic regression model of clinical seizure status in the LPD dataset including AUC/AT (normalized), stroke status, and their interaction as predictors. Points show observed seizure outcomes for individual patients, colored by presence of stroke, and the solid line shows the fitted probability of seizure across AUC/AT values for the reference group. The accompanying table reports coefficient estimates, odds ratios, confidence intervals, and p-values, illustrating that neither AUC/AT nor the stroke interaction term is significantly associated with seizure risk in this cohort.



term	coef	CI_low	CI_high	OR	OR_CI_low	OR_CI_high	pval
const	0.634264	-1.4912	2.75973	1.88563	0.225103	15.7955	0.55863
auc_normalized	-0.00289069	-0.00632055	0.000539168	0.997113	0.993699	1.00054	0.0985629
stroke_present	0.367178	-2.8228	3.55716	1.44365	0.059439	35.0634	0.821514
interaction	0.00189972	-0.00277112	0.00657055	1.0019	0.997233	1.00659	0.425361

Figure 3.9: Logistic regression model of clinical seizure status in the seizure + LPD dataset including AUC/AT (normalized), stroke status, and their interaction as predictors. Points show observed seizure outcomes for individual patients, colored by presence of stroke, and the solid line shows the fitted probability of seizure across AUC/AT values for the reference group. The accompanying table reports coefficient estimates, odds ratios, confidence intervals, and p-values, illustrating that neither AUC/AT nor the stroke interaction term is significantly associated with seizure risk in this cohort.

Chapter 4

Discussion and Conclusion

4.1 Discussion

In this pediatric ECMO cohort, SPaRCNet's seizure output summarized as the area under the curve and above a threshold (AUC/AT) was not a significant predictor of clinical seizure labels at either the segment or patient level. Across all three SPaRCNet-derived feature sets (seizure, LPD, and seizure + LPD), the maximum Youden's J values were modest ($J = 0.27$ for LPD and < 0.08 for the seizure and seizure + LPD features), indicating only weak separation between seizure and non-seizure patients even under threshold optimization.

Consistent with these findings, logistic regression models using AUC/AT as a continuous predictor did not show statistically significant associations with clinical seizure status in any dataset. In several models, higher AUC/AT values were actually associated with lower predicted seizure probability (i.e., the regression slope was in the opposite direction to that expected for a useful biomarker), although the corresponding confidence intervals were wide and all p-values exceeded conventional significance thresholds. This pattern further supports the conclusion that, in this cohort, the SPaRCNet-derived features do not provide a stable, interpretable linkage between segment-level network output and patient-level seizure risk.

There are several plausible explanations for the observed performance, as well as limitations that may have affected results. First, SPaRCNet was trained on mostly adult EEG segments from patients with epilepsy or critical illness, whereas pediatric ECMO EEG exhibits different age-dependent background rhythms, artifact profiles, and maturational patterns. The combination of immature EEG features, ECMO-specific noise (including pump noise, sedation, and frequent interventions), and distinct seizure morphologies may cause an adult-trained model to misclassify both seizure and non-seizure segments. Second, the seizure labels were defined at a patient level, indicating if the patient had a seizure at any point during their EEG recording while on ECMO. Consequently, SPaRCNet's segment-level IIIC probabilities are summarized over long recordings. This mismatch between segment-level model outputs and patient-level binary labels likely introduces label noise and reduces the model's power to detect associations. Future evaluations of SPaRCNet in pediatric patient cohorts would benefit from EEG data where seizure and stroke presence are time-labeled, enabling more precise event-level evaluation of SPaRCNet's predictive power. Third, despite a reasonable balance between seizure and non-seizure patients, the sample size of 44 pediatric

ECMO patients still limits statistical power and the stability of model estimates. Small cohorts are particularly challenging for evaluating complex models whose performance may vary substantially across individuals. Finally, the use of a single global threshold to compute AUC/AT, chosen by optimizing the Youden's J value, may be too crude to capture subtle pediatric seizure patterns; age- or context-specific thresholds or alternative temporal aggregation strategies could yield different results. In the future, in addition to logistic regression, other models including random forest should be explored to evaluate AUC/AT or other SPaRCNet-derived metrics as classifiers for seizure detection.

