RESEARCH ARTICLE

Development of Expert-Level Classification of Seizures and Rhythmic and Periodic Patterns During EEG Interpretation

Jin Jing, PhD,* Wendong Ge, PhD,* Shenda Hong, PhD, Marta Bento Fernandes, PhD, Zhen Lin, Chaoqi Yang, Sungtae An, Aaron F. Struck, MD, Aline Herlopian, MD, Ioannis Karakis, MD, PhD, MSc, Jonathan J. Halford, MD, Marcus C. Ng, MD, Emily L. Johnson, MD, Brian L. Appavu, MD, Rani A. Sarkis, MD, MSc, Gamaleldin Osman, MD, MS, Peter W. Kaplan, MBBS, FRCP, Monica B. Dhakar, MD, MS, Lakshman Arcot Jayagopal, MD, Zubeda Sheikh, MD, MS, Olga Taraschenko, MD, PhD, Sarah Schmitt, MD, Hiba A. Haider, MD, Jennifer A. Kim, MD, PhD, Christa B. Swisher, MD, Nicolas Gaspard, MD, PhD, Mackenzie C. Cervenka, MD, Andres A. Rodriguez Ruiz, MD, Jong Woo Lee, MD, PhD, Mohammad Tabaeizadeh, MD, Emily J. Gilmore, MD, Kristy Nordstrom, AS, Ji Yeoun Yoo, MD, Manisha G. Holmes, MD, Susan T. Herman, MD, Jennifer A. Williams, MB, BAO, Bch, FRCPI, Jay Pathmanathan, MD, PhD, Fábio A. Nascimento, MD, Ziwei Fan, MS, Samaneh Nasiri, PhD, Mouhsin M. Shafi, MD, PhD, Sydney S. Cash, MD, PhD, Daniel B. Hoch, MD, PhD, Andrew J. Cole, MD, Eric S. Rosenthal, MD, Sahar F. Zafar, MD, Jimeng Sun, PhD,† and M. Brandon Westover, MD, PhD†

Correspondence
Dr. Westover
mwestover@mgh.harvard.edu

Neurology® 2023;100:e1750-e1762. doi:10.1212/WNL.0000000000207127

Abstract

Background and Objectives

Seizures (SZs) and other SZ-like patterns of brain activity can harm the brain and contribute to in-hospital death, particularly when prolonged. However, experts qualified to interpret EEG data are scarce. Prior attempts to automate this task have been limited by small or inadequately labeled samples and have not convincingly demonstrated generalizable expert-level performance. There exists a critical unmet need for an automated method to classify SZs and other SZ-like events with expert-level reliability. This study was conducted to develop and validate a computer algorithm that matches the reliability and accuracy of experts in identifying SZs and SZ-like events, known as "ictal-interictal-injury continuum" (IIIC) patterns on EEG, including SZs, lateralized and generalized periodic discharges (LPD, GPD), and lateralized and generalized rhythmic delta activity (LRDA, GRDA), and in differentiating these patterns from non-IIIC patterns.

Methods

We used 6,095 scalp EEGs from 2,711 patients with and without IIIC events to train a deep neural network, *SPaRCNet*, to perform IIIC event classification. Independent training and test data sets were generated from 50,697 EEG segments, independently annotated by 20 fellowship-trained neurophysiologists. We assessed whether *SPaRCNet* performs at or above the sensitivity, specificity, precision, and calibration of fellowship-trained neurophysiologists for identifying IIIC events. Statistical performance was assessed by the calibration index and by the percentage of experts whose operating points were below the model's receiver operating characteristic curves (ROCs) and precision recall curves (PRCs) for the 6 pattern classes.

Results

SPaRCNet matches or exceeds most experts in classifying IIIC events based on both calibration and discrimination metrics. For SZ, LPD, GPD, LRDA, GRDA, and "other" classes, *SPaRCNet*

RELATED ARTICLES

Editorial
Putting the "Big" in Big Data: Learning to Be Just as (Un)certain as a Clinician at EEG
Page 799

Research Article
Interrater Reliability of Expert Electroencephalographers Identifying Seizures and Rhythmic and Periodic Patterns in EEGs
Page 806

MORE ONLINE

Class of Evidence
Criteria for rating therapeutic and diagnostic studies
NPub.org/coe

*These authors contributed equally to this work as first authors.

†These authors contributed equally to this work as senior authors.

From the Department of Neurology (J.J., W.G., M.B.F., M.T., K.N., F.A.N., Z.F., S.N., S.S.C., D.B.H., A.J.C., E.S.R., S.F.Z., M.B.W.), Massachusetts General Hospital, Harvard Medical School, Boston; Massachusetts General Hospital Clinical Data Animation Center (CDAC) (J.J., W.G., M.B.F., M.T., F.A.N., Z.F., S.N., S.S.C., D.B.H., S.F.Z., M.B.W.), Boston; National Institute of Health Data Science (S.H.), Peking University, Beijing, China; College of Computing (Z.L., C.Y., J.S.), University of Illinois at Urbana-Champaign; College of Computing (S.A.), Georgia Institute of Technology, Atlanta; Department of Neurology (A.F.S.), University of Wisconsin-Madison; William S. Middleton Memorial Veterans Hospital (A.F.S.), Madison, WI; Yale New Haven Hospital (A.H., J.A.K., E.J.G.), Yale University, CT; Emory University School of Medicine (I.K., A.A.R.R.), Atlanta, GA; Medical University of South Carolina (J.J.H., S.S.), Charleston; University of Manitoba (M.C.N.), Winnipeg, Canada; Johns Hopkins School of Medicine (E.L.J., P.W.K., M.C.C.), Baltimore, MD; University of Arizona College of Medicine (B.L.A.), Phoenix; Brigham and Women's Hospital (R.A.S., J.W.L.), Boston, MA; Mayo Clinic (G.O.), Rochester, MN; Warren Alpert School of Medicine (M.B.D.), Brown University, Providence, RI; University of Nebraska Medical Center (L.A.J., O.T.), Omaha; West Virginia University Hospitals (Z.S.), Morgantown; University of Chicago (H.A.H.), IL; Atrium Health (C.B.S.), Charlotte, NC; Hôpital Erasme (N.G.), Université Libre de Bruxelles, Belgium; Icahn School of Medicine (J.Y.Y.), Mount Sinai, NY; NYU Grossman School of Medicine (M.G.H.), New York; Barrow Neurological Institute (S.T.H.), Phoenix, AZ; Mater Misericordiae University Hospital (J.A.W.), Dublin, Ireland; University of Pennsylvania (J.P.), Philadelphia; and Beth Israel Deaconess Medical Center (M.M.S.), Harvard Medical School, Boston, MA.

Go to Neurology.org/N for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Glossary

EUPRC = experts under PR curve; **EUROC** = experts under ROC curve; **FPR** = false-positive rate; **GPD** = generalized periodic discharge; **GRDA** = generalized rhythmic delta activity; **IIC** = ictal-interictal continuum; **IIIC** = ictal-interictal-injury continuum; **LPD** = lateralized periodic discharge; **LRDA** = lateralized rhythmic delta activity; **PPV** = positive predicted value; **PR** = precision recall; **ROC** = receiver operating characteristic; **SZ** = seizure; **TPR** = true-positive rate.

exceeds the following percentages of 20 experts—ROC: 45%, 20%, 50%, 75%, 55%, and 40%; PRC: 50%, 35%, 50%, 90%, 70%, and 45%; and calibration: 95%, 100%, 95%, 100%, 100%, and 80%, respectively.

