Mahidol University
Faculty of Medicine Siriraj Hospital

SI-NEURO

AI Fundamentals and Breakthrough Applications in Epilepsy

Kanokwan Boonyapisit, MD.

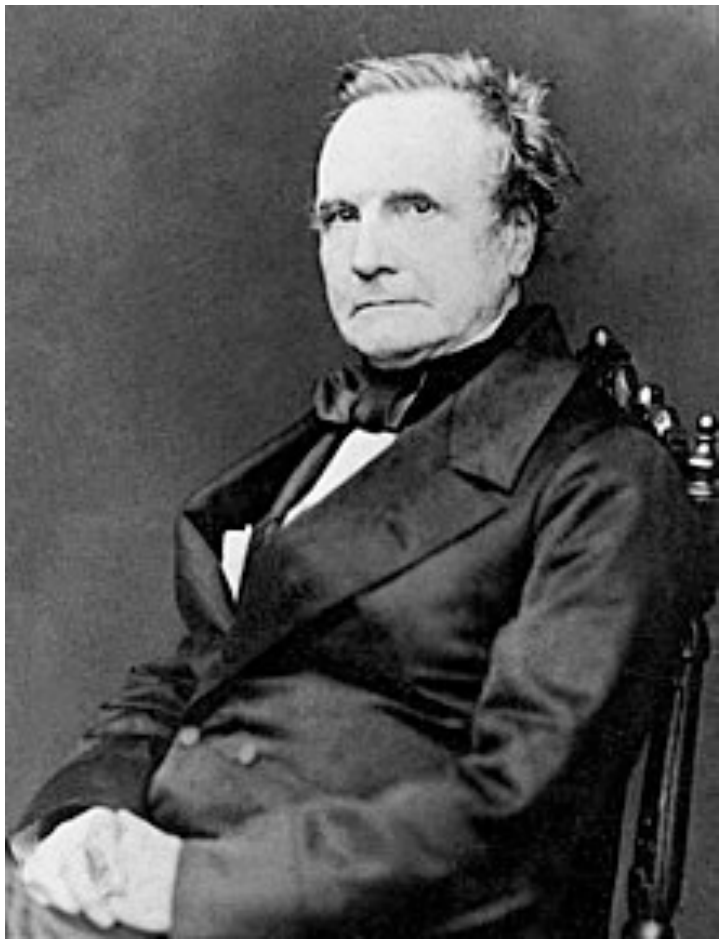
Department of Medicine

Siriraj Hospital

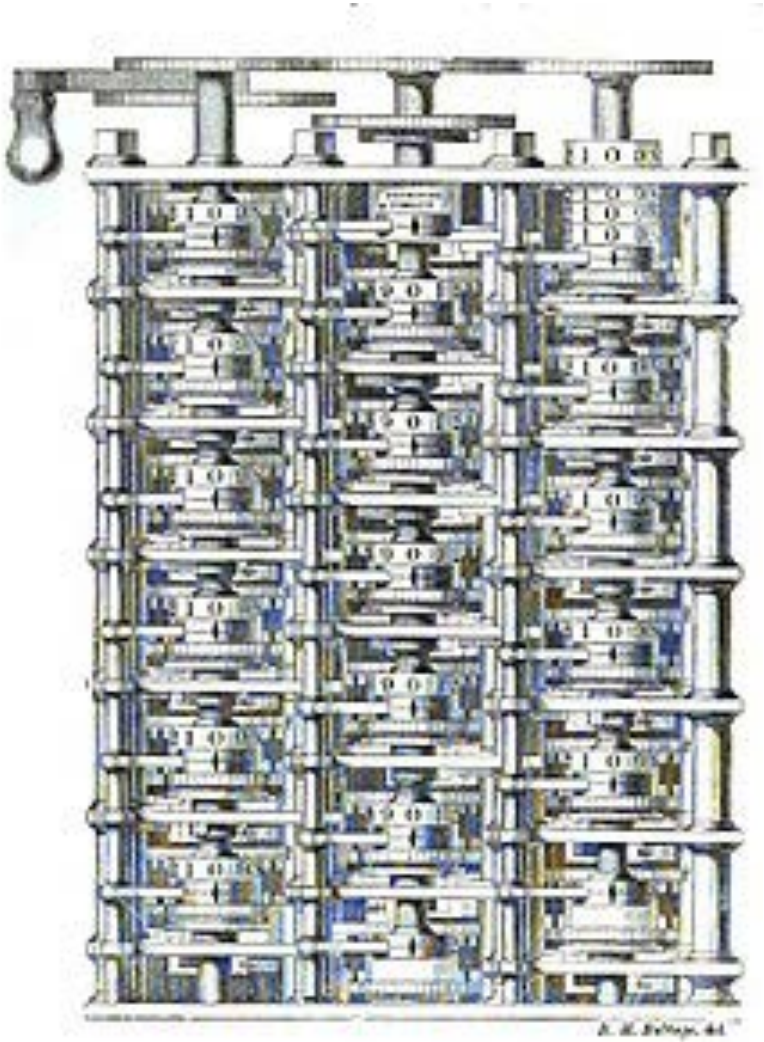
Artificial intelligence (AI)

- “A Field of computer science focused on creating systems to perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making and language translation”

- “A field of research in computer science that develops and studies methods and software that enable machines to perceive their environment and use learning and intelligence to take actions that maximize their chances of achieving defined goals”



Charles Babbage
1791 –1871

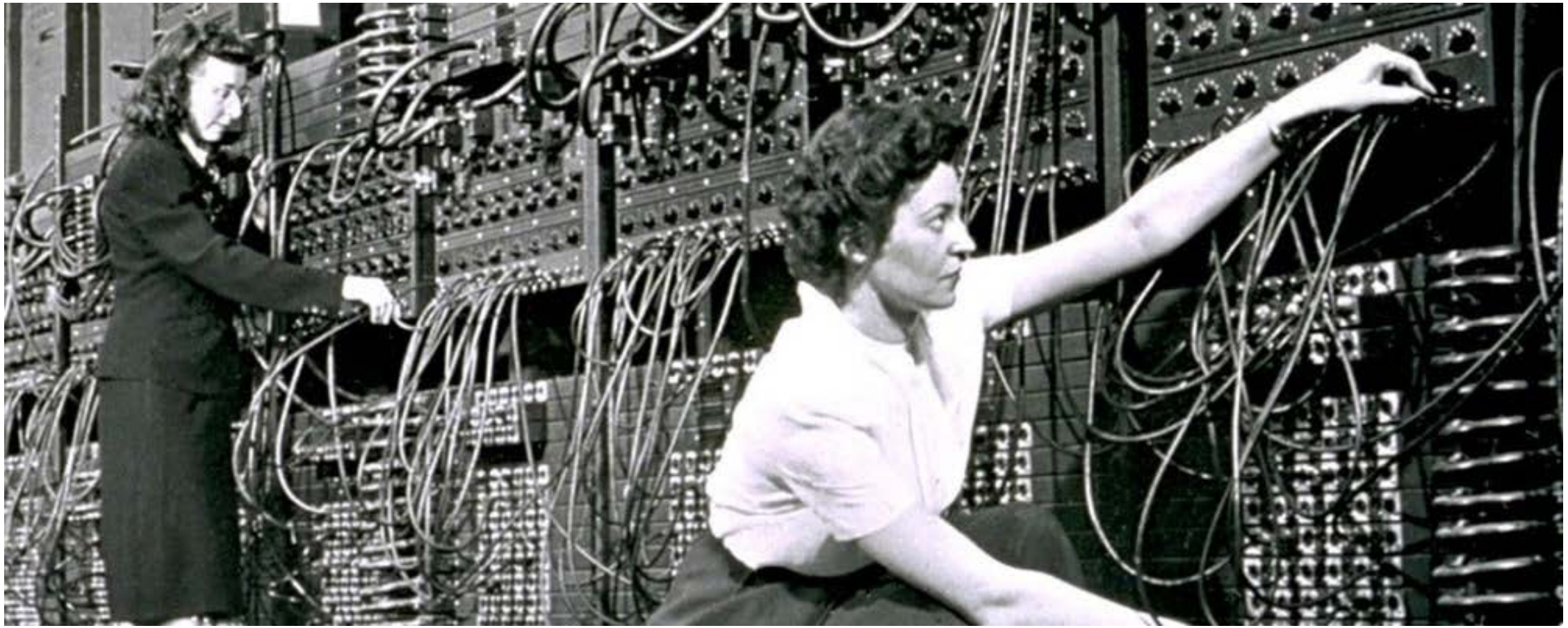


Invented the first mechanical computer, the Difference Engine, that eventually led to more complex electronic designs

Alan Turing
(23 June 1912 – 7 June 1954)

Military Model Enigma I,
in use from 1930

Automatic C
a British earl
stored-progr
by Alan Turi



The **ENIAC, or Electronic Numerical Integrator and Computer**, was the result of a U.S. government-funded project during World War II to build an electronic computer that could be programmed. The project was based out of the University of Pennsylvania's Moore School of Engineering. The design team included engineer J. Presper Eckert Jr. and physicist John Mauchly under the leadership of Herman Goldstine. The team began work on the project in 1943. John von Neumann, a noted mathematician of the day, began consulting on the project in 1944.

Example of AI

Web search engines

Google

Recommendation systems

YouTube NETFLIX

Interaction via speech

amazon alexa Siri

Generative and creative tools

ChatGPT Gemini Claude Llama

Use “large language model”

Games eg. computer chess

Terms

Machine Learning

- A subset of artificial intelligence that involves development of algorithms and statistical models that enable computers to improve their performance on a specific task through experience or data without being explicitly programmed for that task
- Involves using data to train a computer algorithm to maximize its performance based on a single quantitative metric (e.g., accuracy)

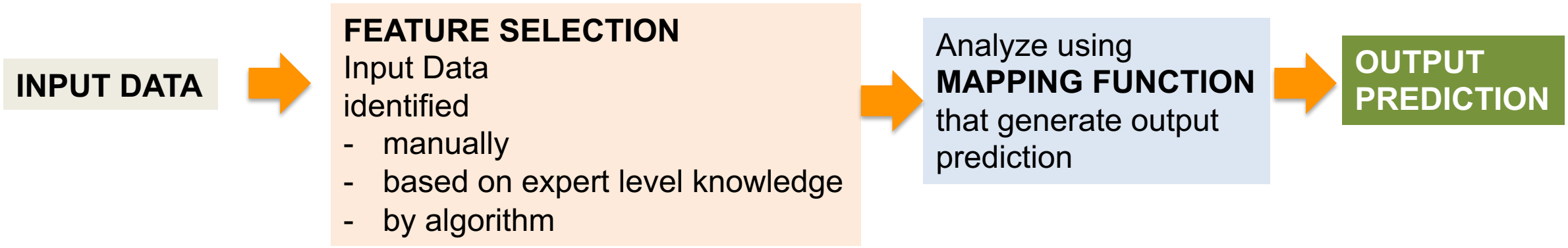
Artificial intelligence

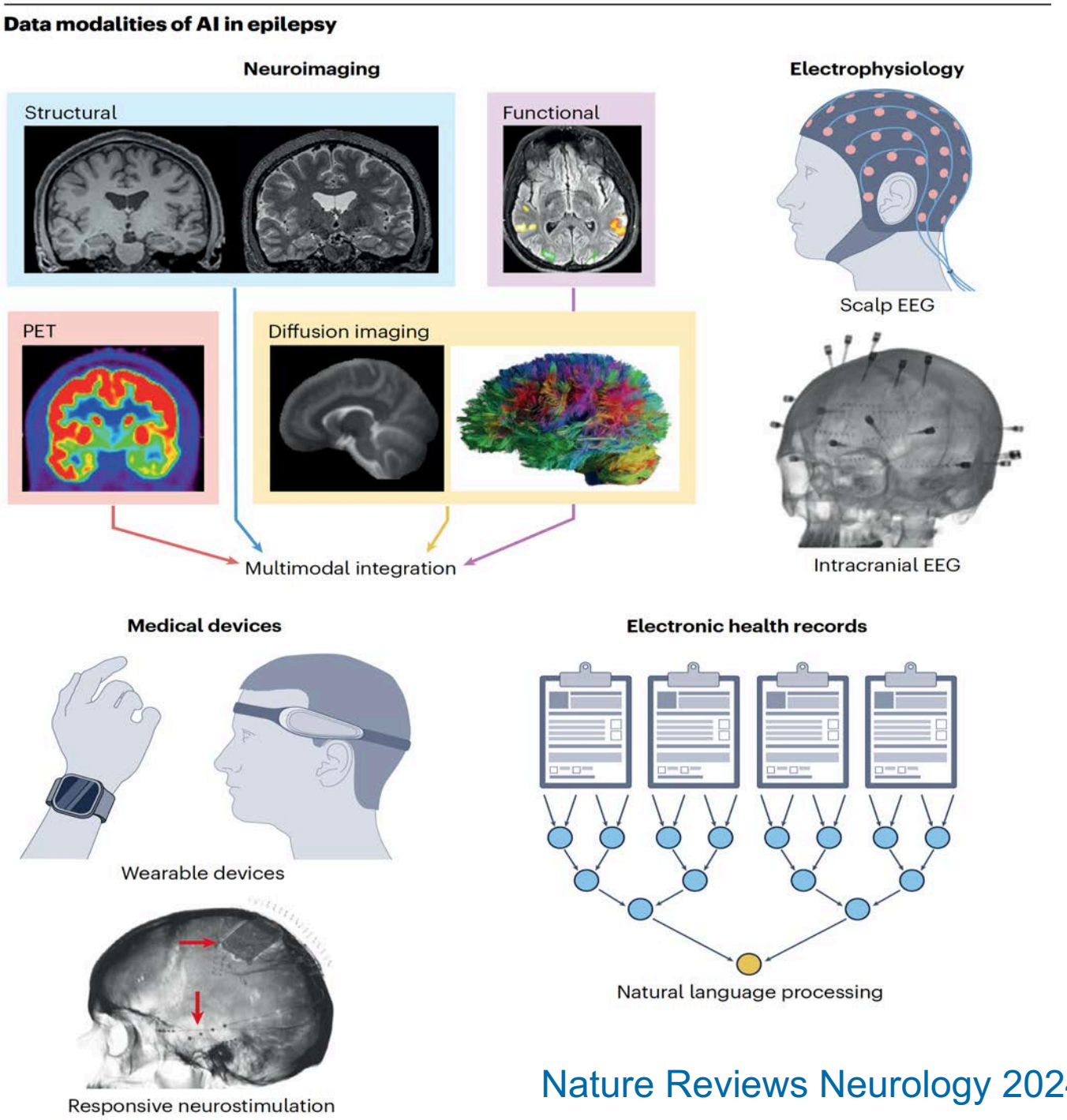
- A Field of computer science focused on creating systems to perform tasks that typically require human intelligence
- Aims to perform a broad range of tasks, including tasks for which there has not been explicit training, and can do so using multiple ML tools

Deep learning

- A machine learning technique that uses layered neural networks to analyse and interpret vast amounts of data