4.2 Indications for the Pediatric ICU

In this pediatric ECMO cohort, a state-of-the-art adult ICU seizure-detection algorithm (SPaRCNet) showed limited ability to distinguish patients with and without clinically identified seizures. These findings indicate that adaptation of seizure-detection models trained and validated exclusively in adult critical-care EEG should be approached cautiously when applied to pediatric populations. An adult-derived model such as SPaRCNet cannot be assumed to perform reliably in pediatric ECMO patients without dedicated pediatric validation and, likely, pediatric-specific retraining or recalibration.

In the clinical setting, misclassification of neurological injury has direct implications for timely care and outcomes. If an automated seizure-detection system has a high false-negative burden, electrographic seizures may go unrecognized in real time, delaying escalation of anti-seizure therapy and potentially prolonging seizure exposure. Conversely, if the false-positive rate is high, frequent spurious seizure alerts can contribute to alarm fatigue and may erode clinicians' trust in the tool. In a busy PICU, neurologists and intensivists may not have the capacity to review every alert in detail, reducing the practical usefulness of a detection model whose outputs are not well-calibrated to pediatric EEG.

Despite these limitations, automated tools such as SPaRCNet for seizure detection and the Correlate of Injury to the Nervous System (COIN) for stroke detection have the potential to support pediatric intensive care unit workflow by making continuous EEG-derived neurological monitoring more accessible to non-expert interpreters. However, their implementation in the PICU should be preceded by pediatric-specific validation and framed as an adjunct to, rather than a replacement for, expert clinical EEG interpretation.

4.3 Future Work

Future work should focus on several complementary directions. First, models such as SPaRCNet should be fine-tuned using pediatric EEG data, including ECMO and ICU cohorts, to allow the models to better understand pediatric-specific indications of ictal-interictal patterns indicative of seizure. As newer ICU EEG models such as MORGOTH enter clinical workflows, they should likewise undergo formal evaluation and, where needed, pediatric-specific recalibration rather than being assumed to generalize from adult training data.

Second, pediatric-aware preprocessing may present another opportunity for improved model performance, including band-pass choices and artifact handling. Distinct preprocessing

pipelines may be required for neonatal, infant, and older pediatric patients, reflecting known developmental changes in EEG background and seizure morphology.

Third, there is currently no large, open-source EEG dataset for pediatric patients in the ICU. Analogous to the adult IIIC dataset curated for SPaRCNet, the field would benefit greatly from the development of a large-scale multi-center pediatric EEG corpus with expert annotations of seizures and other injury patterns. The assembly of such datasets would enable robust training, external validation, and head-to-head comparison of pediatric-specific seizure-detection models, including adapted versions of SPaRCNet and MORGOTH. Ultimately, prospective studies embedding these tools into PICU workflow will be needed to determine whether pediatric-optimized models improve time to seizure detection, reduce EEG interpretation burden, and meaningfully affect clinical outcomes.

4.4 Conclusion

Pediatric patients supported with ECMO face a substantial risk of electrographic seizures that are difficult to recognize in real time with conventional EEG workflows. This work establishes the first benchmark for automated seizure detection in pediatric ECMO and highlights the urgent need for pediatric-specific deep learning models. These results indicate that pediatric-specific fine-tuning or model design is likely necessary for reliable automated seizure detection in pediatric ECMO patients. By developing a pediatric-focused preprocessing and analysis pipeline around SPaRCNet, this work establishes practical methods, open-source code, and quantitative benchmarks that can support future multi-center studies. Ultimately, pediatric-adapted seizure-risk models, trained and validated on larger pediatric critical-care EEG datasets, will be needed to provide robust, real-time neurocritical case decision support and to help deliver safer, more responsive care for children on ECMO.

Appendix A

GitHub Repository

All analysis scripts used to generate the results in this thesis are available at the accompanying GitHub repository:

https://github.com/nbarnouw/sparcnet-pediatric-ecmo-analysis

Appendix B

AUC vs. Threshold Value for LPD and Seizure + LPD Datasets

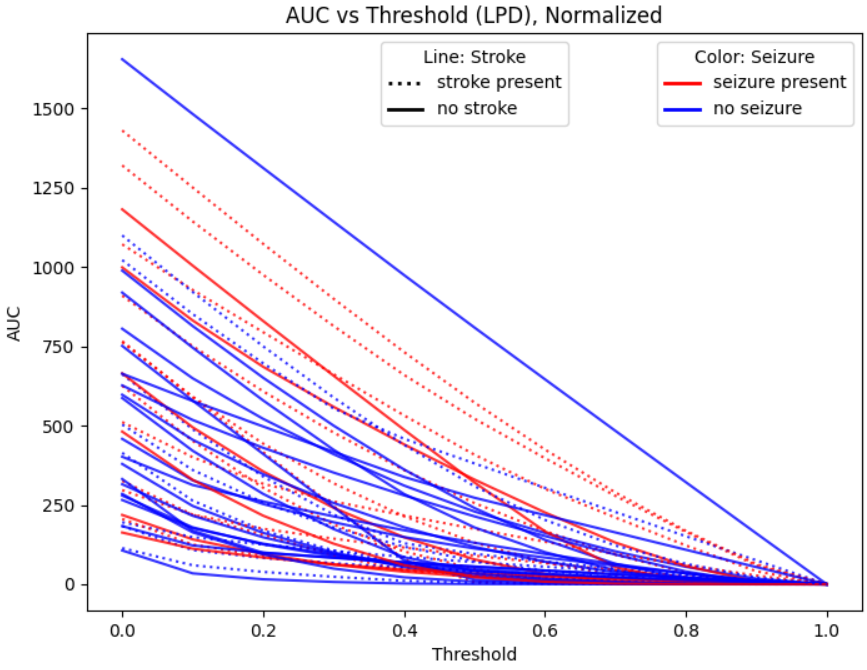


Figure B.1: Patient-level AUC values plotted against threshold values for LPD dataset. Color indicates seizure presence, while line style (dashed or solid) indicates stroke presence.

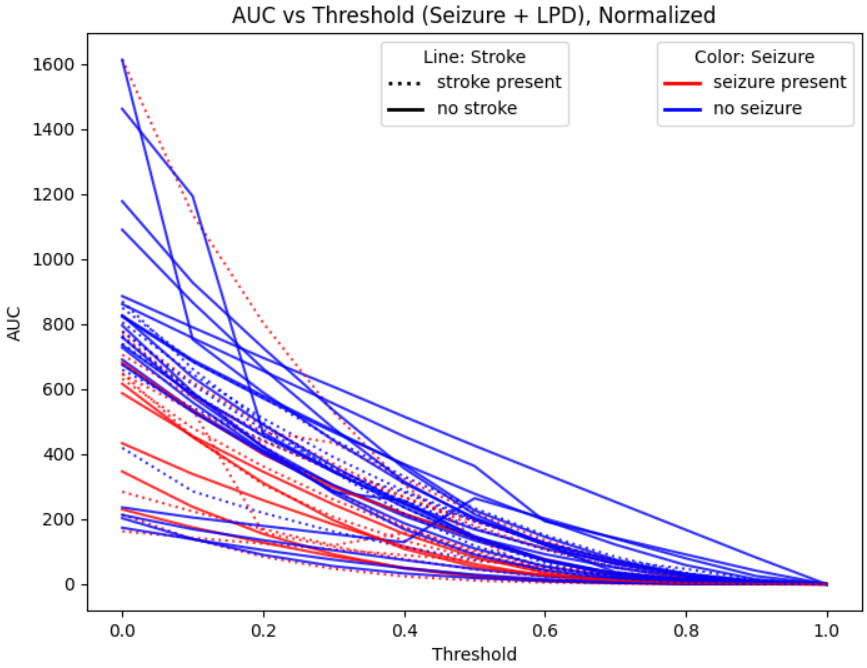


Figure B.2: Patient-level AUC values plotted against threshold values for seizure + LPD dataset. Color indicates seizure presence, while line style (dashed or solid) indicates stroke presence.

References

[1] S. Sahin. "Seizures in a Pediatric Intensive Care Unit: A Prospective Study." *Journal of Tropical Pediatrics | Oxford Academic*, **62**(2), Feb. 2016, pp. 94–100. DOI: 10.1093/tropej/fmv076.

[2] M. Waak. "Real-time seizure detection in paediatric intensive care patients: the RESET child brain protocol." *British Medical Journal*, **12**(6), June 2022. DOI: 10.1136/Bmjopen-2021-05930.

[3] B. Westover. "Development of Expert-Level Classification of Seizures and Rhythmic and Periodic Patterns During EEG Interpretation." *Neurology*, **100**(17), Apr. 2023. DOI: 10.1212/WNL.0000000000207127.

[4] M. Nan Lin. *Epilepsy: Disorder Overview*. Child Neurology Foundation. Aug. 2023.