Discussion

SPaRCNet is the first algorithm to match expert performance in detecting SZs and other SZ-like events in a representative sample of EEGs. With further development, *SPaRCNet* may thus be a valuable tool for an expedited review of EEGs.

Classification of Evidence

This study provides Class II evidence that among patients with epilepsy or critical illness undergoing EEG monitoring, *SPaRCNet* can differentiate (IIIC) patterns from non-IIIC events and expert neurophysiologists.

Seizures (SZs) and other SZ-like types of brain activities known as "ictal-interictal-injury continuum" (IIIC) patterns (aka "ictal-interictal continuum" [IIC] patterns) are seen commonly during brain monitoring with EEG in patients with epilepsy or critical illness.[1-5] IIIC events can damage the brain, especially when prolonged,[4,6-12] and are biomarkers of impending delayed cerebral ischemia[1,13,14] and risk for posttraumatic epilepsy.[13] Fellowship-trained clinical neurophysiologists are the gold standard for identifying IIIC events. However, subspecialists are scarce, and in most of the world, brain monitoring services are unavailable. There is a critical need for methods that detect IIIC events automatically without compromising accuracy.

Prior efforts to automate SZ detection have been limited by lack of large well-annotated data sets to train and evaluate algorithms (eAppendix 1 and eFigure 1, links.lww.com/WNL/C668), and there have been only a few attempts to detect IIIC events other than SZs[15] (eTable 1). To address this gap, we created a set of 50,697 IIIC and non-IIIC events from 2,711 patients' (6,095 EEGs) and obtained independent annotations from 124 raters, 20 of whom annotated sufficient data to compare against algorithm performance and qualified as experts (physicians with subspecialty EEG training). We then trained a computer program to classify IIIC events and distinguish them from non-IIIC events with accuracy matching clinical experts. By providing expert-level automatic classification of SZs and highly epileptiform brain activity on EEG, we can augment brain monitoring in patients at high risk for such events, particularly in settings without continuously available human expertise.

Methods

Standard Protocol Approvals, Registrations, and Patient Consents

We used 6,095 EEGs recorded at Massachusetts General Hospital during clinical care from 2,711 patients with and without IIIC events (median 18.1 hours; interquartile range 21.5 hours) (Table 1). All data were deidentified. This study was conducted under a protocol approved by the Mass General Brigham Institutional Review Board (Protocol Number: 2013P001024), which waived the requirement for informed consent for this retrospective data analysis study.

IIIC Annotation by Experts

The process to annotate IIIC and non-IIIC patterns is described in the companion paper and in eAppendices 2, 3, and eFigure 2 (links.lww.com/WNL/C668). There were 5 pattern classes: SZs, lateralized and generalized periodic discharges (LPD, GPD), lateralized and generalized rhythmic delta activities (LRDA, GRDA), and "other" (all non-IIIC patterns). Note that EEG segments that were noisy and too poor in quality to classify were included in the "other" class. In total, 180,000 ten-second EEG segments received labels from at least 3 experts. We designated a subset of 71,982 segments as having "high-quality labels" that had received at least 10 independent annotations from a group of 20 experts who had both completed clinical neurophysiology fellowship training and had each annotated ≥1,000 EEG segments (eFigures 3 and 4). Algorithm performance was measured relative to the annotations from these 20 experts. The remaining segments were considered low quality (labeled by too few experts or by raters without specialty training). Of note, we group all SZ, GPD, LPD, LRDA, and GRDA, following Pohlmann-Eden et al.,[16] and call these "IIIC" as in the work of Chong and Hirsch[5] to distinguish this from the more restrictive "IIC" in the recent revision of the ACNS nomenclature, which includes only rhythmic and periodic patterns with specific additional features.

Training and Test Data Sets

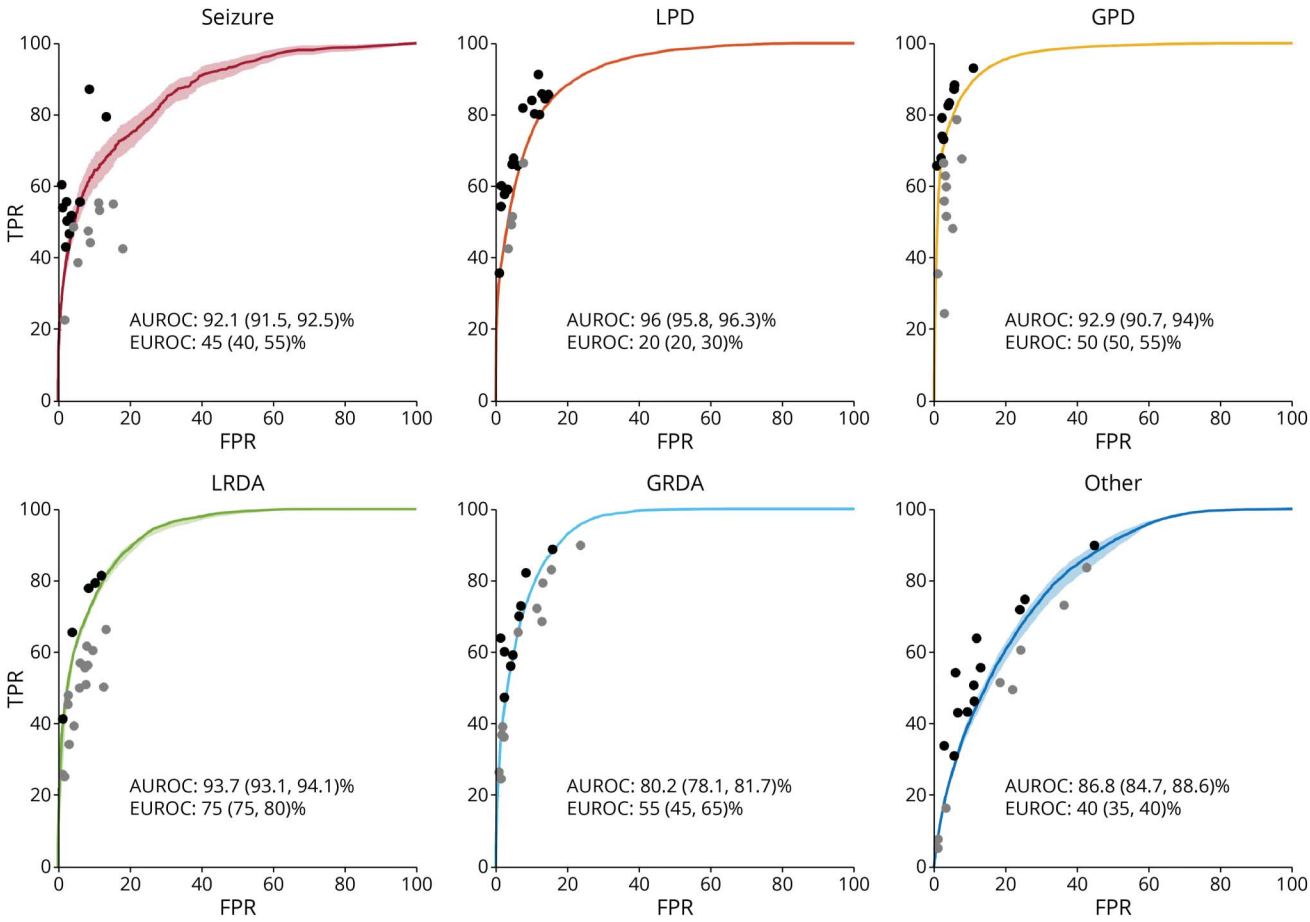
We split the 71,982 EEG segments with high-quality labels into approximately equal-sized training and test sets such that