Data pipeline



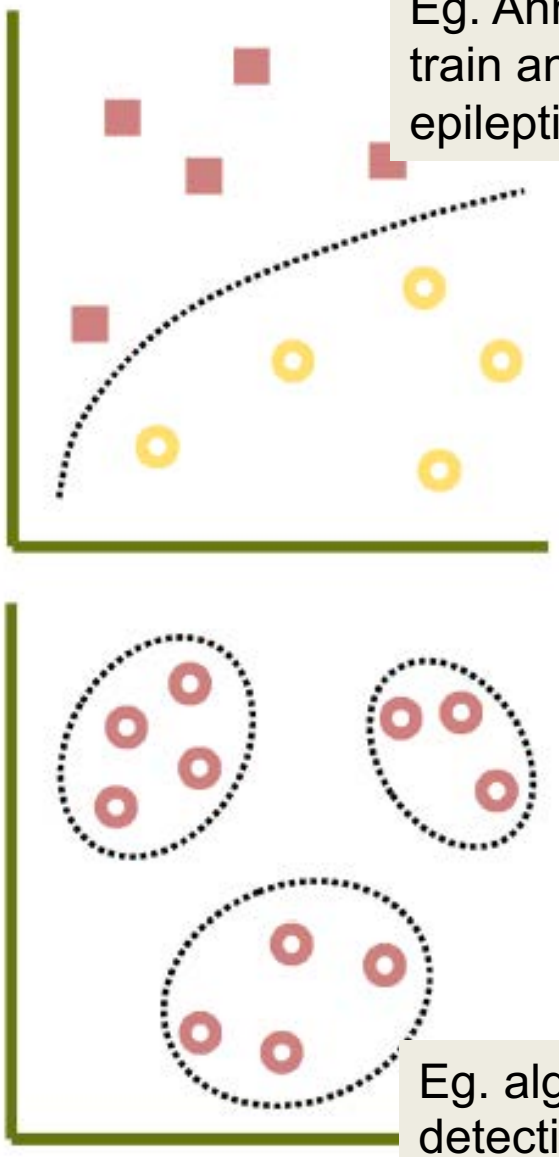


Nature Reviews Neurology 2024;20:319-36

Overview of machine learning concepts

- Supervised learning
- Unsupervised learning

A

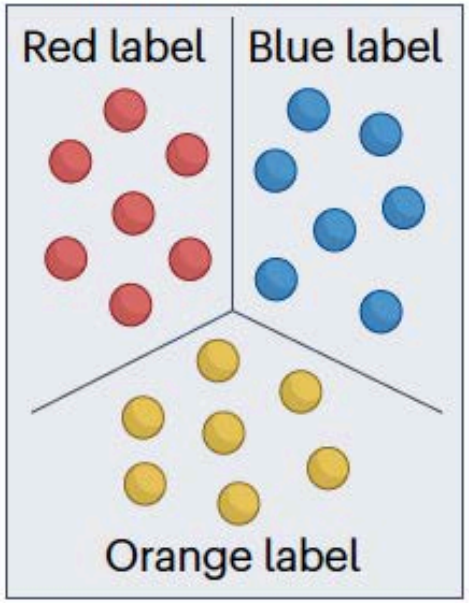


Eg. Annotated EEG recordings may be used to train an algorithm to automatically detect epileptiform discharges.

Supervised learning

t.

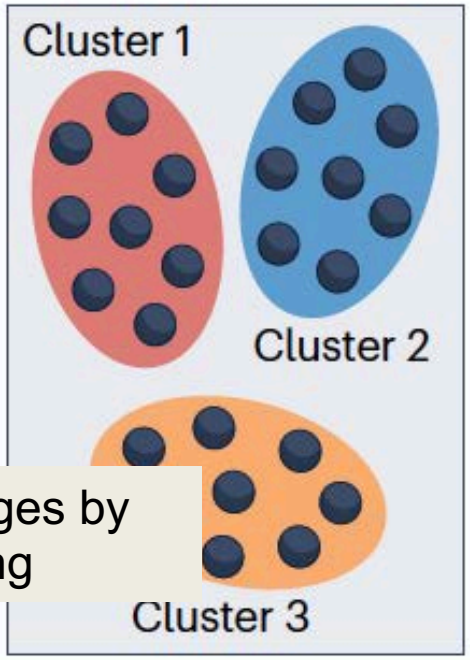
The algorithm learns by analysing these data and making predictions or decisions that are refined according to accuracy against the known labels



Unsupervised learning

An algorithm is trained with an unlabelled dataset, allowing autonomous identification of patterns and structures in the data. Can uncover patterns and intrinsic structures,

without guidance



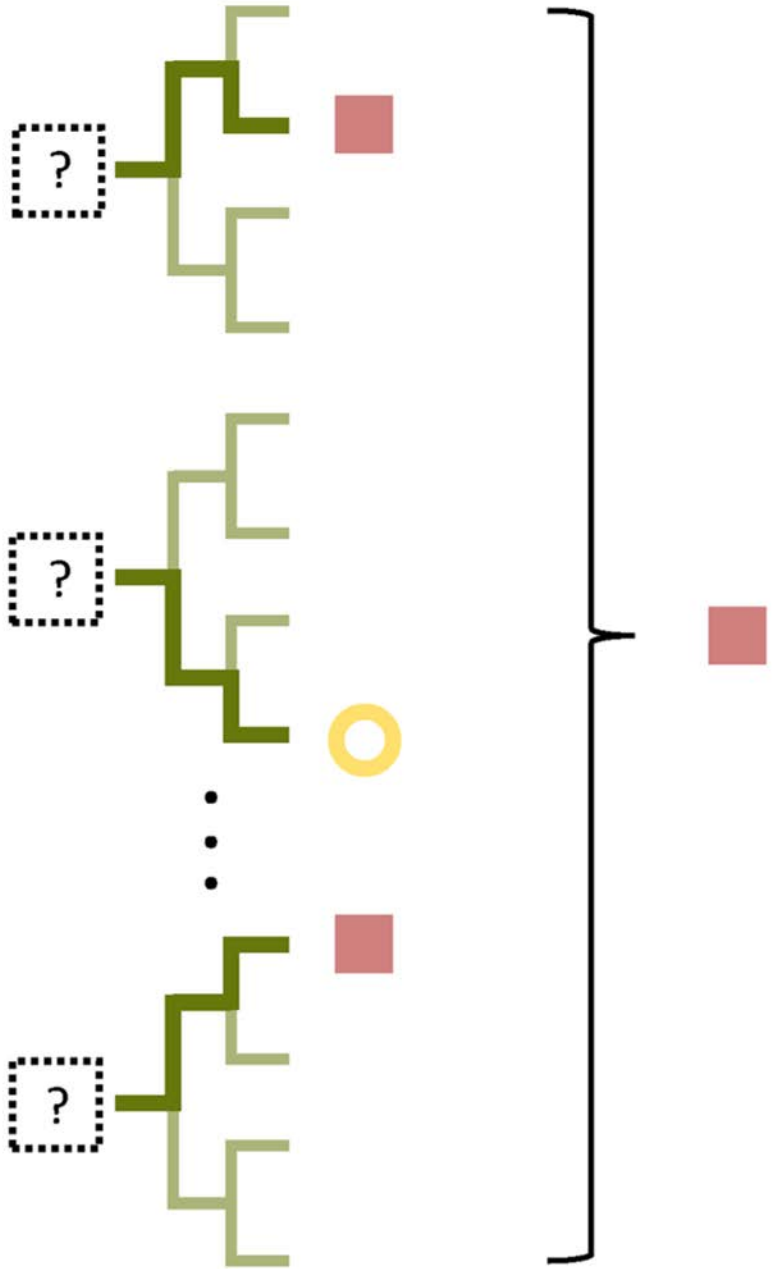
Eg. algorithm identify candidate epileptiform discharges by detecting outliers from the background EEG recording

Nature Reviews Neurology 2024;20:319-36

Overview of machine learning concepts

- Commonly used mapping functions
  - Random forest
  - K nearest neighbor (k-NN) classification
  - Support vector machine

B

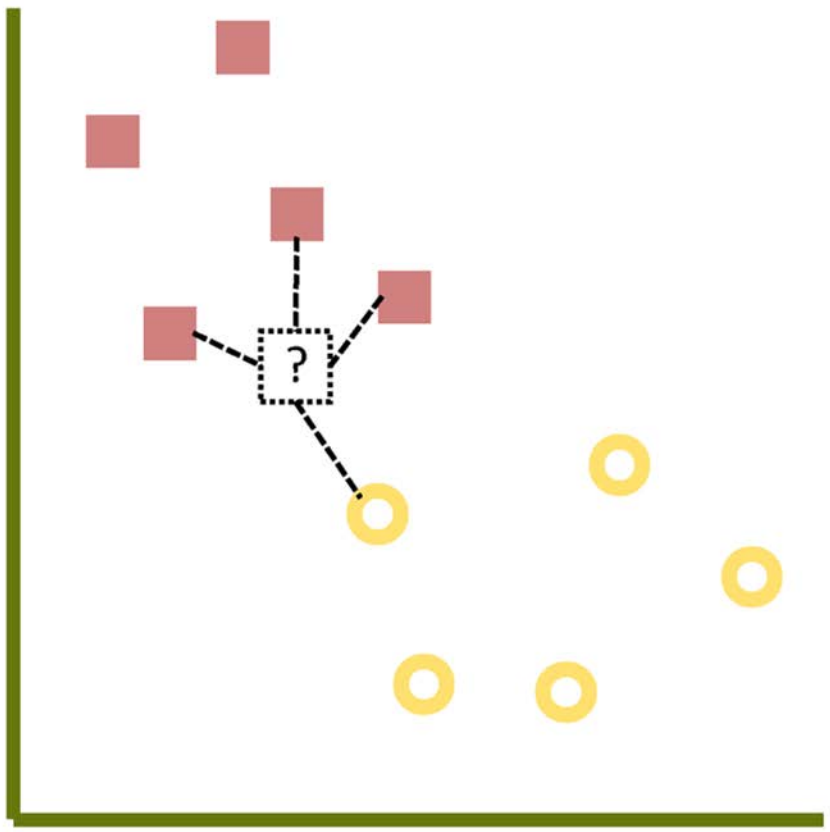


The random forest

Algorithm generates a forest of decision trees,
each utilizing subsets of input features
as bifurcation points to differentiate the
training data into expected outputs
the output of the ensemble (eg, the majority vote) is
reported for new inputs.

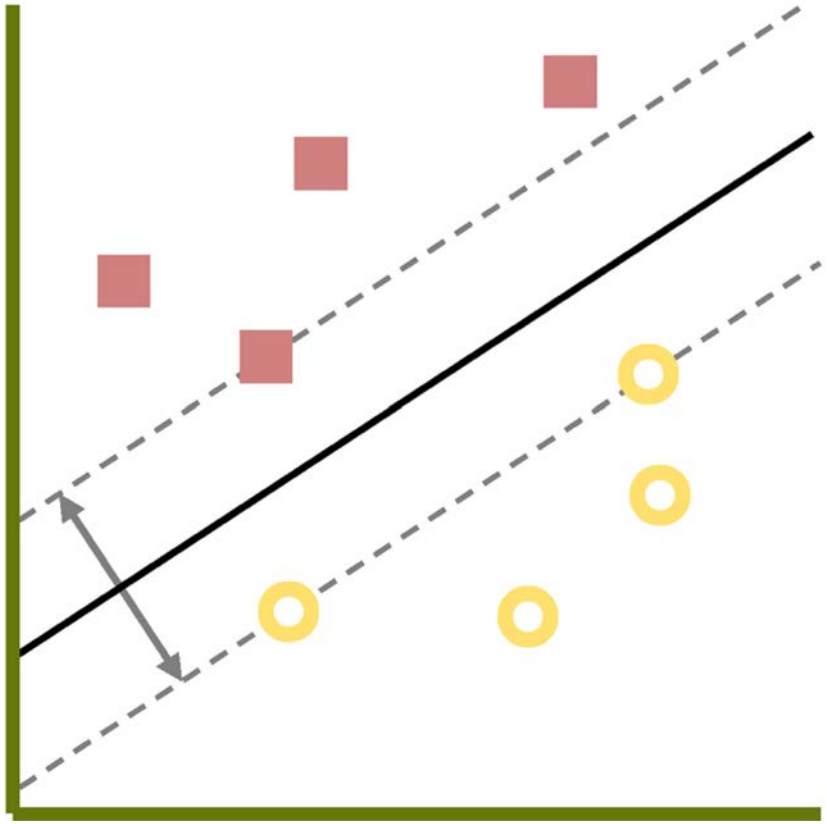
Nature Reviews Neurology 2024;20:319-36

c



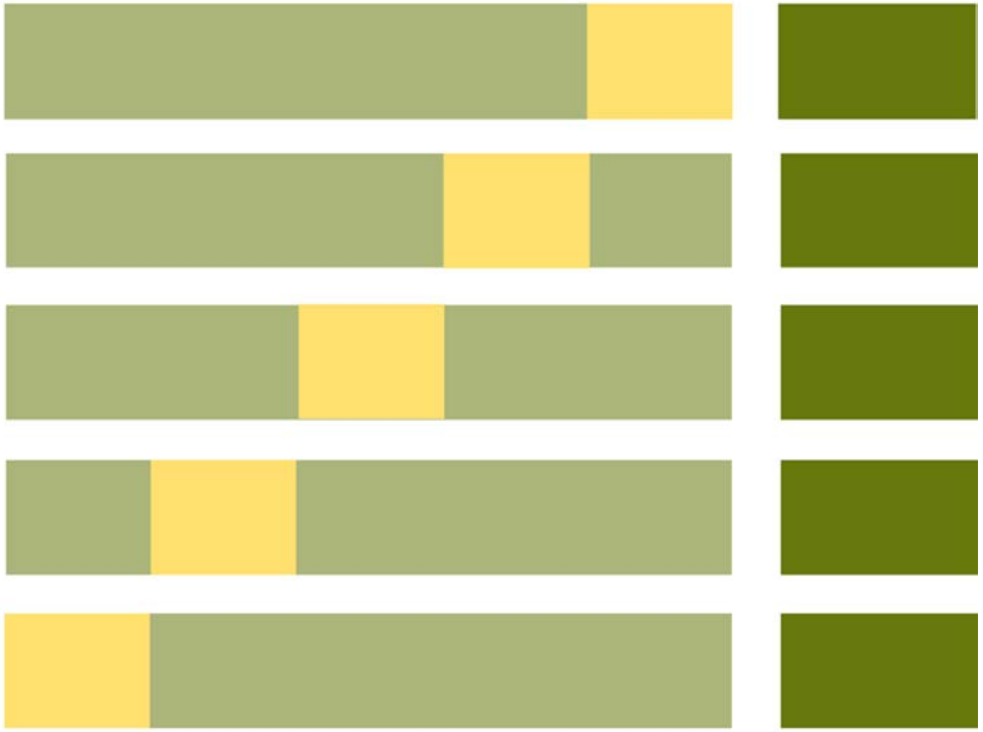
k-nearest neighbor classification
an input is plotted as a vector within a feature space alongside labeled data, and is subsequently assigned to the class of its k nearest neighbors (here, k = 4).