[5] C. Ezetendu. "Stroke in pediatric ECMO patients: analysis of the National Inpatient Sample (NIS) database." *Pediatric Research*, **92**(3), June 2022, pp. 754–761. DOI: 10.1038/s41390-022-02088-7.

[6] G. Lin. "Seizures in children undergoing extracorporeal membrane oxygenation: a systematic review and meta-analysis." *Pediatric Research*, **93**(4), July 2022, pp. 755–762. DOI: 10.1038/s41390-022-02187-5.

[7] D. deCampo. "Continuous EEG (cEEG) protocol improves seizure detection in children on extracorporeal membrane oxygenation (ECMO)." *Journal of Child Neurology*, **38**(10), Apr. 2024, pp. 581–589. DOI: 10.1177/08830738231190145.

[8] D. Hanalioglu. "Neurophysiologic Features Reflecting Brain Injury During Pediatric ECMO Support." *Neurocritical Care*, **40**(2), Sept. 2023, pp. 759–768. DOI: 10.1007/s12028-023-01836-9.

[9] N. Jette. "Frequency and predictors of nonconvulsive seizures during continuous electroencephalographic monitoring in critically ill children." *Archives of Neurology*, **63**(12), Sept. 2006, pp. 1750–1755. DOI: 10.1001/archneur.63.12.1750.

[10] W. T. Kerr. "The present and future of seizure detection, prediction, and forecasting with machine learning, including the future impact on clinical trials." *Frontiers*, **15**, July 2024. DOI: 10.3389/fneur.2024.1425490.

[11] A. Shoeibi. "A comprehensive comparison of handcrafted features and convolutional autoencoders for epileptic seizures detection in EEG signals." *Expert Systems with Applications*, **163**, Jan. 2021. DOI: 10.1016/j.eswa.2020.113788.

[12] M. A. Alsuwaiket. "Feature Extraction of EEG Signals for Seizure Detection Using Machine Learning Algorthims." *Engineering, Technology & Applied Science Research*, **12**(5), Oct. 2022, pp. 9247–9251. DOI: 10.48084/etasr.5208.

[13] A. S. Abdulhussien. "Automatic seizure detection with different time delays using SDFT and time-domain feature extraction." *The Journal of Biomedical Research*, **36**(1), Jan. 2022, pp. 48–57. DOI: 10.7555/JBR.36.20210124.

[14] R. Mourad. "Machine and deep learning methods for epileptic seizure recognition using EEG data: A systematic review." *Brain Research*, **1864**, Oct. 2025. DOI: 10.1016/j.brainres.2025.149797.

[15] M. Alarfaj. "Deep learning approaches for diagnosing seizure based on EEG signal analysis." *Frontiers in Human Neuroscience*, **19**, Nov. 2025. DOI: 10.3389/fnhum.2025.1669919.

[16] W. Zhao. "Residual and bidirectional LSTM for epileptic seizure detection." *Frontiers in Human Neuroscience*, **18**, June 2024. DOI: 10.3389/fncom.2024.1415967.

[17] A. Abdelhameed. "A Deep Learning Approach for Automatic Seizure Detection in Children With Epilepsy." *Frontiers in Computational Neuroscience*, **15**, Apr. 2021. DOI: 10.3389/fncom.2021.650050.

[18] W. Jennekens. "Topography of maturational changes in EEG burst spectral power of the preterm infant with a normal follow-up at 2 years of age." *Clinical Neurophysiology*, **123**(11), May 2012, pp. 2130–2138. DOI: 10.1016/j.clinph.2012.03.018.

[19] E. Pavlidis. "A review of important EEG features for the assessment of brain maturation in premature infants." *Acta Paediatricia*, **106**(9), June 2017. DOI: 10.1111/apa.13956.

[20] M. R. Mourid. "Artificial Intelligence in Pediatric Epilepsy Detection: Balancing Effectiveness With Ethical Considerations for Welfare." *Health Science Reports*, **8**(1), Jan. 2025. DOI: 10.1002/hsr2.70372.

[21] C. Rubinos. "Automating Critical Care EEG: Enhancing Accuracy, Efficiency, and Patient Outcomes in Critical Care." *Epilepsy Current*, **15**, Mar. 2026. DOI: 10.1177/15357597261426958.

[22] M. Caffarelli. "A quantitative EEG index for the recognition of arterial ischemic stroke in children." *Clinical Neurophysiology*, **156**, Dec. 2023, pp. 113–124. DOI: 10.1016/j.clinph.2023.10.001.