Table 1 Characteristics of Patients/EEGs Used in Training and Testing of *SPaRCNet*

	Training data		Test data	
	Data set 1 ≥3 labels	**Data set 2 Pseudo**	**Data set 3 ≥10 labels**	**Data set 4 ≥10 labels**
Patients	1,950	1,936	761	761
Age, y, mean (SD)	47.73 (25.52)	47.69 (25.57)	52.55 (24.61)	52.49 (25.72)
Female, n (%)	963 (49.38)	957 (49.43)	395 (51.91)	367 (48.23)
Segments, n (%)	111,095 (100)	103,818 (100)	36,242 (100)	35,740 (100)
Seizure	20,866 (18.78)	3,305 (3.18)	545 (1.50)	549 (1.54)
LPD	15,553 (14.00)	17,338 (1.67)	7,578 (20.90)	7,589 (21.23)
GPD	16,228 (14.61)	16,983 (16.36)	10,040 (27.70)	9,737 (27.24)
LRDA	16,180 (14.56)	12,515 (12.05)	6,077 (16.77)	5,946 (16.64)
GRDA	18,239 (16.42)	11,449 (11.03)	5,093 (14.05)	5,067 (14.18)
Others	24,029 (21.63)	42,228 (40.68)	6,909 (19.06)	6,852 (19.17)
Expert performance relative to consensus of others (%), mean (range)				
Seizure		TPR	52 (23 to 81)	TPR 52 (23 to 87)
		FPR	6 (2 to 19)	FPR 7 (1 to 18)
		PPV	17 (5 to 33)	PPV 17 (4 to 47)
		CAL	−11 (−65 to 29)	CAL −7 (−64 to 24)
LPD		TPR	68 (42 to 93)	TPR 68 (36 to 91)
		FPR	7 (1 to 16)	FPR 7 (1 to 15)
		PPV	75 (61 to 89)	PPV 74 (6 to 92)
		CAL	−4 (−57 to 42)	CAL −4 (−61 to 35)
GPD		TPR	68 (34 to 87)	TPR 67 (25 to 93)
		FPR	4 (1 to 11)	FPR 4 (1 to 11)
		PPV	87 (74 to 96)	PPV 88 (75 to 97)
		CAL	−1 (−58 to 43)	CAL −1 (−55 to 47)
LRDA		TPR	53 (21 to 88)	TPR 54 (25 to 81)
		FPR	7 (1 to 18)	FPR 7 (1 to 13)
		PPV	61 (38 to 84)	PPV 62 (26 to 85)
		CAL	−5 (−68 to 43)	CAL −7 (−65 to 33)
GRDA		TPR	62 (21 to 93)	TPR 61 (25 to 90)
		FPR	7 (1 to 18)	FPR 7 (1 to 24)
		PPV	62 (33 to 87)	PPV 62 (32 to 84)
		CAL	−3 (−68 to 52)	CAL −3 (−65 to 71)
Other		TPR	51 (7 to 95)	TPR 50 (5 to 90)
		FPR	16 (1 to 44)	FPR 16 (1 to 45)
		PPV	49 (29 to 75)	PPV 5 (28 to 78)
		CAL	−14 (−94 to 6)	CAL −11 (−94 to 58)

Abbreviations: GPD = generalized periodic discharge; GRDA = generalized rhythmic delta activity; LPD = lateralized periodic discharge; LRDA = lateralized rhythmic delta activity.
TPR: true-positive rate = TP/(TP + FN), aka sensitivity; FPR: false-positive rate = FP/(FP + TN), aka 1 − specificity; PPV: positive predictive value = TP/(TP + FP), aka precision. TP = number of true positives; FP = number of false positives; FN = number of false negatives; TN = number of true negatives. CAL: calibration index, the percentage of the maximal possible over-calling or under-calling that an expert's calibration curve exhibits.

Figure 1 ROC Curves



Solid curves are median ROC curves that show model performance; shading indicates 95% confidence bands. Expert operating points (x, y) on the ROC curve are shown as solid circles with (x, y) = (false-positive rate [FPR, aka 1 − specificity], true-positive rate [TPR, aka sensitivity]). Markers are colored in black when they lie above the median ROC curve of the model (better than model performance) and in gray when they lie below (inferior to model performance). EUROC = % of experts under the ROC curve; GPD = generalized periodic discharge; GRDA = generalized rhythmic delta activity; LPD = lateralized periodic discharge; LRDA = lateralized rhythmic delta activity; PPV = positive predicted value; ROC = receiver operating characteristic.

all data from any patient appeared only in the training or test set, and both sets were approximately equal in number of patients, proportion of each IIIC and non-IIIC event type, and agreement among experts (Table 1). Additional EEG segments with labels from ≥3 experts were used to augment the training set.

Model Development

We trained a deep convolutional neural network named *SPaRCNet* ("SPaRC" stands for Seizures, Periodic and Rhythmic pattern Continuum) to classify IIIC events and differentiate these from non-IIIC events using all training data, including the segments with both low-quality and high-quality labels. Technical details are provided in eAppendix 4 (eFigure 5) and eAppendix 5 (eFigure 6 and 7) (links.lww.com/WNL/C668).

Model Evaluation

We evaluated *SPaRCNet* on the test data in 2 ways. (1) Discrimination (eAppendix 6, links.lww.com/WNL/C668): We evaluated *SPaRCNet's* ability to discriminate each IIIC pattern using receiver operating characteristic (ROC) and precision recall (PR) curves. For each IIIC class, we left 1 expert out and calculated the operating point as (x, y) = (false-positive rate [FPR, aka 1 − specificity], true-positive rate [TPR, aka sensitivity]) for ROC analysis and (x, y) = (TPR, positive predicted value [PPV]) for precision-recall analysis, relative to the consensus (majority label) of other 19 experts. We note that PPV depends on TPR and FPR and on the prevalence (Pr) of an event: PPV = (TPR × Pr)/(TPR × Pr + FPR × [1 − Pr]). We then calculated ROC and PR curves for the model relative to the same consensus of 19. We repeated this 20 times for each expert with bootstrapping 10,000 times to derive CIs. *SPaRCNet* outperformed a given expert if its ROC and PR curves were above that expert's operating point, and we quantified overall discrimination by the % of experts *SPaRCNet* outperformed. Specifically, we calculate EUROC (% of experts under *SPaRCNet's* ROC curve) and EUPRC (% of experts under *SPaRCNet's* PR curve). (2) Calibration (eAppendix 7): We evaluated the accuracy of probabilities

assigned by *SPaRCNet* vs experts using calibration curves, as in the companion paper. Model calibration was quantified as the % of experts with the worse (larger) calibration index than *SPaRCNet* (eFigure 8 and 9). Additional metrics are presented in the supplement, including expert-algorithm reliability metrics (eAppendix 8) and confusion matrices (eAppendix 9, eFigure 10, and eTable 2). CIs and *p* values for performance statistics were obtained by bootstrapping. We define *SPaRCNet* as having performance superior to experts on a given metric (e.g., experts under *SPaRCNet's* PR curve [EUPRC] or experts under *SPaRCNet's* ROC curve [EUROC]) if the lower limit of its 95% CI exceeds 50%. We define noninferior to mean that the lower limit of the 95% CI exceeds 20%.