Nature Reviews Neurology 2024;20:319-36



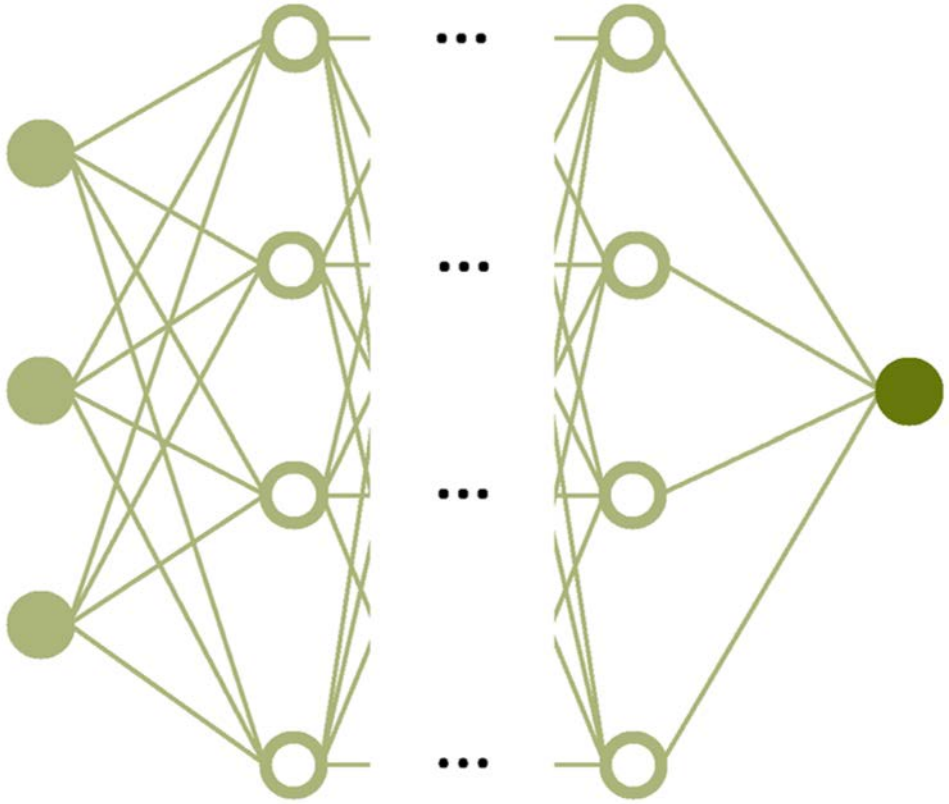
Support vector machine
generate a hyperplane in a higher-dimensional feature space to maximally separate clusters of labeled training data, providing a decision boundary for classifying new inputs

Nature Reviews Neurology 2024;20:319-36



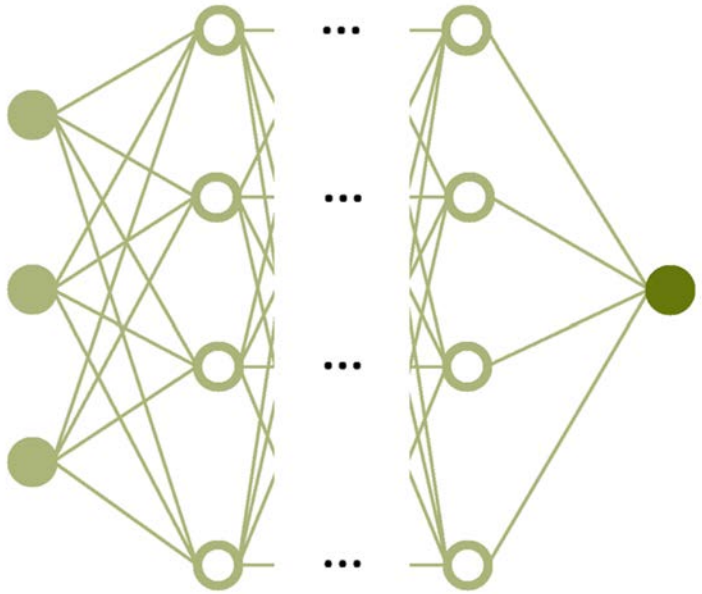
In cross-validation
a subset of the training data is withheld as the validation set (yellow),
allowing for fine-tuning of an algorithm parametrized on the training set (light green);
after multiple iterations (here showing K-fold cross-validation with K = 5),
the algorithm may be tested on an initially withheld testing set (dark green)
to assess accuracy and generalizability of the finalized model

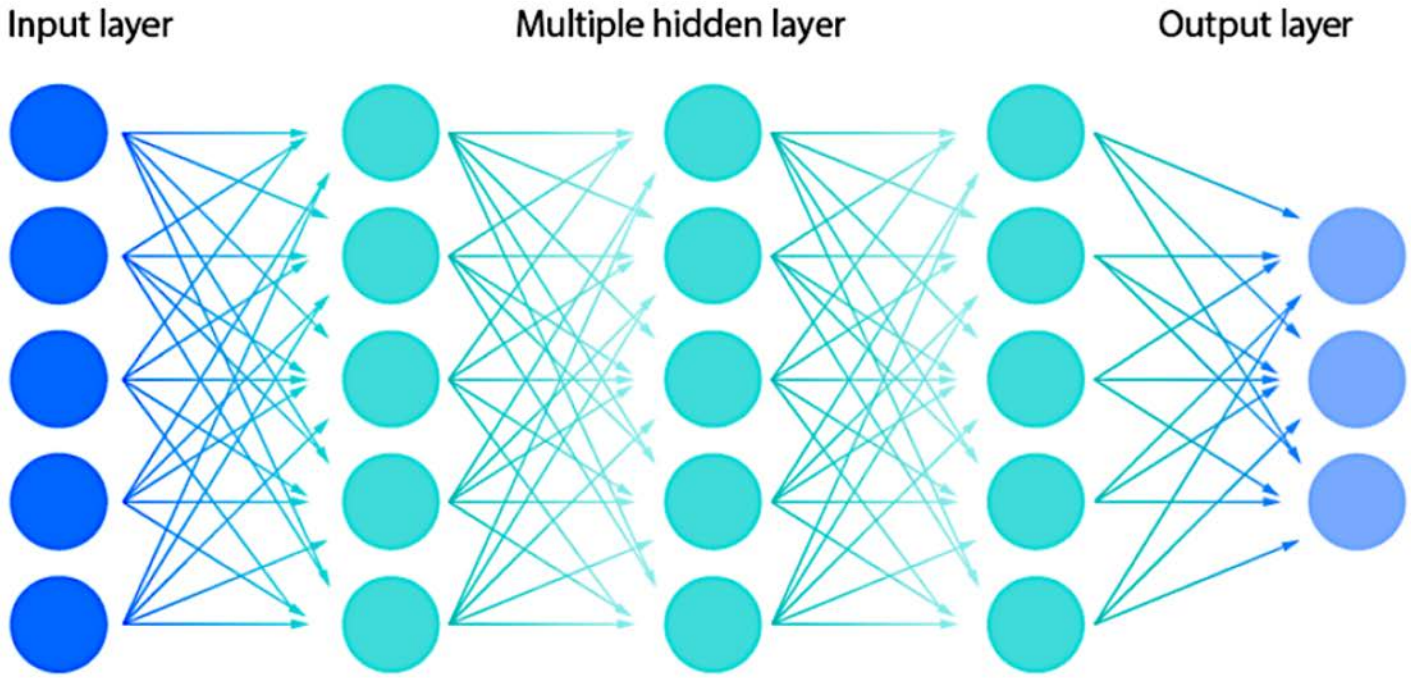
Nature Reviews Neurology 2024;20:319-36



Multilayers artificial neural networks
process data through layers of nodes, in each of which weighted inputs are summated and passed through a nonlinear activation function to yield intermediary outputs;
these may in turn proceed through additional layers of nodes as desired, ultimately reaching output nodes

Nature Reviews Neurology 2024;20:319-36





See video “Stat Quest” on YOUTUBE

Multilayers artificial neural networks
process data through layers of nodes, in each of which weighted inputs are summated and passed through a nonlinear activation function to yield intermediary outputs;
these may in turn proceed through additional layers of nodes as desired, ultimately reaching output nodes

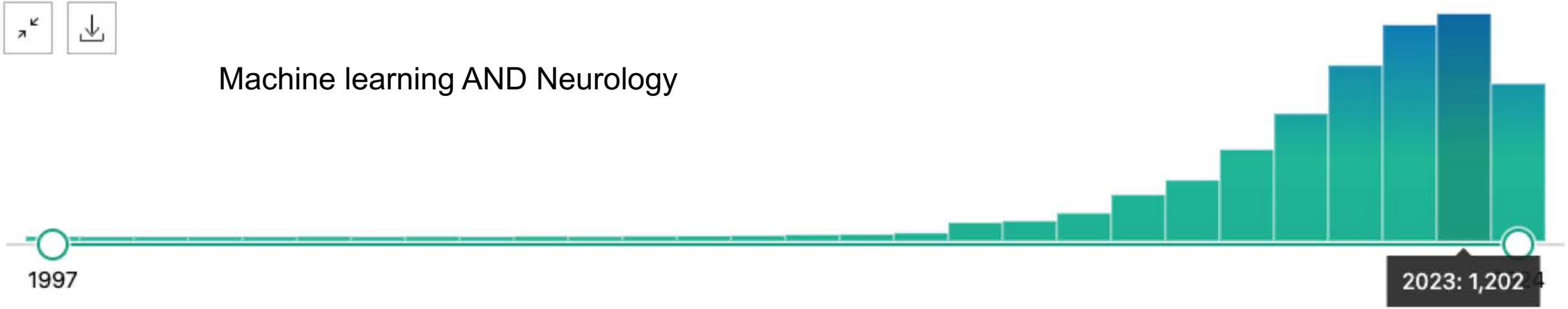
Nature Reviews Neurology 2024;20:319-36

RESULTS BY YEAR

5,398 results

Page 1 of 108

Machine learning AND Neurology



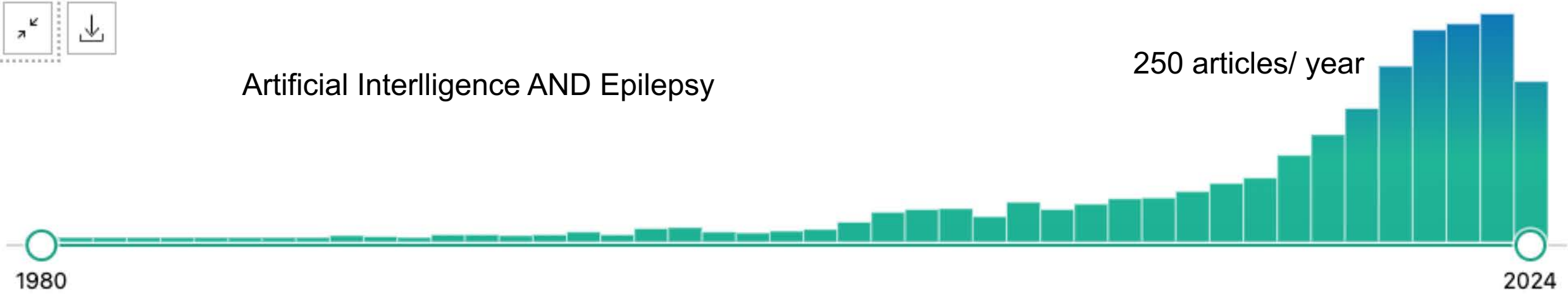
RESULTS BY YEAR

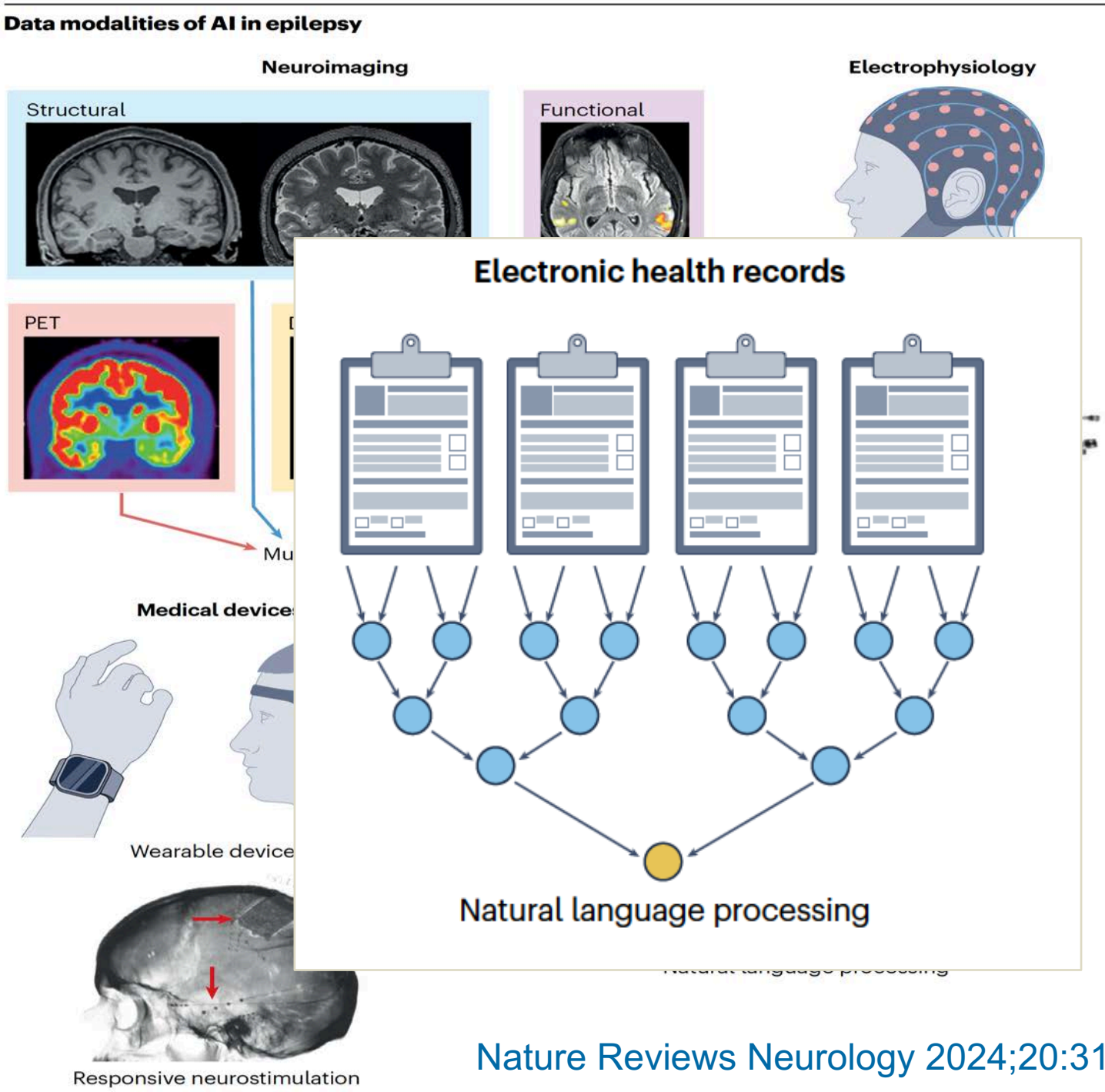
1,757 results

Page 1 of 36

Artificial Interlligence AND Epilepsy

250 articles/ year





Nature Reviews Neurology 2024;20:319-36

Received: 24 June 2023 | Accepted: 29 July 2023

DOI: 10.1002/epi4.12800

Epilepsia Open™ Open Access

ORIGINAL ARTICLE

Epilepsy classification using artificial intelligence: A web-based application

Ali A. Asadi-Pooya[1,2] | Davood Fattahi[1] | Nahid Abolpour[1] | Reza Boostani[3] | Mohsen Farazdaghi[1] | Mehrdad Sharifi[4,5,6]

To evaluate the feasibility of using easily accessible and applicable clinical information (based on history taking and physical examination) in order to make a reliable differentiation between idiopathic generalized epilepsy (IGE) versus focal epilepsy using machine learning (ML) methods.