Visualization of Model Outputs

We generated a 2-dimensional visualization ("embedding map," Figure 3) of the relationships between IIIC patterns and non-IIIC patterns learned by the model through the UMAP algorithm[17] (see eAppendix 10 and eFigure 6, links.lww.com/WNL/C668).

Data Availability

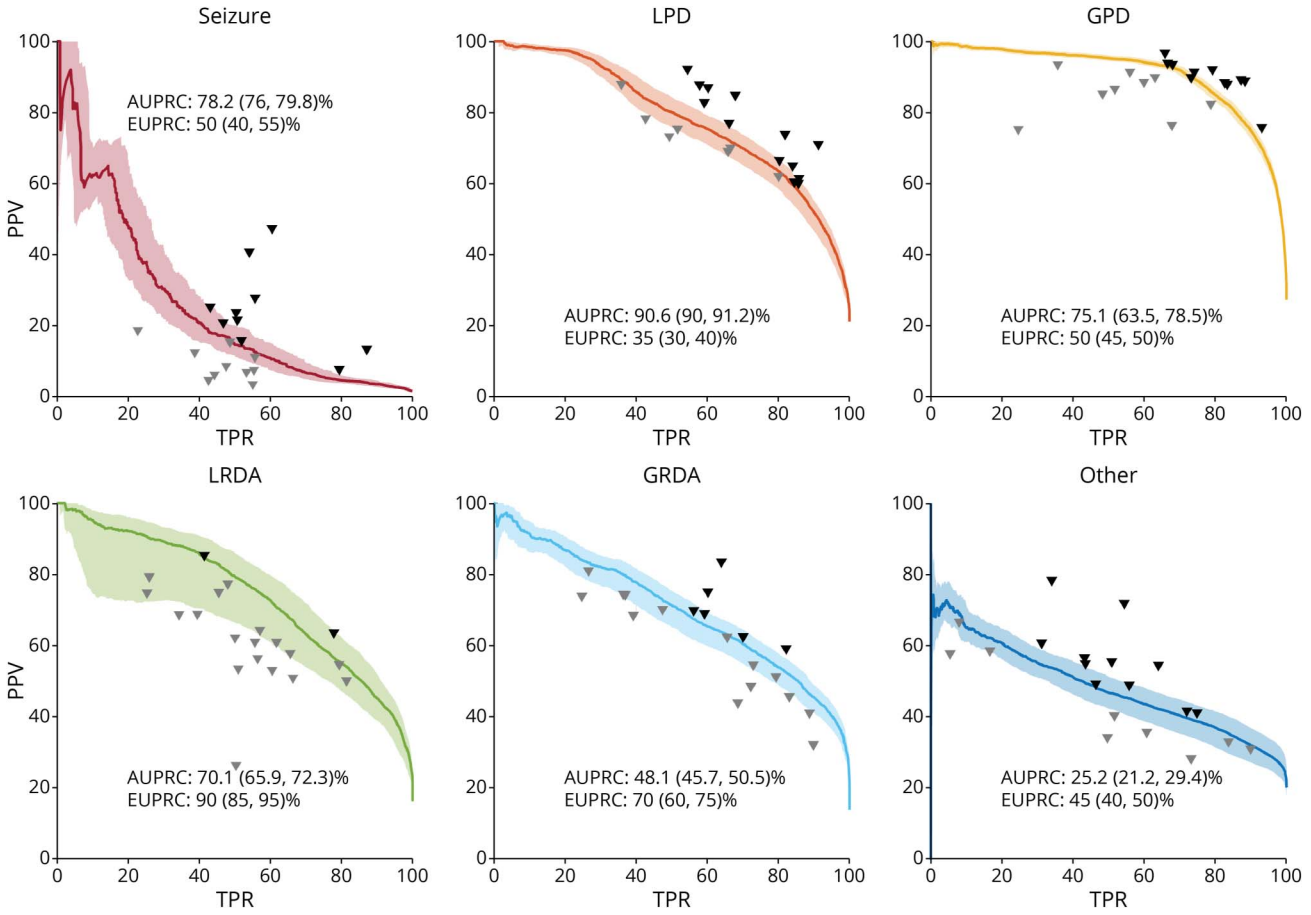
The data and code in this study will be available after approval of a data access agreement, pledging to not reidentify individuals or share the data with a third party. All data inquiries should be addressed to the corresponding author.

Results

Detection of SZs and Other IIIC Patterns

As listed in Table 1, the mean expert performance metrics (TPR, FPR, PPV) for SZ, LPD, GPD, LRDA, GRDA, and "other" classes are (52, 7, 17)%, (68, 7, 74)%, (67, 4, 88)%, (54, 7, 62)%, (61, 7, 62)%, and (50, 16, 5)%. *SPaRCNet* achieves performance comparable with or exceeding experts, as demonstrated by ROC (Figure 1) and PR (Figure 2) curves for all classes. For SZ, LPD, GPD, LRDA, GRDA, and "other" classes, the percentages (median with 95% CI) of the 20 experts for which *SPaRCNet's* ROC curve lies above expert operating points are 45 (40, 55)%, 20 (20, 30)%, 50 (50, 55)%, 75 (75, 80)%, 55 (45, 65)%, and 40 (35, 40)%,

Figure 2 PR Curves



Solid curves are median PR curves that show model performance; shading indicates 95% confidence bands. Expert operating points (x, y) on the PR curve are shown as solid triangles with (x, y) = (TPR, precision [aka positive predictive value (PPV)]). Markers are colored in black when they lie above the median PR curve of the model (better than model performance) and in gray when they lie below (inferior to model performance). EUPRC = % of experts under the PR curve; GPD = generalized periodic discharge; GRDA = generalized rhythmic delta activity; LPD = lateralized periodic discharge; LRDA = lateralized rhythmic delta activity; PR = precision recall.