Epilepsia Open. 2023;8:1362–1368

Data pipeline

INPUT DATA → **FEATURE SELECTION**
Input Data identified
- manually
- based on expert level knowledge
- by algorithm

→ Analyze using **MAPPING FUNCTION** that generate output prediction → **OUTPUT PREDICTION**

All patients with an electro-clinical diagnosis of IGE or focal epilepsy, at the outpatient epilepsy clinic at Shiraz University, Shiraz, Iran, from 2008 until 2022, were included

1445 patients; 964 with focal epilepsy and 481 with IGE

→ The first author selected a set of clinical features

→ Different types of classifiers were assessed and the final classification was made based on their best results using the stacking method

First phase of the study was a retrospective study of a prospectively developed and maintained database.

Epilepsia Open. 2023;8:1362–1368

Feature Selection

The first author selected a set of clinical features that are
(1) easily obtainable even by people who are not experts in the field and
(2) helpful in making a diagnosis of epilepsy type/syndrome (differentiating focal epilepsy from IGE) based on the previous literature.

Other clinical features [eg, an exact diagnosis of seizure types (eg, focal seizure with impaired awareness vs absence seizures)] that are very helpful in differentiating focal epilepsy from IGE, but need a skillful expert were not included

The study did not include EEG and imaging findings.

Epilepsia Open. 2023;8:1362–1368

Epilepsy Classifier: IGE vs Focal

Age at onset: (years, e.g. 12, 13, etc.)

Sex: (1 for Male, 2 for Female)

Febrile convulsion: (1 for Yes, 2 for No)

Family history of epilepsy: (1 for Yes, 2 for No)

Major head injury: (1 for Yes, 2 for No)

Medical comorbidity: (1 for Yes, 2 for No)

Aura: (an integer from 1 to 17, see HERE)

Exam: (1 for Normal, 2 for Abnormal)

Tongue biting: (1 for Yes, 2 for No)

Classify

Description:

Project details: The present online application aims to utilize clinical information of patients with epilepsy (PWE) to differentiate focal epilepsy from idiopathic generalized epilepsy (IGE) by application of machine learning methods. Nine easily obtainable clinical features (based on a detailed history and physical examination) are utilised as the inputs. The classification framework benefits from multiple classifiers and their best results are exploited by a Stacking classifier to perform the final classification. The training procedure is carried out on a large database of PWE built over 14 years at the epilepsy center at Shiraz University of Medical Sciences, Iran, from 2008 until 2022. More technical details can be found in the related publication.

Input parameters: including age at seizure onset, sex, a history of febrile convulsion, a family history of epilepsy, a history of severe head injury, a history of medical comorbidity, aura with seizures, ictal-related tongue biting, and abnormal physical examination.

Aura types: 1 = No aura, 2 = Indescribable feeling, 3 = Dizziness, 4 = Fear / Nervousness / Anxiety / Adrenaline rush, 5 = Cognitive / Deja vu / Jamais vu / Forced thinking, 6 = Epigastric / Abdominal / Nausea, 7 = Elementary visual, 8 = Complex visual, 9 = Elementary auditory, 10 = Complex auditory, 11 = Olfactory, 12 =

Aura types: 1 = No aura, 2 = Indescribable feeling, 3 = Dizziness, 4 = Fear / Nervousness / Anxiety / Adrenaline rush, 5 = Cognitive / Deja vu / Jamais vu / Forced thinking, 6 = Epigastric / Abdominal / Nausea, 7 = Elementary visual, 8 = Complex visual, 9 = Elementary auditory, 10 = Complex auditory, 11 = Olfactory, 12 = Gustatory / Taste, 13 = Left focal sensory, 14 = Right focal sensory, 15 = Other sensory, 16 = Headache, 17 = Other.

Epilepsia Open. 2023;8:1362–1368

http://www.epiclass.ir/f-ige.

	Precision			Sensitivity			Specificity			F1-score		
Classifiers	FE	IGE	Avg	FE	IGE	Avg	FE	IGE	Avg	FE	IGE	Avg
Stack	0.87	0.71	0.81	0.85	0.74	0.81	0.74	0.85	0.77	0.86	0.72	0.81
SVM	0.83	0.68	0.78	0.85	0.66	0.79	0.66	0.85	0.72	0.84	0.67	0.78
LogReg	0.83	0.67	0.78	0.84	0.66	0.78	0.66	0.84	0.72	0.84	0.67	0.78
KNN	0.87	0.66	0.80	0.76	0.80	0.78	0.76	0.80	0.78	0.84	0.71	0.79
RanFor	0.85	0.69	0.80	0.84	0.70	0.79	0.70	0.84	0.77	0.85	0.69	0.80
GradBoost	0.88	0.68										
AdaBoost	0.87	0.69										
Bagging	0.86	0.68										
ExtRa Trees	0.82	0.71	0.78	0.89	0.60	0.79	0.60	0.89	0.70	0.85	0.66	0.79

Note: Each row represents a classifier while their precision, sensitivity, specificity, and F1-score are in the columns for focal epilepsy (FE), idiopathic generalized epilepsy (IGE), and their average.

Also, in order to enable and facilitate future external validation studies by other peers and professionals, the developed and trained ML model was implemented and published via an online web-based application that is freely available at http://www.epiclass.ir/f-ige.

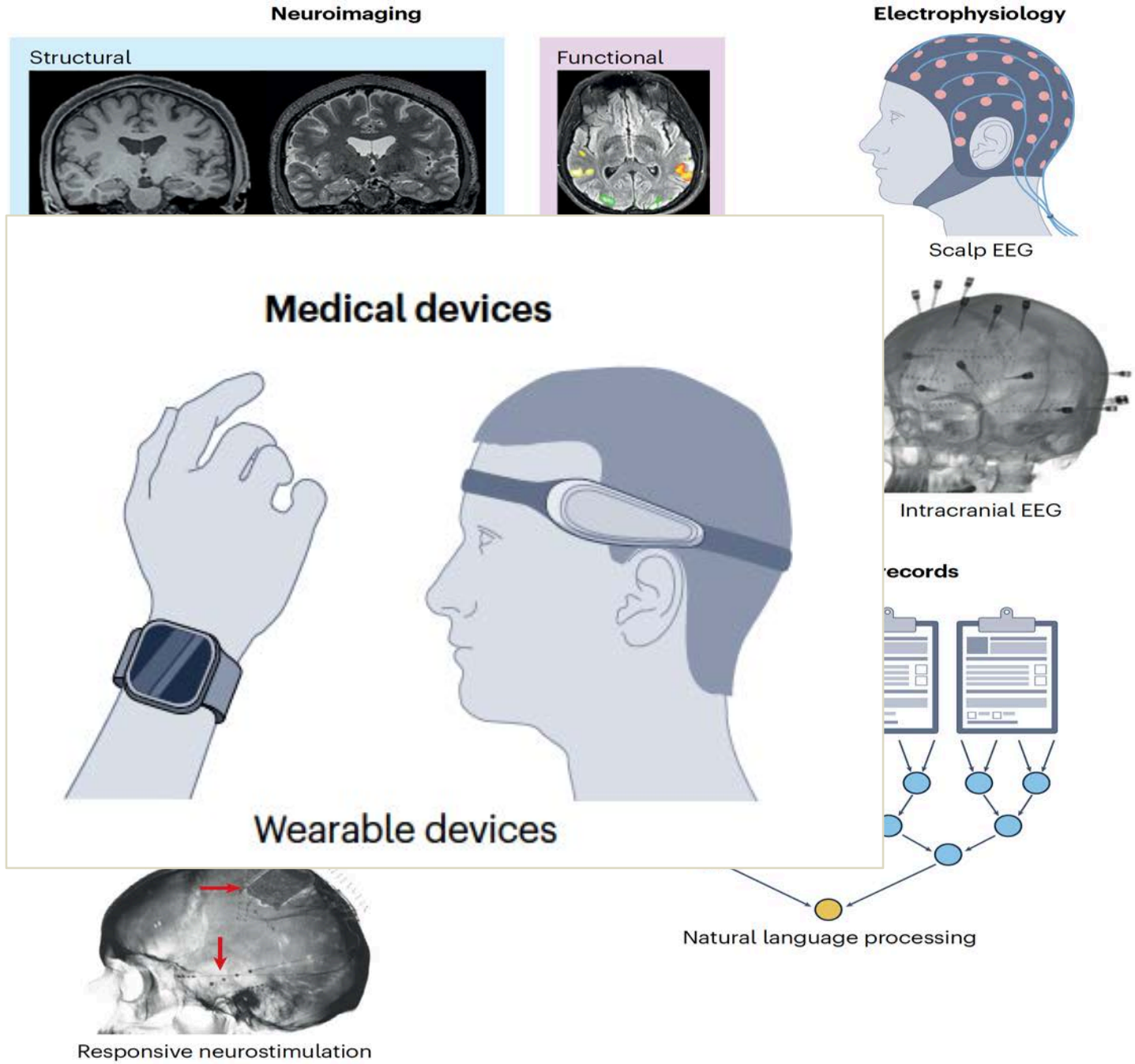
This study developed a pragmatic algorithm aimed at epilepsy classification (IGE vs focal epilepsy) for individuals whose epilepsy begins at age 10 years and older

The algorithm has the precision: 0.81, sensitivity: 0.81, and specificity: 0.77.

This algorithm is that it could be used by people who are not experts in epilepsy diagnosis (eg, internists, etc.)

Epilepsia Open. 2023;8:1362–1368

Data modalities of AI in epilepsy



Received: 30 July 2022 | Accepted: 30 September 2022

DOI: 10.1111/ane.13716

Acta Neurologica Scandinavica

WILEY

REVIEW ARTICLE

Seizure detection based on wearable devices: A review of device, mechanism, and algorithm

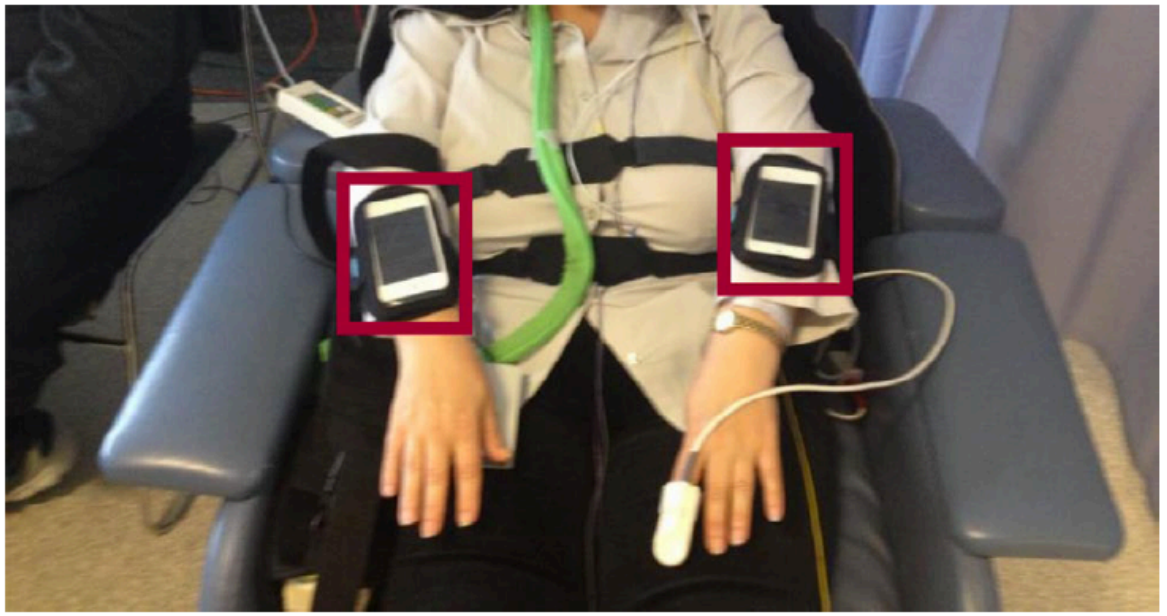
Wen Li[1] | Guangming Wang[1] | Xiyuan Lei[1] | Duozheng Sheng[1] | Tao Yu[2] | Gang Wang[1]

Li W, et al. Acta Neurol Scand. 2022;146:723–731

Data pipeline

INPUT DATA

FEATURE SE
Input Data identified
- manually
- based on e
- by algorith



OUTPUT PREDICTION

ACM: acceralometer
sEMG
EKG
EDA: electrodermal activity
PPG: photoplethysmography
EEG: behind ears

Input data selected based on different criteria
: Time
Frequency

Different types of classifiers
KNN
SVM
RF
gradient tree boosting

Epilepsia Open. 2019;4:309–317

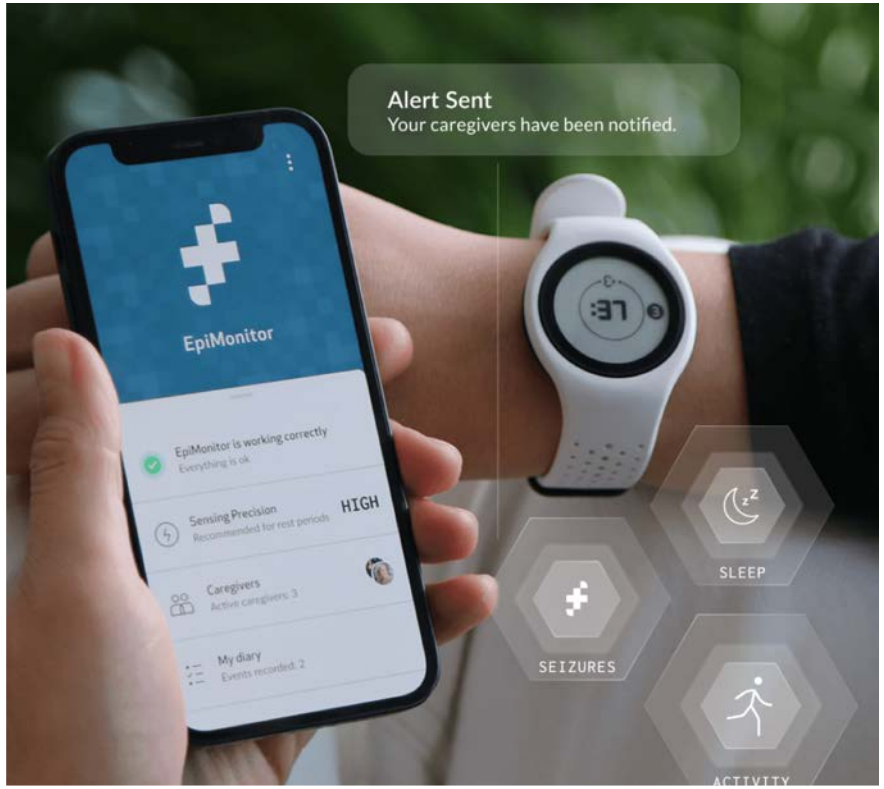
First author	Subjects (*n*)	Seizures	Signal	Feature set	Algorithm	SEN (%)	FDR (/h)
Johansson[56]	11	37 TCS	ACM	Time, frequency	KNN	100	0.05
Dong[57]	5	379 GTCS	ACM	Time, frequency	RF	88.01	0.01
Dong[58]	7	547 MS	ACM	Time, frequency	two-layer ensemble model	76.84	0.04
Conradsen[33]	11	22 GTCS	sEMG	Time	threshold	100	0.04
Beniczky[23]	11	32 GTCS	sEMG	Time	threshold	93.80	0.03
Baumgartner[59]	20	47 MS	sEMG	Frequency, time–frequency	threshold	72	/
You[60]	12	56 MS	behind-the-ear EEG	Time–frequency	GAN, threshold	96.30	0.14
Frankel[25]	10	24 FS	behind-the-ear EEG	Time, frequency, time–frequency, nonlinear	RF	90	0.09
You[61]	16	52 MS	behind-the-ear EEG	frequency	VAE based on RNN	90.40	0.83
Vandecasteele[64]	11	47 FS	ECG, PPG	Time	SVM	70	2.11
Forooghifar[65]	18	154 FS	ECG	Time, frequency, nonlinear	RF		
Cooman[66]	24	227 FS	ECG	Time	SVM		
Baghersalimi[67]	29	277 FS, FTCb	ECG	Time	Res1DCNN		
Poh[18]	7	16 GTCS	ACM, EDA	Time, frequency, nonlinear	SVM	94	0.03
Milošević[68]	7	22 GTCS	ACM, sEMG	Time, frequency, time–frequency, nonlinear	SVM	91	0.04
Onorati[19]	22	55 MS	ACM, EDA	Time, frequency, nonlinear	SVM	94.55	0.01
Vandecasteele[63]	135	896 FS	behind-the-ear EEG, ECG	Time, frequency, nonlinear	RF	92	1.85
Böttcher[69]	10	21 TCS	ACM, EDA	Time, frequency	Gradient tree boosting	91	0.01
Nasseri[62]	10	19 MS	ACM, EDA, PPG, temperature	Time	LSTM	93	0.10
Böttcher[70]	9	20 MS	ACM, EDA, PPG	Time, frequency	Gradient tree boosting	75	0.56

Abbreviations: FMS, focal motor seizure; FS, focal seizure; FTCb, focal to bi

Sensitivity of 72-100%
FDR of 0.01-2.11/hr

FDR= false discovery rate

Li W, et al. Acta Neurol Scand. 2022;146:723–731



Embrace2

: monitor ACM, EDA, PPG and temperature

SPEAC® System
Brain Sentinel® Monitoring and Alerting System

non-EEG physiological signal-based seizure monitoring system: record sEMG

ACM: acceralometer, EDA: electrodermal activity, PPG: photoplethysmography

TABLE 1 Wearable seizure monitor available in the market

Device	Company	Wear site	Signal	Targeted seizure	Certificate	Battery life
Smart Watch[13,14]	SmartMonitor	Wrist/ankle	ACM	CS, FMS	Not reported	30 h
Epi-Care[16,17]	Danish Care Technology	Wrist	ACM	GTCS	CE	24 h
Embrace2[21]	Empatica	Wrist	ACM, EDA, PPG and temperature	GTCS	FDA and CE	48 h+
Epilert[22]	Tekru Technologies	Wrist	ACM, EDA, PPG and temperature	TCS	Not reported	48 h
EDDI[23,33]	IctalCare	Brachial biceps muscles	sEMG	GTCS	CE	8 h
SPEAC[23,34]	Brain Sentinel	Biceps and triceps brachii	sEMG	GTCS	FDA	12 h
ePatch[24]	BioTelemetry	Left ribs	ECG	CS and non-CS with autonomic changes	Not reported	72 h
Epilog[25]	Epitel	Scalp below hairline	EEG	Depend on paired software	No FDA	168 h
Sensor Dot[26]	Byteflies	Behind ear (optional)	EEG	Typical absence	CE	24 h
Nightwatch[27,35]	LivAssured BV	Armband	ACM and PPG	CS at night	CE	–
IMEC[36,37]	imec/Holst Centre	Armband and patches on chest	ACM, EDA, ECG, and sEMG	FS	Not reported	96 h

Abbreviations: CS, convulsive seizure; FMS, focal motor seizure; FS, focal seizure; TCS, tonic–clonic seizure.

Li W, et al. Acta Neurol Scand. 2022;146:723–731

Received: 31 March 2020 | Revised: 23 April 2020 | Accepted: 27 April 2020

DOI: 10.1111/epi.16541

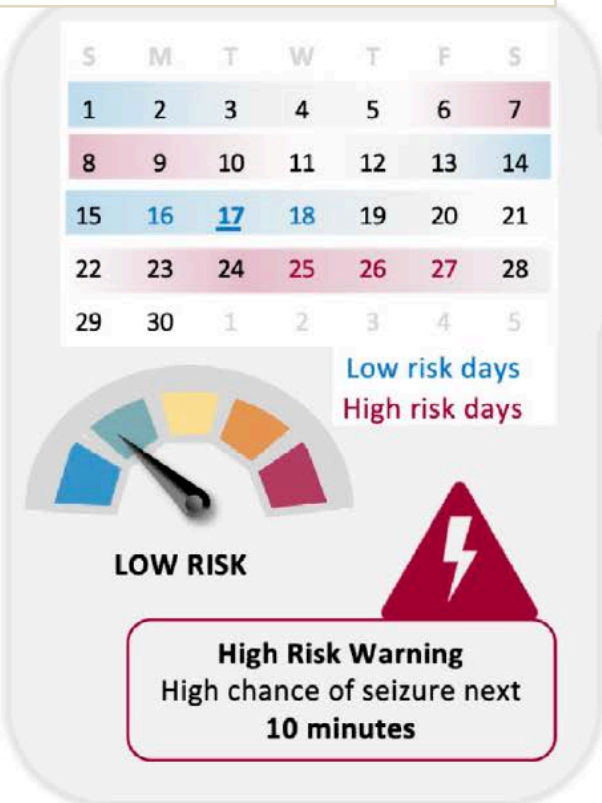
SUPPLEMENT ARTICLE

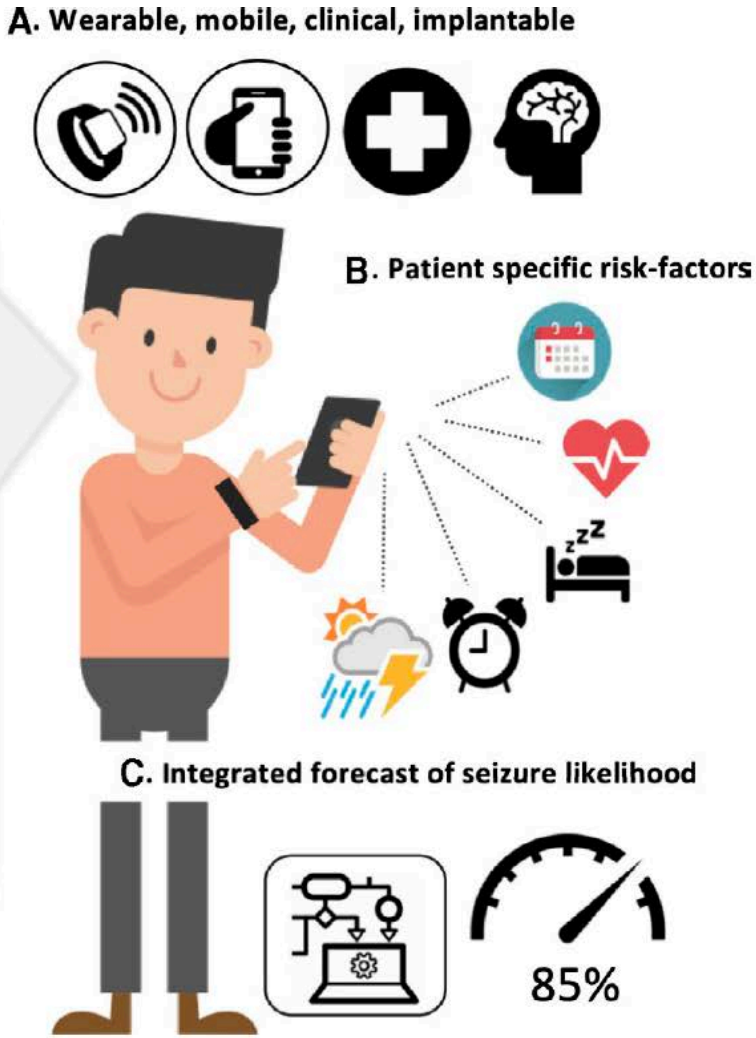
Epilepsia®

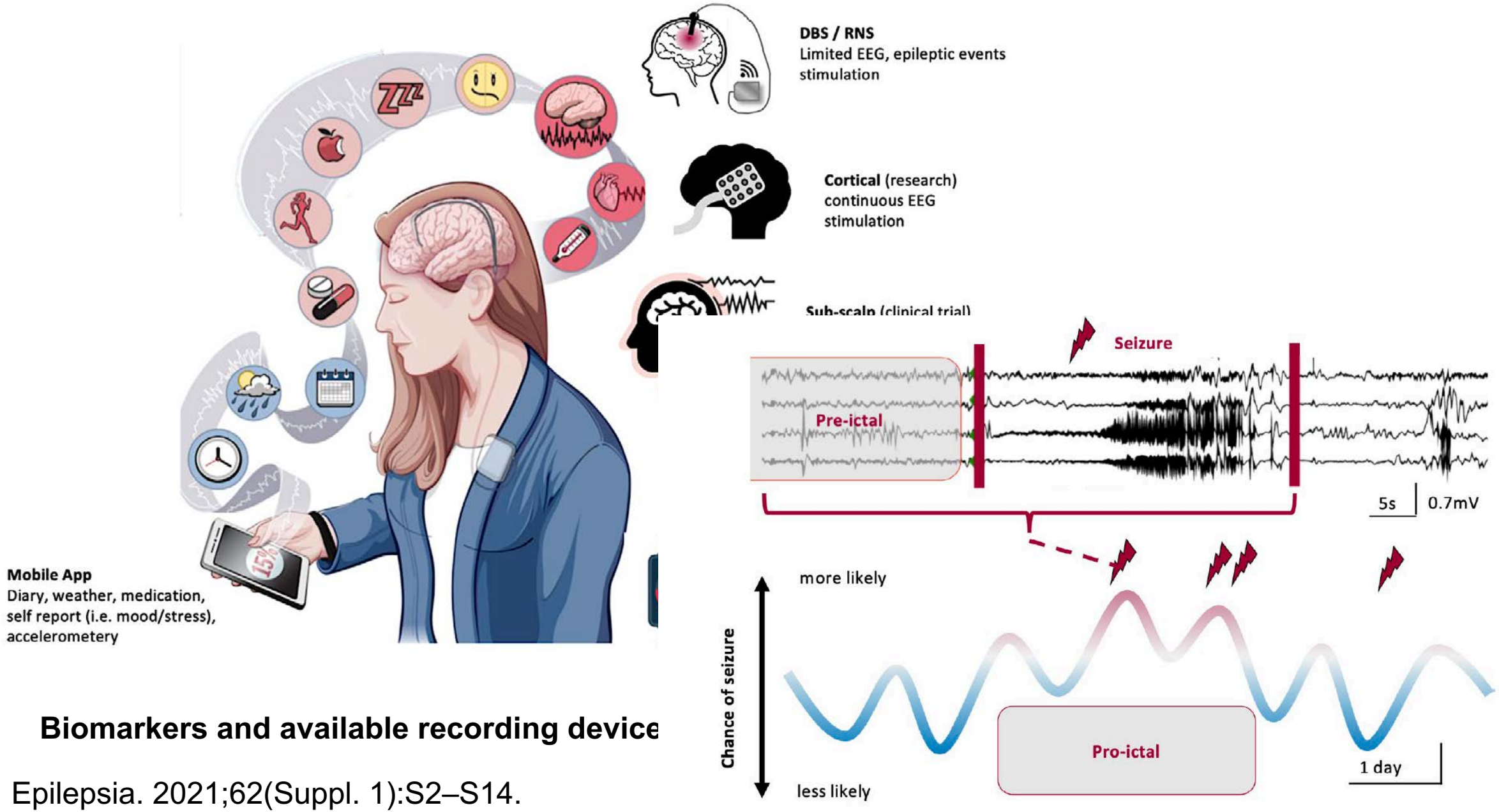
Seizure forecasting and cyclic control of seizures

Rachel E. Stirling[1] | Mark J. Cook[2] | David B. Grayden[1] | Philippa J. Karoly[1,2]

Epilepsia. 2021;62(Suppl. 1):S2–S14.

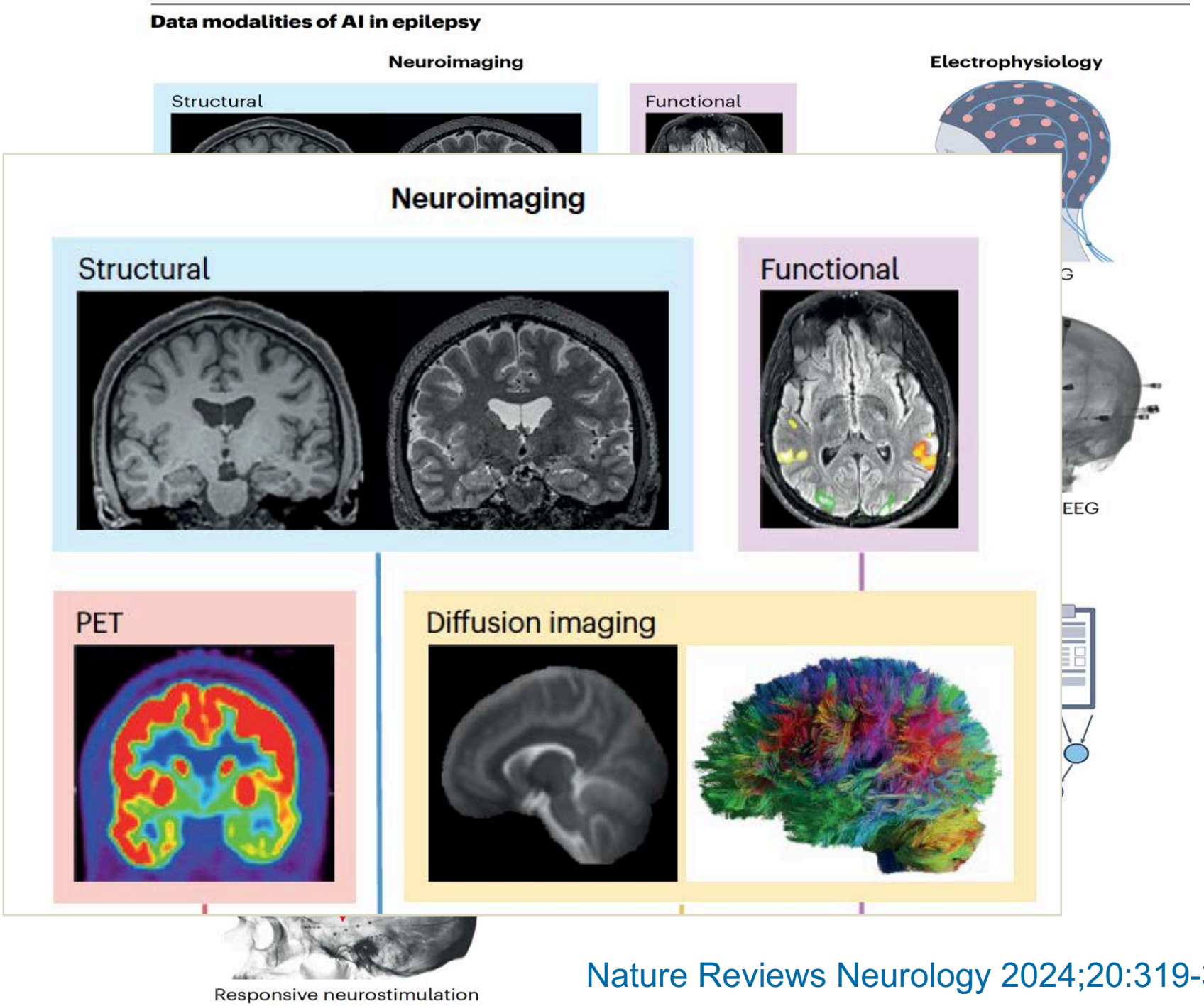






Biomarkers and available recording device

Epilepsia. 2021;62(Suppl. 1):S2–S14.