Figure 3 Maps of the Ictal-Interictal-Injury Continuum Learned by SPaRCNet

Two-dimensional coordinates are calculated by an algorithm (UMAP) such that patterns assigned similar probabilities for each class by the model are near each other in the map. The map learned by *SparCNet* (model) forms a "starfish" pattern, with the 5 IIIC patterns (SZ, LPD, GPD, LRDA, and GRDA) emanating as arms from a central region containing non-IIIC patterns. The coloring of the map indicates the model's classification decisions and closely matches the pattern obtained by overlaying expert-consensus labels (human). Model uncertainty (uncertainty), indicating the degree to which the model assigns similar probabilities to multiple patterns, is greatest near the central region and decreases toward the tips of the "starfish" arms. The probability that an EEG segment represents a seizure or any one of the 4 most highly epileptiform patterns (the sum of the probabilities of SZ, LPD, GPD, or LRDA is shown in SZ burden and IIIC burden). GPD = generalized periodic discharge; GRDA = generalized rhythmic delta activity; IIIC = ictal-interictal-injury continuum; LPD = lateralized periodic discharge; LRDA = lateralized rhythmic delta activity; SZ = seizure.

respectively. For the PR curves, these percentages are 50 (40, 55)%, 35 (30, 40)%, 50 (45, 50)%, 90 (85, 95)%, 70 (60, 75)%, and 45 (40%, 50%). Calibration results are shown in eFigure 8 (links.lww.com/WNL/C668). *SPaRCNet* demonstrates better calibration than 80% of experts for all 6 classes and better than 95% of experts for 5 classes (SZ, LPD, GPD, LRDA, and GRDA).

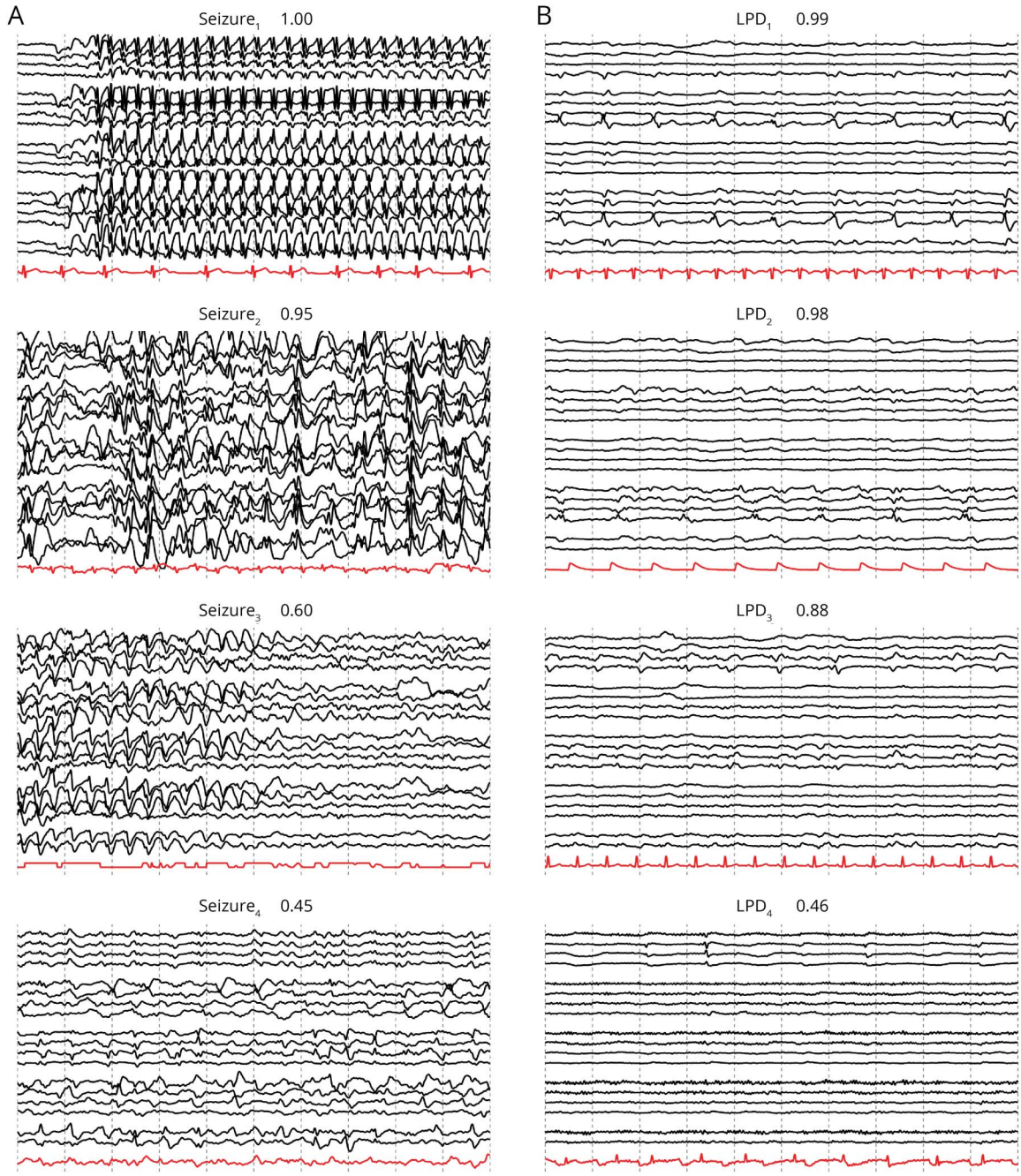
Based on these results, for EUROC, *SpaRCNet* achieves performance superior to experts on GPD and LRDA and

noninferior on SZ, LPD, GRDA, and "other." For EUPRC, *SpaRCNet* achieves performance superior to experts on LRDA and GRDA, and noninferior on SZ, LPD, GPD, and "other." Overall, these results indicate that *SPaRCNet* can classify SZs and other IIIC events and distinguish them from non-IIIC events at least and human experts, and with calibration better than most individual experts, with performance comparable with the consensus of a committee of experts.