Nature Reviews Neurology 2024;20:319-36

RESEARCH ARTICLE

Multicenter Validation of a Deep Learning Detection Algorithm for Focal Cortical Dysplasia

Ravnoor Singh Gill, PhD-cand, Hyo-Min Lee, PhD-cand, Benoit Caldairou, PhD, Seok-Jun Hong, PhD, Carmen Barba, MD, Francesco Deleo, MD, Ludovico D'Incerti, MD, Vanessa Cristina Mendes Coelho, MD, Matteo Lenge, PhD, Mira Semmelroch, PhD, Dewi Victoria Schrader, MD, Fabrice Bartolomei, MD, Maxime Guye, MD, PhD, Andreas Schulze-Bonhage, MD, Horst Urbach, MD, Kyoo Ho Cho, MD, Fernando Cendes, MD, PhD, Renzo Guerrini, MD, Graeme Jackson, MD, R. Edward Hogan, MD, Neda Bernasconi, MD, PhD,* and Andrea Bernasconi, MD*

Correspondence
Dr. Bernasconi
andrea.bernasconi@mcgill.ca

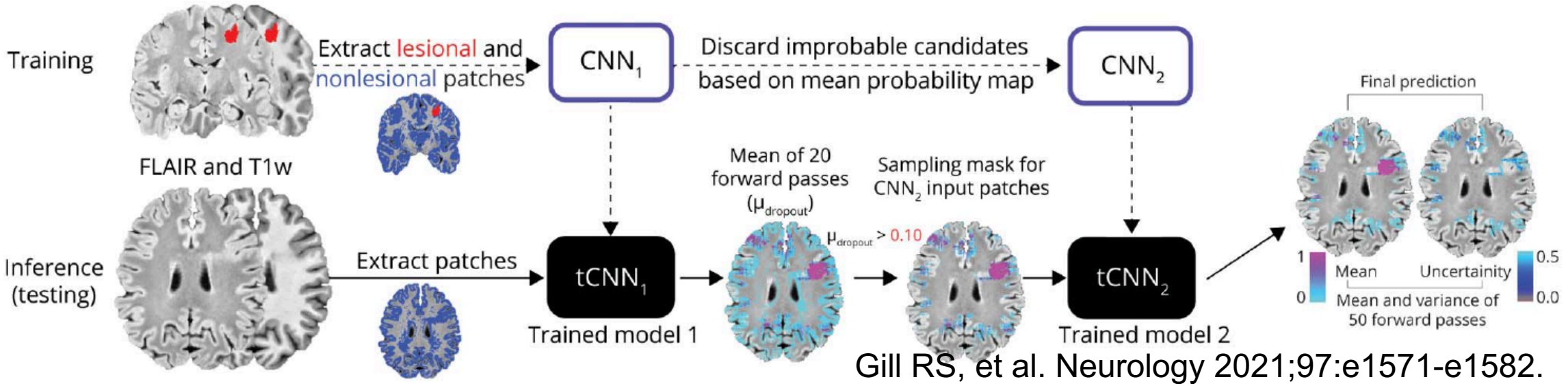
Neurology® 2021;97:e1571-e1582. doi:10.1212/WNL.0000000000012698

Gill RS, et al. Neurology 2021;97:e1571-e1582.

INPUT DATA → FEATURE SELECTION → Analyze using **MAPPING FUNCTION** → **OUTPUT PREDICTION**

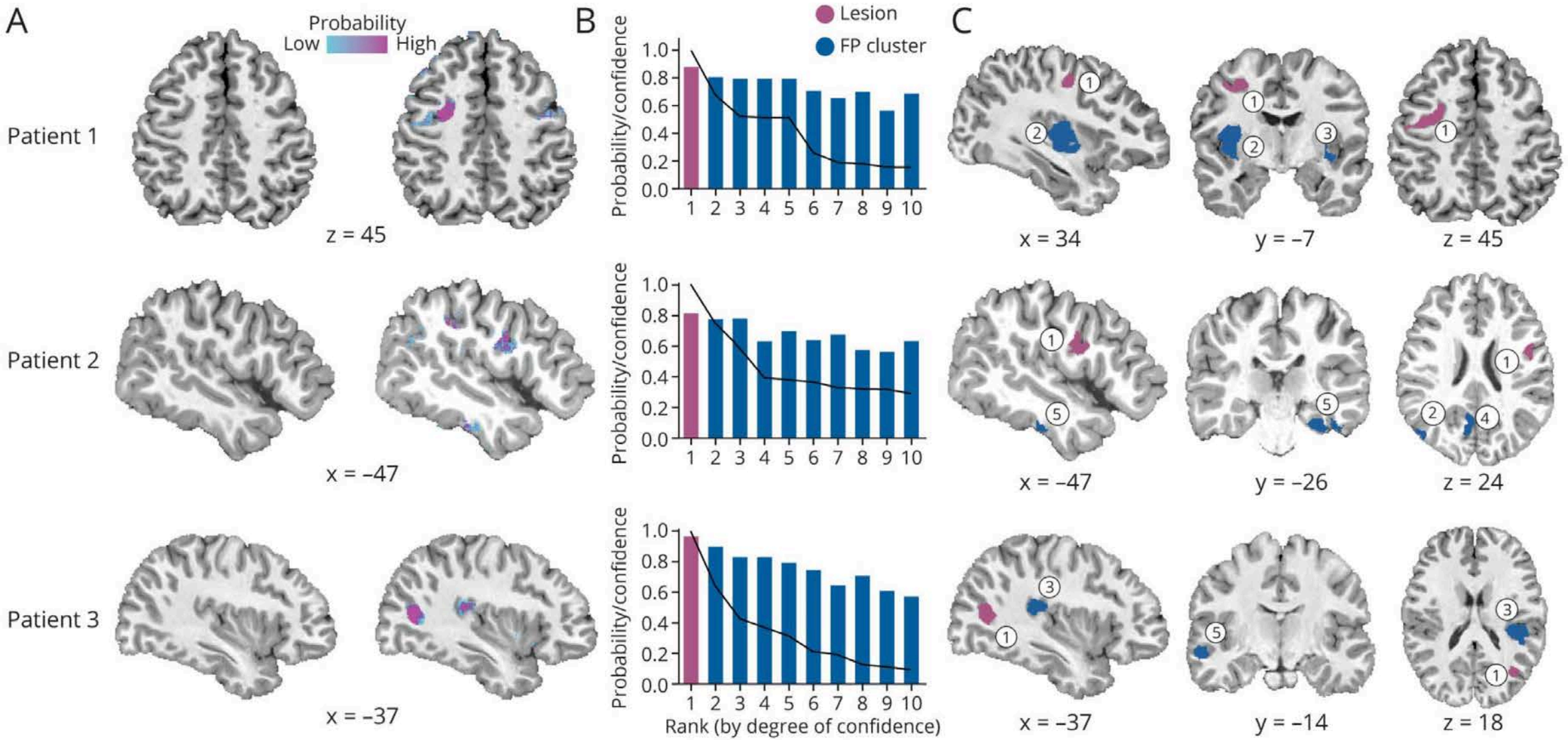
Data pipeline

3D T1-weighted and 3D FLAIR MRI of 148 patients with histologically verified FCD at 9 centers → Feature selection : image processing → Deep convolutional neural network (CNN) classifier



Gill RS, et al. Neurology 2021;97:e1571-e1582.

To evaluate performance, detection maps were compared to expert FCD manual labels.
Sensitivity was tested in an independent cohort of 23 cases with FCD (13 ± 10 years).
Applying the algorithm to 42 healthy controls and 89 controls with temporal lobe epilepsy tested specificity.



(A) T1-weighted MRI and prediction probability maps with the lesion circled. (B) Probability of the lesion and false-positive (FP) clusters sorted by their rank; superimposed line indicates the degree of confidence for each cluster. (C) Location of the focal cortical dysplasia (FCD) lesion (rank 1, highest confidence; purple) and FP clusters (ranks 2–5; blue). In these cases, the lesion has both the highest confidence (rank 1) and high probability (>0.8). S = site.

Gill RS, et al. Neurology 2021;97:e1571-e1582.

Site	No.	Age, mean ± SD, y	Female, %	MRI+/MRI−, n	Sensitivity, n (%)		FPs
					All patients	MRI−	
S1-I	45	27 ± 9	49	13/32	39/45 (87)	26/32 (81)	7 ± 4
S1-II	17	18 ± 9	65	2/15	15/17 (88)	13/15 (87)	7 ± 4
S2	08	11 ± 6	25	5/3	8/8 (100)	3/3 (100)	6 ± 5
S3	05	22 ± 17	80	2/3	5/5 (100)	3/3 (100)	1 ± 1
S4	11	8 ± 7	36	11/0	11/11 (100)	NA	8 ± 6
S5-I	10	23 ± 14	30	8/2	9/10 (90)	1/2 (50)	10 ± 6
S5-II							6 ± 7
S6							3 ± 3
S7							8 ± 6
S8							6 ± 5
S9							1 ± 2
Total	148	23 ± 13	47	49/51%	137/148 (93)	64/75 (85)	6 ± 5
Independent	23	13 ± 10	48	30/70%	19/23 (83)	12/16 (75)	5 ± 3

- The overall sensitivity of the classifier cross-validation was 93% (137 of 148 FCD lesions detected), with 6 ± 5 FP clusters per patient
- 85% of MRI-negative and 100% of MRI-positive lesions were detected.

When the classifier was tested on the independent cohort

- overall sensitivity was 83% (19 of 23 FCD lesions detected, 5 ± 3 FP clusters per patient)
- 75% of MRI-negative lesions and 100% of MRI-positive detected

Abbreviations: FPs: false positive rate per cohort; NA = not applicable; S = site.
I and II refer to different MRI scanners for the same site. Independent refers to validation cohort from S1 and S2. Number refers to sample size.

Gill RS, et al. Neurology 2021;97:e1571-e1582.

RESEARCH ARTICLE

Convolutional Neural Network Algorithm to Determine Lateralization of Seizure Onset in Patients With Epilepsy

A Proof-of-Principle Study

Erik Kaestner, PhD,* Jun Rao, MS,* Allen J. Chang, MS, Zhong Irene Wang, PhD, Robyn M. Busch, PhD, Simon S. Keller, PhD, Theodor Rüber, MD, Daniel L. Drane, PhD, Travis Stoub, PhD, Ezequiel Gleichgerrcht, Leonardo Bonilha, MD, PhD, Kyle Hasenstab, PhD,† and Carrie McDonald, PhD†

Correspondence
Dr. McDonald
camcdonald@
health.ucsd.edu

Neurology® 2023;101:e324-e335. doi:10.1212/WNL.0000000000207411

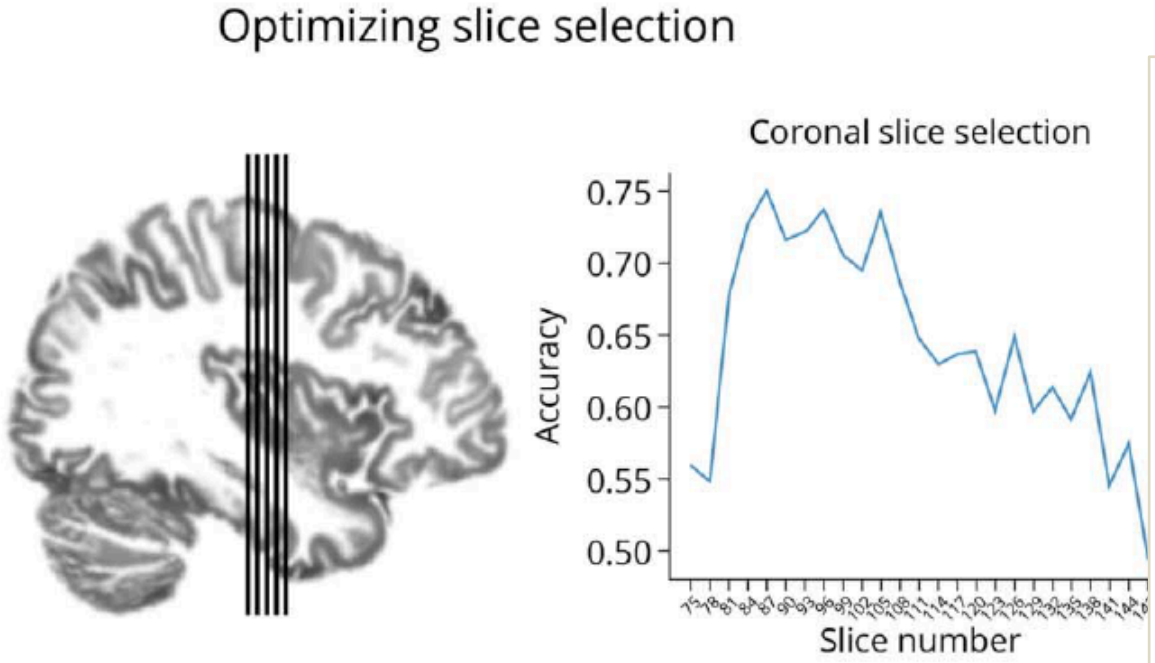
Using a dataset of 359 patients with temporal lobe epilepsy (TLE) from 7 surgical centers
Tested whether aCNN-based on T1-weighted images could classify seizure laterality concordant with clinical team consensus.
This CNN was compared with a randomized model (comparison with chance) and
a hippocampal volume logistic regression (comparison with current clinically available measures).