Visualization of the IIIC

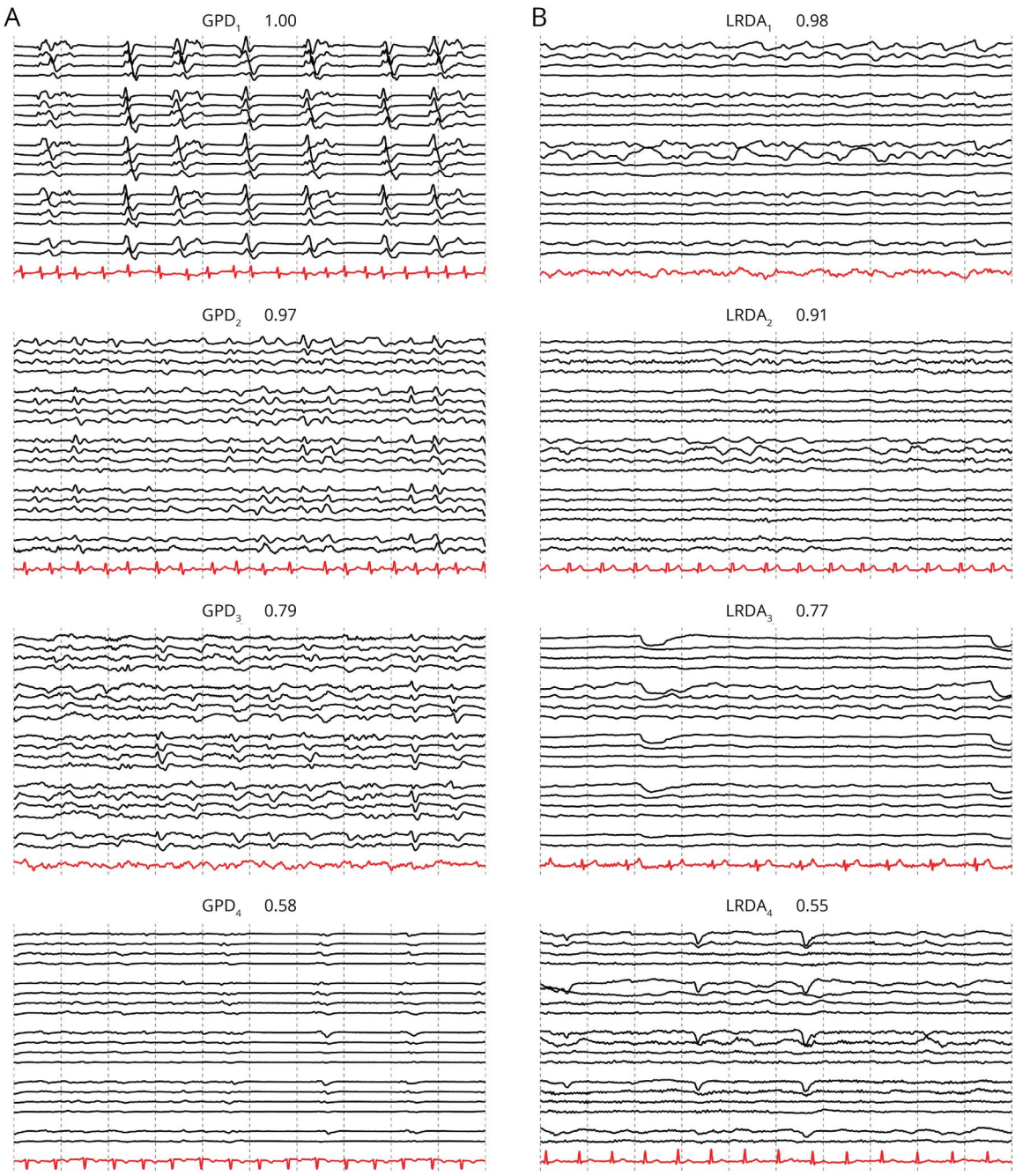
A 2-dimensional visualization ("embedding map") of the relationships between IIIC patterns learned by the model is shown in Figure 3. In this starfish shape, each of the 5 IIIC patterns emanates like an arm from a central region of non-IIIC or "other" patterns. Expert labels (Figure 3, human) agree closely with the classifications assigned by *SPaRCNet* (Figure 3, model). Model uncertainty (Figure 3, uncertainty) increases moving inward along the starfish arms, matching the

Figure 4 Examples of Smooth Pattern Transition for SZ (A) and LPD (B)



Samples are selected at different levels of model uncertainty ranging from the "starfish" arm tips toward the central area. IIIC patterns transition smoothly between distinct prototype patterns at the starfish arm tips into less distinct patterns near the body, lending credence to the concept of a "continuum" between ictal and interictal EEG patterns. IIIC = ictal-interictal-injury continuum; LPD = lateralized periodic discharge; SZ = seizure.

Figure 5 Examples of Smooth Pattern Transition for GPD (A) and LRDA (B)

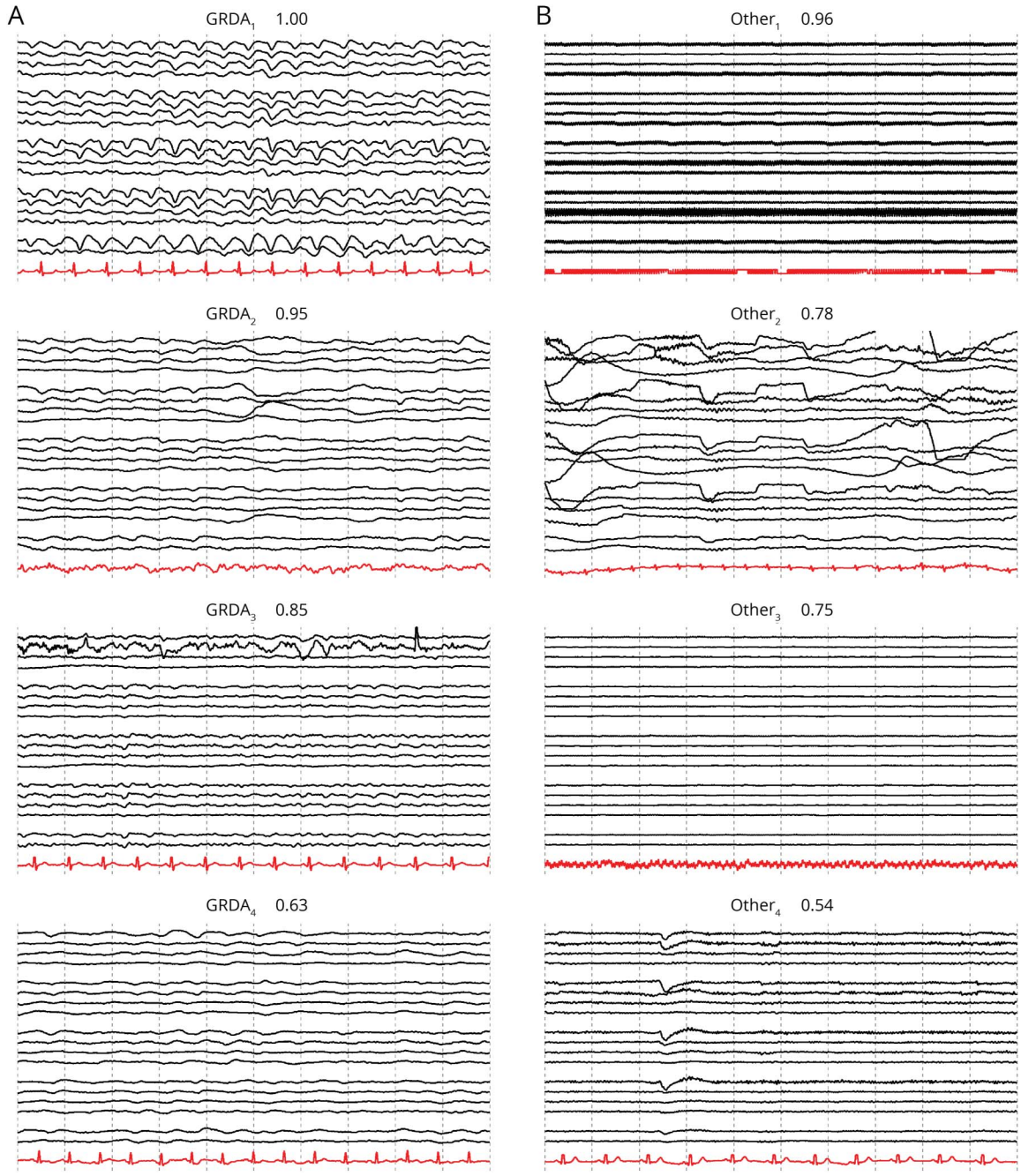


Samples are selected at different levels of model uncertainty ranging from the "starfish" arm tips toward the central area. GPD = generalized periodic discharge; LRDA = lateralized rhythmic delta activity.

increasing haphazardness of expert labels near the central region (Figure 3, model). Plotting the probability of SZs (Figure 3, SZ burden) or of being any IIIC pattern (Figure 3, IIIC burden) and viewing examples from various points along the starfish arms (Figures 4–6) demonstrate that in the map learned by *SPaRCNet*, IIIC patterns transition smoothly between distinct prototype patterns at the starfish arm tips into less distinct patterns near the body, lending credence to the concept of a "continuum" between ictal and interictal EEG patterns.