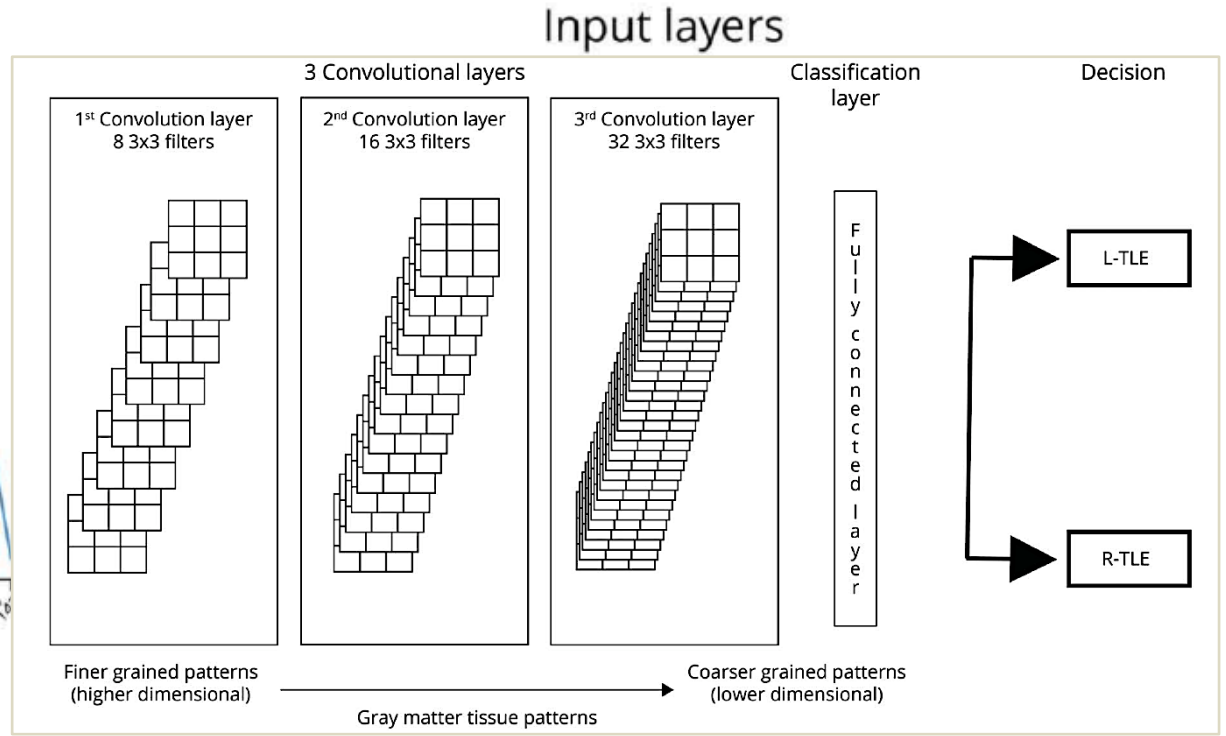
Neurology 2023;101:e324-e335.

INPUT DATA → FEATURE SELECTION → Analyze using **MAPPING FUNCTION** → **OUTPUT PREDICTION**

Data pipeline

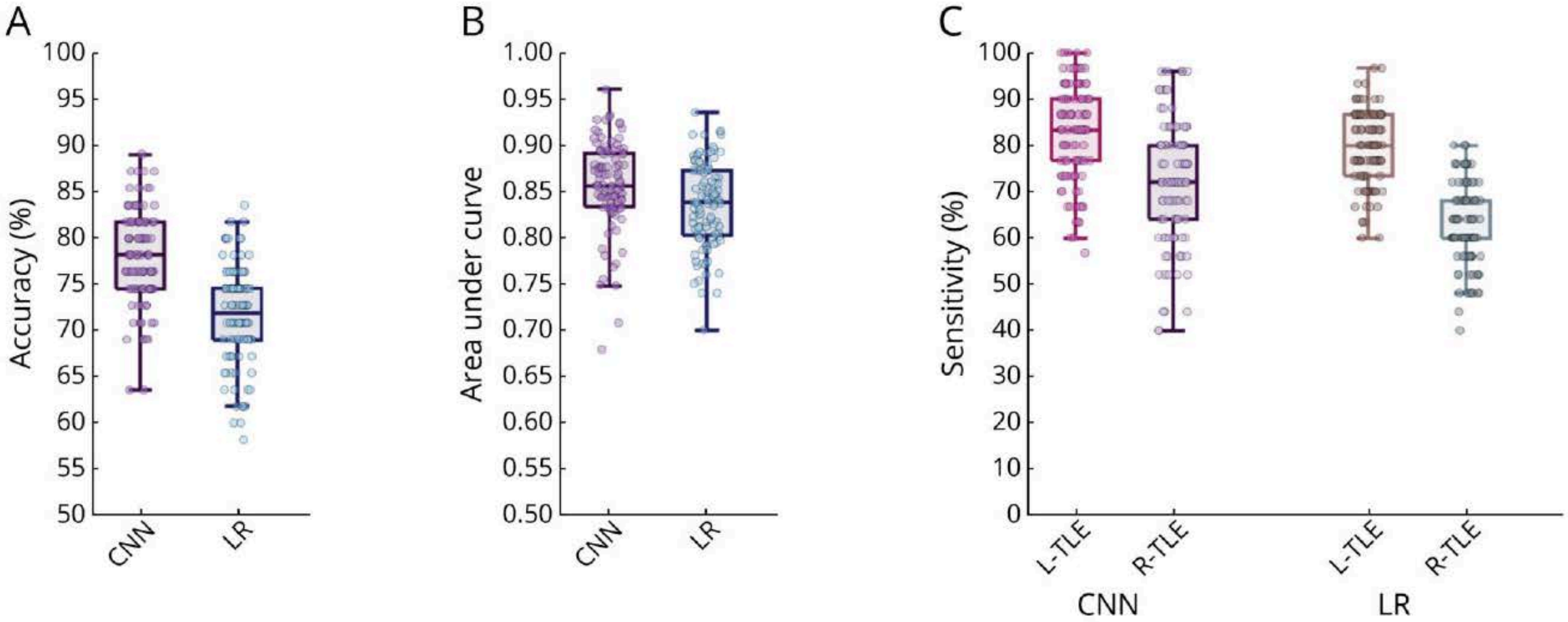
Dataset of 359 patients with temporal lobe epilepsy (TLE) from 7 surgical centers → Feature selection : image processing → Deep convolutional neural network (CNN) classifier





Neurology 2023;101:e324-e335.

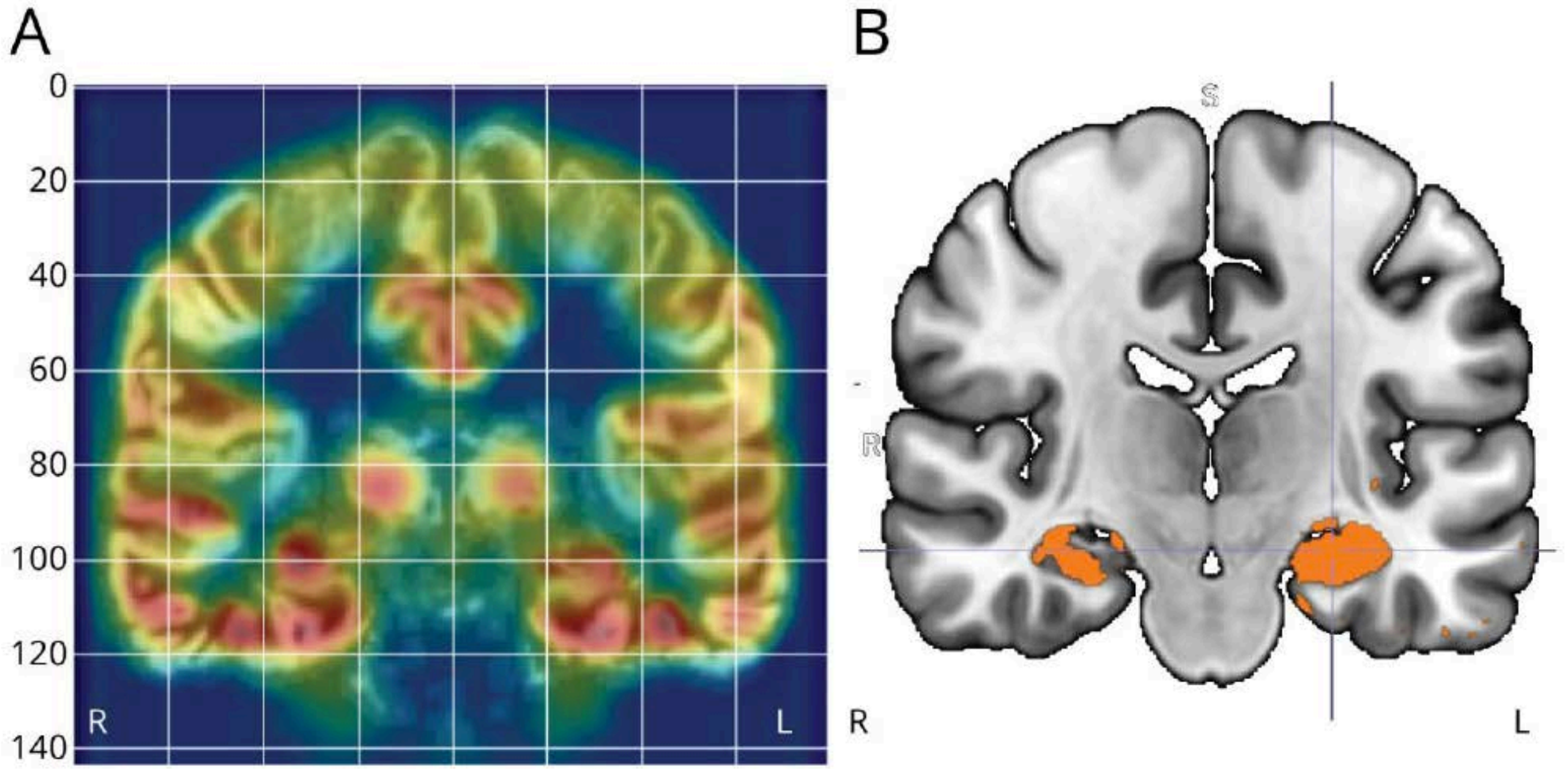
CNN Model vs Hippocampal-Logistic Model



Across 100 runs, the CNN model was concordant with clinician lateralization on average 78% (SD = 5.1%) of runs with the best-performing model achieving 89% concordance.
The CNN outperformed the hippocampal volume model (average concordance of 71.7%) on 85% of runs with an average improvement of 6.25%.

Neurology 2023;101:e324-e335.

Feature Visualization of Differences Between L-TLE and R-TLE Patient Groups



Feature visualization maps revealed that in addition to the medial temporal lobe, regions in the lateral temporal lobe, cingulate, and precentral gyrus aided in classification.

Neurology 2023;101:e324-e335.

RESEARCH ARTICLE

Convolutional Neural Network Algorithm to Determine Lateralization of Seizure Onset in Patients With Epilepsy

A Proof-of-Principle Study

Erik Kaestner, PhD,* Jun Rao, MS,* Allen J. Chang, MS, Zhong Irene Wang, PhD, Robyn M. Busch, PhD, Simon S. Keller, PhD, Theodor Rüber, MD, Daniel L. Drane, PhD, Travis Stoub, PhD, Ezequiel Gleichgerrcht, Leonardo Bonilha, MD, PhD, Kyle Hasenstab, PhD,† and Carrie McDonald, PhD†

Correspondence
Dr. McDonald
camcdonald@
health.ucsd.edu

Neurology® 2023;101:e324-e335. doi:10.1212/WNL.0000000000207411

Classification of Evidence

This study provides Class II evidence that in patients with drug-resistant unilateral temporal lobe epilepsy, a convolutional neural network algorithm derived from T1-weighted MRI can correctly classify seizure laterality.

Neurology 2023;101:e324-e335.

Received: 9 October 2022 | Revised: 27 January 2023 | Accepted: 30 January 2023

DOI: 10.1111/epi.17522

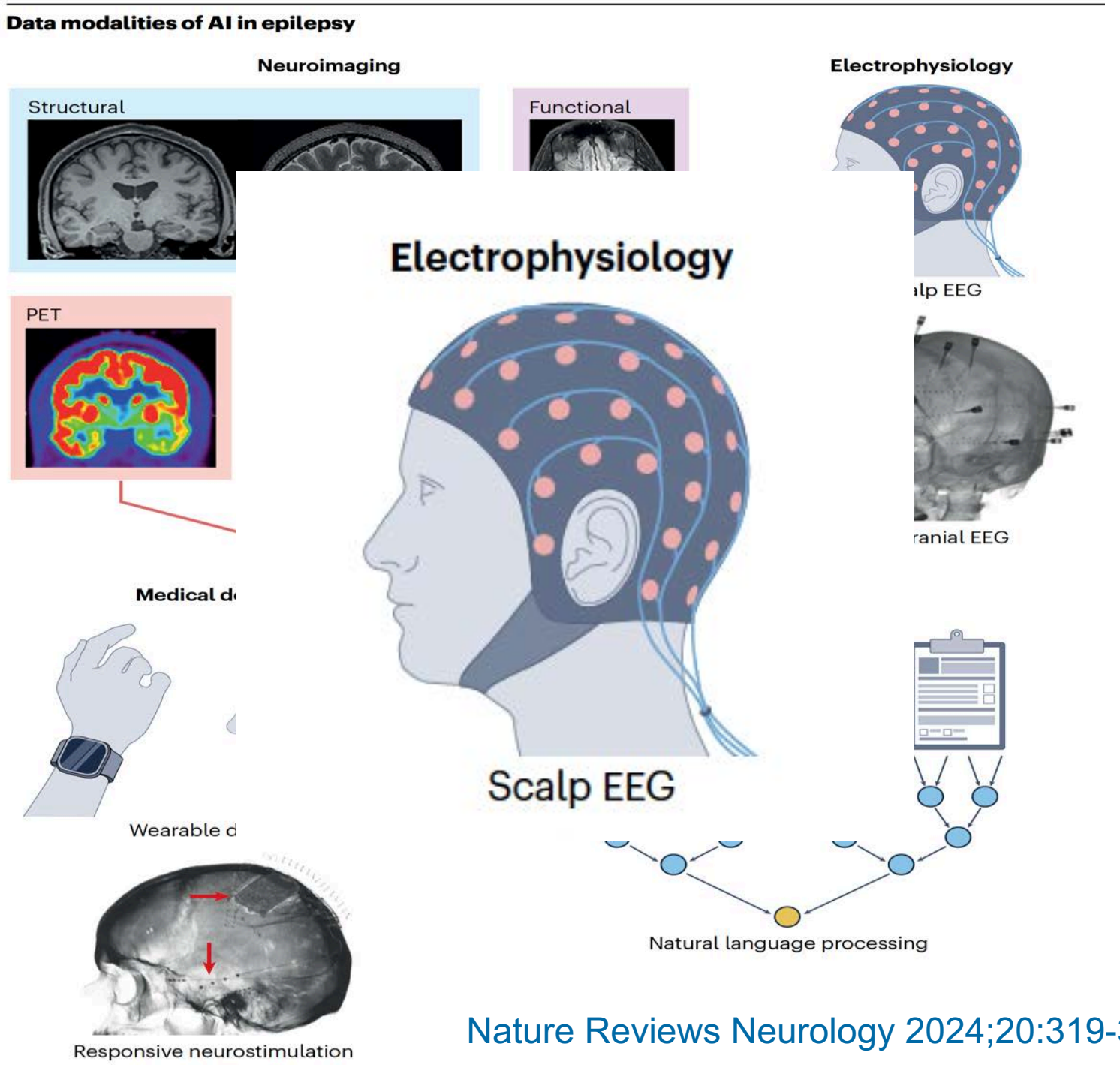
Epilepsia®

CRITICAL REVIEW

Artificial intelligence for the detection of focal cortical dysplasia: Challenges in translating algorithms into clinical practice

Lennart Walger[1] | Sophie Adler[2] | Konrad Wagstyl[3] | Leonie Henschel[4] | Bastian David[1] | Valeri Borger[5] | Elke Hattingen[6] | Hartmut Vatter[5] | Christian E. Elger[1] | Torsten Baldeweg[2] | Felix Rosenow[7,8] | Horst Urbach[9] | Albert Becker[10] | Alexander Radbruch[11] | Rainer Surges[1] | Martin Reuter[4,12,13] | Fernando Cendes[14] | Zhong Irene Wang[15] | Hans-Jürgen Huppertz[16] | Theodor Rüber[1]

Epilepsia. 2023;64:1093–1112.