In addition, we conducted a preliminary investigation of how automated clustering and displaying the outputs of *SPaRCNet* in a manner that allows experts to review the EEG based on features rather than temporal order (i.e., conventional page-by-consecutive-page review) may guide which snippets should be examined by experts, allow rapid identification of similar snippets, and allow efficient review of an entire continuous EEG recording (eAppendix 11 and eFigure 11, links.lww.com/WNL/C668). We found that the mean time for an expert to review model predictions for a 12-hour EEG is 2.12 minutes,

Figure 6 Examples of Smooth Pattern Transition for GRDA (A) and "Other" (B)



Samples are selected at different levels of model uncertainty ranging from the "starfish" arm tips toward the central area. GRDA = generalized rhythmic delta activity.

with standard deviation of 0.43 minutes (eTable 3). These preliminary results suggest that the model can help facilitate rapid review of continuous EEG data. Further work is needed to establish this claim on a larger set of recordings. To go beyond this hybrid approach to a fully automated method for continuous EEGs will require assembling a large set of continuous EEGs, each labeled by multiple experts to allow rigorous model evaluation and may require further model development.

This study provides Class II evidence that among patients with epilepsy or critical illness undergoing EEG monitoring, *SPaRCNet* can differentiate (IIIC) patterns from non-IIIC events and expert neurophysiologists.

Discussion

Our results demonstrate that *SPaRCNet* identifies SZs and other SZ-like patterns of brain activity and distinguishes these

from non–SZ-like events at least and neurologists with subspecialty training in clinical neurophysiology. *SPaRCNet* achieves this by training on a large set of examples annotated by multiple raters, effectively learning to simulate a committee of experts. Given the large sample, it is likely that the full range of EEG abnormalities is encompassed in our analysis. Although several previous studies have attempted to automate SZ detection (eAppendix 1, links.lww.com/WNL/C668), and one has attempted to automated detection of other IIIC patterns,[15] expert-level performance in a representative data set labeled by a large number of experts has not been demonstrated. By automating a challenging diagnostic task previously limited to specialists, *SPaRCNet* opens a path for expanding brain monitoring to a broader range of patients with epilepsy and critical illness.

An important limitation is that *SPaRCNet* does not identify all EEG patterns of clinical relevance. Examples of other key patterns include burst suppression, nonrhythmic slowing, and nonperiodic epileptiform discharges. In addition, *SPaRCNet* does not attempt to further characterize patterns. For example, it does not localize the onset of SZs, determine the frequency of discharges within GPDs or LPDs, and attempt to determine the morphology of GPDs (e.g., triphasic vs nontriphasic), or other additional more detailed features specified in the recent ACNS ICU EEG nomenclature,[18] all of which likely have clinical relevance. Finally, *SPaRCNet* categorizes all non-IIIC patterns as "other," whereas for clinically deployment, it is important to discriminate between physiologic non-IIIC patterns (e.g., "normal" vs burst suppression vs focal slowing) and to identify nonphysiologic patterns such as segments that are corrupted by artifact. Detecting such noisy segments is important clinically to allow staff to readjust the EEG electrodes to improve signal quality or to comment on technical limitations of the study.

Each of these problems is worth pursuing in future studies and could be integrated into *SPaRCNet* in the future. Finally, our data do not allow us to directly report metrics, such as false alarms per hour or detection latency. These metrics require continuous EEGs comprehensively labeled by multiple experts and are the focus of studies currently underway.

Future work should address how IIIC detection can be integrated into practice. Initial work will likely focus on systems that augment clinical neurophysiologists' capabilities. In the short term, expert involvement is still desirable to review and confirm findings, and to interpret that level of uncertainty in the clinical context. Subsequent work should focus on reducing and ultimately removing the need for expert review to make brain monitoring as widely accessible as possible.

SPaRCNet identifies SZs and other IIIC events better than typical clinical experts with specialty training in clinical neurophysiology. This is demonstrated in a large and diverse set of clinical EEGs. *SPaRCNet* may thus be a valuable tool for accelerated or automated review of EEGs.

Study Funding

M.B. Westover received funding from the Glenn Foundation for Medical Research and American Federation for Aging Research (Breakthroughs in Gerontology Grant); American Academy of Sleep Medicine (AASM Foundation Strategic Research Award); Football Players Health Study (FPHS) at Harvard University; Department of Defense through a subcontract from Moberg ICU Solutions, Inc; NIH (R01NS102190, R01NS102574, R01NS107291, RF1AG064312, and R01AG062989); and NSF (award SCH-2014431). A.F. Struck received funding from the NIH (R01NS111022). S.S. Cash was funded by NIH National Institute of Neurological Disorders and Stroke (NINDS) R01 NS062092 and NIH NINDS K24 NS088568. M.B. Dhakar received funding from NIH NINDS NS11672601 and an American Epilepsy Society Infrastructure Research Award and got clinical trial support from Marinus Pharmaceuticals and Parexel Inc. J.A. Kim received support from NIH-NINDS (R25NS06574), AHA, Bee Foundation. O. Taraschenko was supported by research grants from the NIH (P20GM130447) and the American Epilepsy Society-NORSE Institute Seed grant. M.C. Cervenka receives or has received research grants from Nutricia, Vitaflo, Glut1 Deficiency Foundation, and BrightFocus Foundation; honoraria from Nutricia and Vitaflo/Nestle Health Sciences; royalties from Demos/Springer Publishing Company; and consulting fees from Nutricia and Glut1 Deficiency Foundation.

Disclosure

The authors report no disclosures relevant to the manuscript. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* May 13, 2022. Accepted in final form January 12, 2023. Submitted and externally peer reviewed. The handling editor was Associate Editor Barbara Jobst, MD, PhD, FAAN.

Appendix Authors

Name	Location	Contribution
Jin Jing, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data
Wendong Ge, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data
Shenda Hong, PhD	National Institute of Health Data Science, Peking University, Beijing, China	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design

Continued

Appendix *(continued)*

Name	Location	Contribution
Marta Bento Fernandes, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design
Zhen Lin	College of Computing, University of Illinois at Urbana-Champaign	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design
Chaoqi Yang	College of Computing, University of Illinois at Urbana-Champaign	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design
Sungtae An	College of Computing, Georgia Institute of Technology, Atlanta	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design
Aaron F. Struck, MD	Department of Neurology, University of Wisconsin-Madison; William S Middleton Memorial Veterans Hospital, Madison, WI	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Aline Herlopian, MD	Yale New Haven Hospital, Yale University, CT	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Ioannis Karakis, MD, PhD, MSc	Emory University School of Medicine, Atlanta, GA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Jonathan J. Halford, MD	Medical University of South Carolina, Charleston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Marcus C. Ng, MD	University of Manitoba, Winnipeg, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Emily L. Johnson, MD	Johns Hopkins School of Medicine, Baltimore, MD	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Brian L. Appavu, MD	University of Arizona College of Medicine, Phoenix	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design

Appendix *(continued)*

Name	Location	Contribution
Rani A. Sarkis, MD, MSc	Brigham and Women's Hospital, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Gamaleldin Osman, MD, MS	Mayo Clinic, Rochester, MN	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Peter W. Kaplan, MBBS, FRCP	Johns Hopkins School of Medicine, Baltimore, MD	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Monica B. Dhakar, MD, MS	Warren Alpert School of Medicine, Brown University, Providence, RI	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Lakshman Arcot Jayagopal, MD	University of Nebraska Medical Center, Omaha	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Zubeda Sheikh, MD, MS	West Virginia University Hospitals, Morgantown	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Olga Taraschenko, MD, PhD	University of Nebraska Medical Center, Omaha	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Sarah Schmitt, MD	Medical University of South Carolina, Charleston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Hiba A. Haider, MD	University of Chicago, IL	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Jennifer A. Kim, MD, PhD	Yale New Haven Hospital, Yale University, CT	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Christa B. Swisher, MD	Atrium Health, Charlotte, NC	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design

Appendix *(continued)*

Name	Location	Contribution
Nicolas Gaspard, MD, PhD	Hôpital Erasme, Université Libre de Bruxelles, Belgium	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Mackenzie C. Cervenka, MD	Johns Hopkins School of Medicine, Baltimore, MD	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Andres A. Rodriguez Ruiz, MD	Emory University School of Medicine, Atlanta, GA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Jong Woo Lee, MD, PhD	Brigham and Women's Hospital, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Mohammad Tabaeizadeh, MD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Emily J. Gilmore, MD	Yale New Haven Hospital, Yale University, CT	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Kristy Nordstrom, AS	Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Ji Yeoun Yoo, MD	Icahn School of Medicine, Mount Sinai, NY	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Manisha G. Holmes, MD	NYU Grossman School of Medicine, New York	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Susan T. Herman, MD	Barrow Neurological Institute, Phoenix, AZ	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design

Appendix *(continued)*

Name	Location	Contribution
Jennifer A. Williams, MB, BAO, Bch, FRCPI	Mater Misericordiae University Hospital, Dublin, Ireland	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Jay Pathmanathan, MD, PhD	University of Pennsylvania, Philadelphia	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Fábio A. Nascimento, MD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Ziwei Fan, MS	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Samaneh Nasiri, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Mouhsin M. Shafi, MD, PhD	Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Sydney S. Cash, MD, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Daniel B. Hoch, MD, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Andrew J. Cole, MD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Eric S. Rosenthal, MD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design

Continued

Appendix *(continued)*

Name	Location	Contribution
Sahar F. Zafar, MD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
Jimeng Sun, PhD	College of Computing, University of Illinois at Urbana-Champaign	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data
M. Brandon Westover, MD, PhD	Department of Neurology, Massachusetts General Hospital, Harvard Medical School; Massachusetts General Hospital Clinical Data Animation Center (CDAC), Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

References

1. Westover MB, Shafi MM, Bianchi MT, et al. The probability of seizures during EEG monitoring in critically ill adults. *Clin Neurophysiol*. 2015;126(3):463-471. doi:10.1016/j.clinph.2014.05.037
2. Claassen J, Jetté N, Chum F, et al. Electrographic seizures and periodic discharges after intracerebral hemorrhage. *Neurology*. 2007;69(13):1356-1365. doi:10.1212/01.wnl.0000281664.02615.6c
3. Oddo M, Carrera E, Claassen J, Mayer SA, Hirsch LJ. Continuous electroencephalography in the medical intensive care unit. *Crit Care Med*. 2009;37(6):2051-2056. doi:10.1097/CCM.0b013e3181a00604
4. Kurtz P, Gaspard N, Wahl AS, et al. Continuous electroencephalography in a surgical intensive care unit. *Intensive Care Med*. 2014;40(2):228-234. doi:10.1007/s00134-013-3149-8
5. Chong DJ, Hirsch LJ. Which EEG patterns warrant treatment in the critically ill? Reviewing the evidence for treatment of periodic epileptiform discharges and related patterns. *J Clin Neurophysiol*. 2005;22(2):79-91.
6. Claassen J, Hirsch LJ, Frontera JA, et al. Prognostic significance of continuous EEG monitoring in patients with poor-grade subarachnoid hemorrhage. *Neurocrit Care*. 2006;4(2):103-112. doi:10.1385/NCC:4:2:103
7. Ribeiro A, Singh R, Brunnhuber F. Clinical outcome of generalized periodic epileptiform discharges on first EEG in patients with hypoxic encephalopathy postcardiac arrest. *Epilepsy Behav*. 2015;49:268-272. doi:10.1016/j.yebeh.2015.06.010
8. De Marchis GM, Pugin D, Meyers E, et al. Seizure burden in subarachnoid hemorrhage associated with functional and cognitive outcome. *Neurology*. 2016;86(3):253-260. doi:10.1212/WNL.0000000000002281
9. Zafar SF, Postma EN, Biswal S, et al. Effect of epileptiform abnormality burden on neurologic outcome and antiepileptic drug management after subarachnoid hemorrhage. *Clin Neurophysiol*. 2018;129(11):2219-2227. doi:10.1016/j.clinph.2018.08.015
10. Tabaeizadeh M, Aboul Nour H, Shoukat M, et al. Burden of epileptiform activity predicts discharge neurologic outcomes in severe acute ischemic stroke. *Neurocrit Care*. 2020;32(3):697-706. doi:10.1007/s12028-020-00944-0
11. Zafar SF, Rosenthal ES, Jing J, et al. Automated annotation of epileptiform burden and its association with outcomes. *Ann Neurol*. 2021;90(2):300-311. doi:10.1002/ana.26161
12. Payne ET, Zhao XY, Frndova H, et al. Seizure burden is independently associated with short term outcome in critically ill children. *Brain*. 2014;137(pt 5):1429-1438. doi:10.1093/brain/awu042
13. Kim JA, Boyle EJ, Wu AC, et al. Epileptiform activity in traumatic brain injury predicts post-traumatic epilepsy. *Ann Neurol*. 2018;83(4):858-862. doi:10.1002/ana.25211
14. Kim JA, Rosenthal ES, Biswal S, et al. Epileptiform abnormalities predict delayed cerebral ischemia in subarachnoid hemorrhage. *Clin Neurophysiol*. 2017;128(6):1091-1099. doi:10.1016/j.clinph.2017.01.016
15. Herta J, Koren J, Fürbass F, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. *Epilepsy Behav*. 2015;49:273-279. doi:10.1016/j.yebeh.2015.04.064
16. Pohlmann-Eden B, Hoch DB, Cochius JI, Chiappa KH. Periodic lateralized epileptiform discharges: a critical review. *J Clin Neurophysiol*. 1996;13(6):519-530.
17. McInnes L, Healy J, Melville J. Umap: uniform manifold approximation and projection for dimension reduction [Google Scholar]. 2018.
18. Hirsch LJ, Fong MWK, Leitinger M, et al. American Clinical Neurophysiology Society's standardized critical care EEG terminology: 2021 version. *J Clin Neurophysiol*. 2021;38(1):1-29. doi:10.1097/WNP.0000000000000806