Nature Reviews Neurology 2024;20:319-36

Research

JAMA Neurology | Original Investigation

Automated Interpretation of Clinical Electroencephalograms Using Artificial Intelligence

Jesper Tveit, PhD; Harald Aurlien, MD, PhD; Sergey Plis, PhD; Vince D. Calhoun, PhD; William O. Tatum, DO; Donald L. Schomer, MD; Vibeke Arntsen, MD; Fieke Cox, MD, PhD; Firas Fahoum, MD; William B. Gallentine, DO; Elena Gardella, MD, PhD; Cecil D. Hahn, MD; Aatif M. Husain, MD; Sudha Kessler, MD; Mustafa Aykut Kural, MD, PhD; Fábio A. Nascimento, MD; Hatice Tankisi, MD, PhD; Line B. Ulvin, MD; Richard Wennberg, MD, PhD; Sándor Beniczky, MD, PhD

To develop and validate an AI model (Standardized Computer-based Organized Reporting of EEG–Artificial Intelligence [SCORE-AI]) with the ability to distinguish abnormal from normal EEG recordings
to classify abnormal EEG recordings into categories relevant for clinical decision-making: epileptiform-focal, epileptiform-generalized, nonepileptiform-focal, and nonepileptiform-diffuse

JAMA Neurol. 2023;80(8):805-812

INPUT DATA → FEATURE SELECTION → Analyze using **MAPPING FUNCTION** → **OUTPUT PREDICTION**

Data pipeline

30,493 recordings of patients referred for EEG were included into the development data set annotated by 17 experts

→ Feature selection

3 independent test data sets:

- a multicenter data set of 100 EEGs
- evaluated by 11 experts
- a single-center data set of 9785 EEGs evaluated by 14 experts
- a data set of 60 EEGs with external reference standard (for benchmarking with previously published AI models)

→ Convolutional neural network model SCORE-AI → EEG classification

JAMA Neurol. 2023;80(8):805-812

The SCORE-AI achieved high accuracy, with an area under the receiver operating characteristic curve between 0.89 and 0.96 for the different categories of EEG abnormalities, and performance similar to human experts.
Benchmarking against 3 previously published AI models was limited to comparing detection of epileptiform abnormalities. The accuracy of SCORE-AI (88.3%; 95%CI, 79.2%-94.9%) was significantly higher than the 3 previously published models ($P < .001$) and similar to human experts.

Table 1. Gwet AC1 Agreement Coefficients for the 11 Human Experts, SCORE-AI, and the Human Expert Majority Consensus

EEG recording category	Agreement coefficient (95% CI)	
	Agreement among the human experts	Agreement between SCORE-AI and majority consensus of human experts
Normal	0.723 (0.649-0.796)[a]	0.903 (0.820-0.987)[a]
Epileptiform-focal	0.723 (0.643-0.803)	0.757 (0.634-0.880)
Epileptiform-generalized	0.901 (0.854-0.949)	0.928 (0.865-0.991)
Nonepileptiform-diffuse	0.630 (0.539-0.721)	0.738 (0.608-0.868)
Nonepileptiform-focal	0.587 (0.499-0.674)	0.775 (0.657-0.893)
Exact match/multiple abnormalities	0.497 (0.433-0.561)[a]	0.689 (0.611-0.766)[a]

Abbreviations: EEG, electroencephalography; SCORE-AI, Standardized Computer-based Organized Reporting of EEG–Artificial Intelligence.

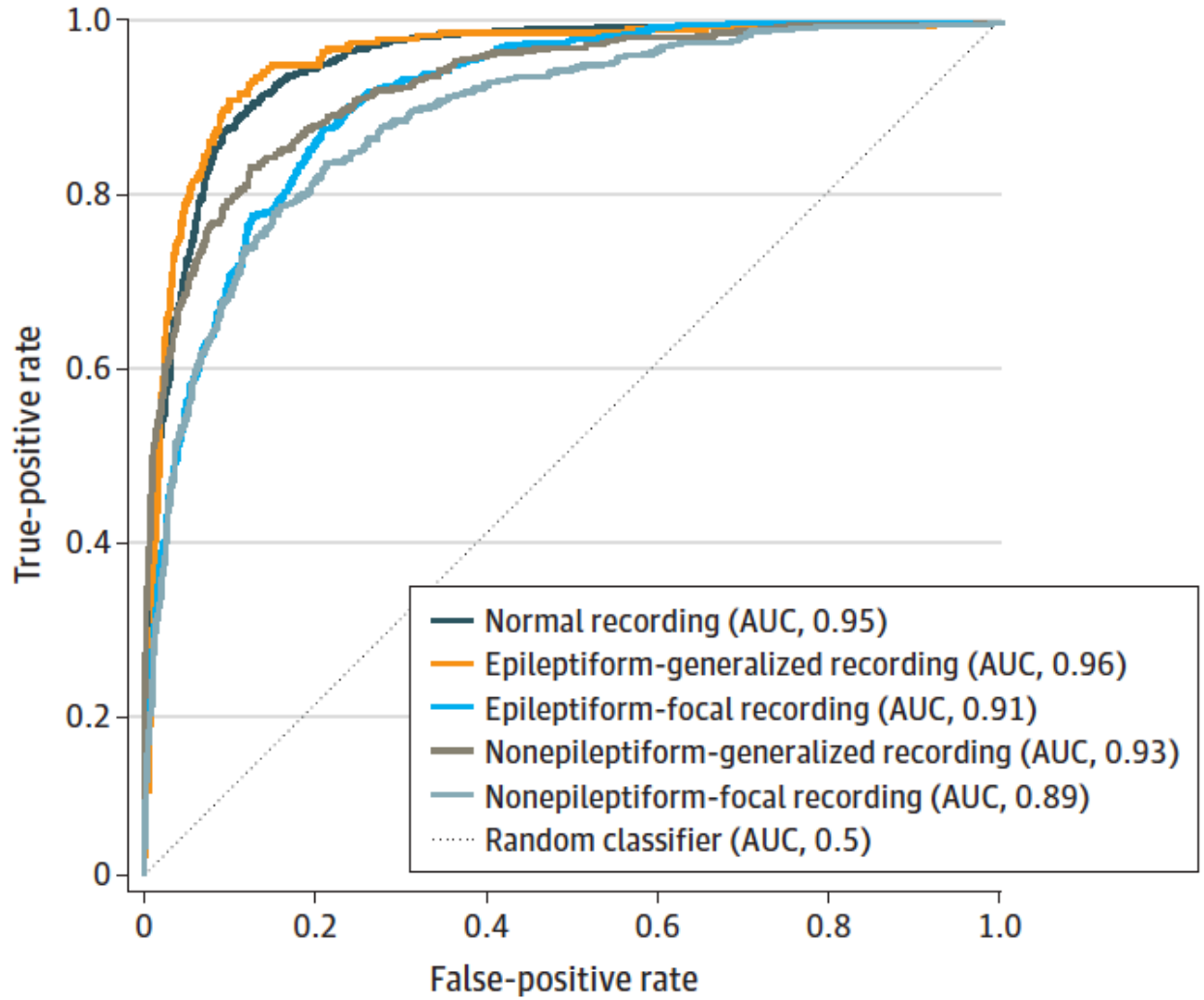
[a] Significant difference. Statistical comparisons were based on the 95% CIs. Significance means there was no overlap between the 95% CIs.

Table 2. Average Accuracy of SCORE-AI and of the Human Experts With Respect to the Human Expert Majority Consensus on 100 EEGs From the Multicenter Test Data Set

EEG recording category	Average accuracy (95% CI)		Difference (*P* value)
	SCORE-AI	Human experts	
Normal	95.00 (89.61-97.88)	91.36 (88.04-94.10)	.09
Epileptiform-focal	84.69 (76.73-90.54)	88.4 (84.35-91.91)	.12
Epileptiform-generalized	94.9 (89.41-97.83)	95.36 (92.51-97.48)	.34
Nonepileptiform-diffuse	84.69 (76.63-90.83)	86.09 (81.99-89.66)	.33
Nonepileptiform-focal	85.71 (77.86-91.41)	85.25 (81.04-88.78)	.47
Exact match/multiple abnormalities	65.31 (54.93-73.60)	66.7 (60.56-72.41)	.33

Abbreviations: EEG, electroencephalography; SCORE-AI, Standardized Computer-based Organized Reporting of EEG–Artificial Intelligence.

Figure. Receiver Operating Characteristics Curves on the Holdout Test EEG Data Set (n = 2549)



AUC indicates area under the curve.

RESEARCH ARTICLE

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

Neurology 2023;100:e1750-e1762

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

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†

Data pipeline

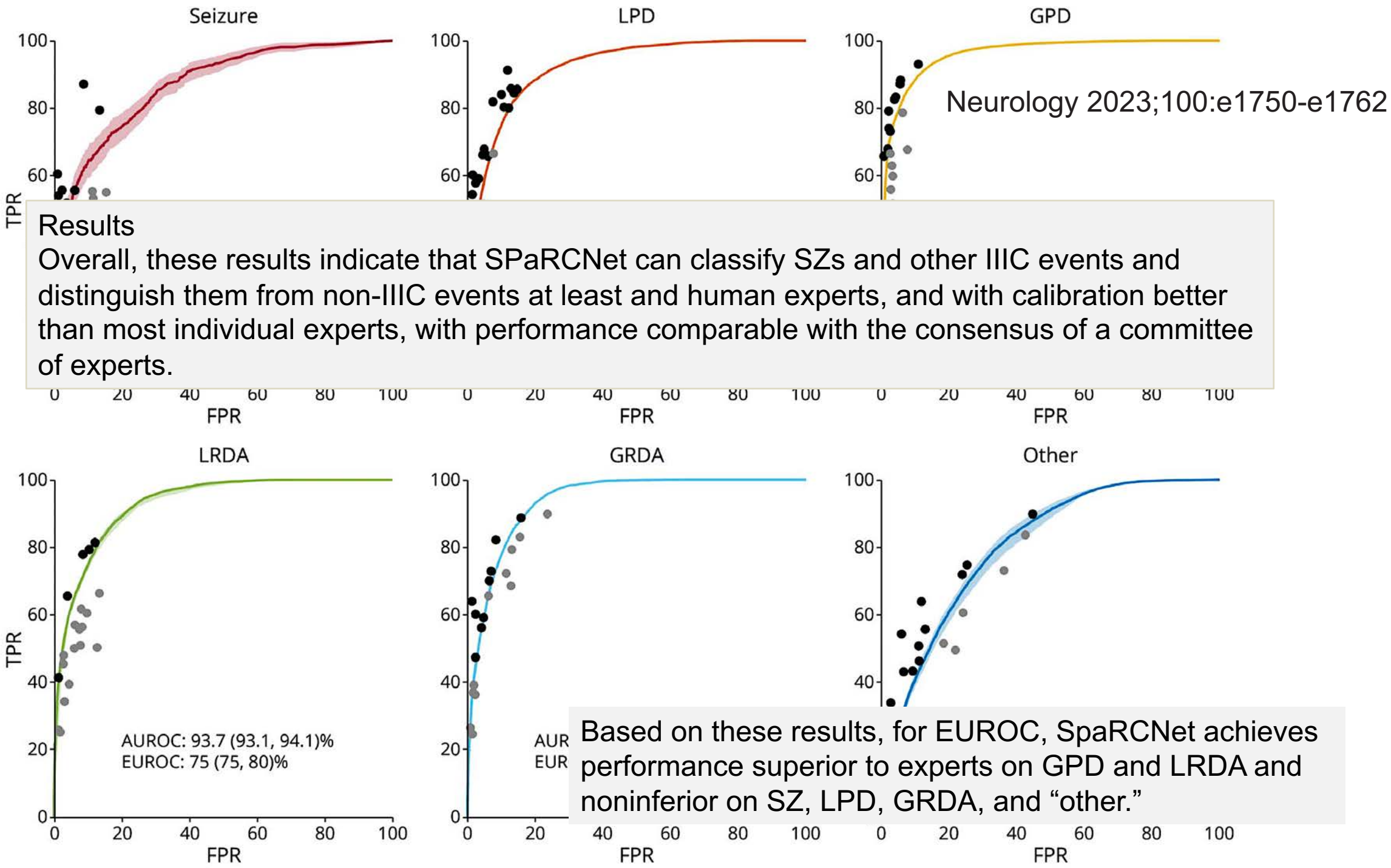
6,095 scalp EEGs from 2,711 patients with and without IIIC events

Feature selection
Independent training and test data sets were generated from 50,697 EEG segments, independently annotated by 20 fellowship-trained neurophysiologists

Deep neural network, SPaRCNet

IIIC event classification

Neurology 2023;100:e1750-e1762

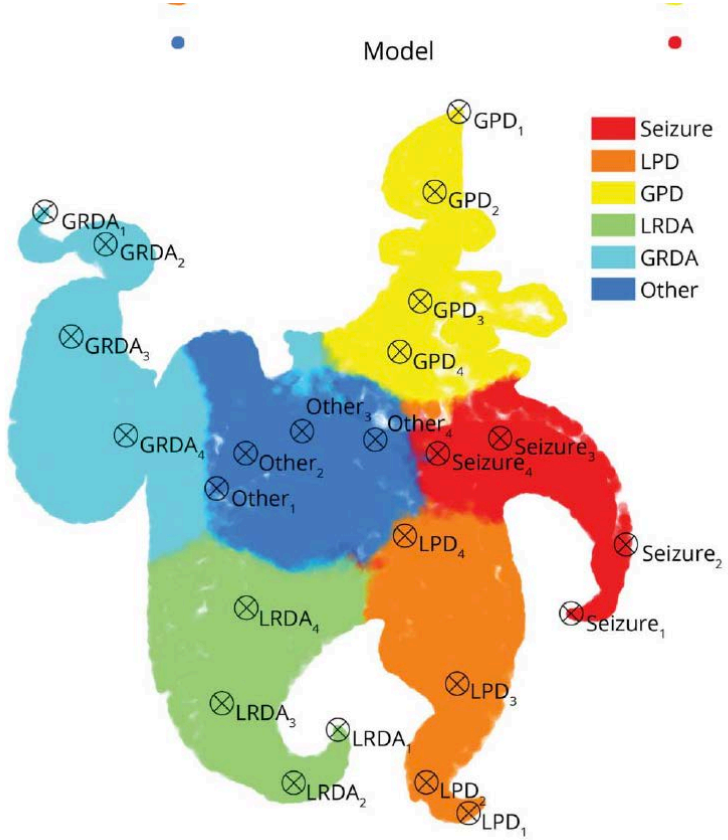


Neurology 2023;100:e1750-e1762

Results

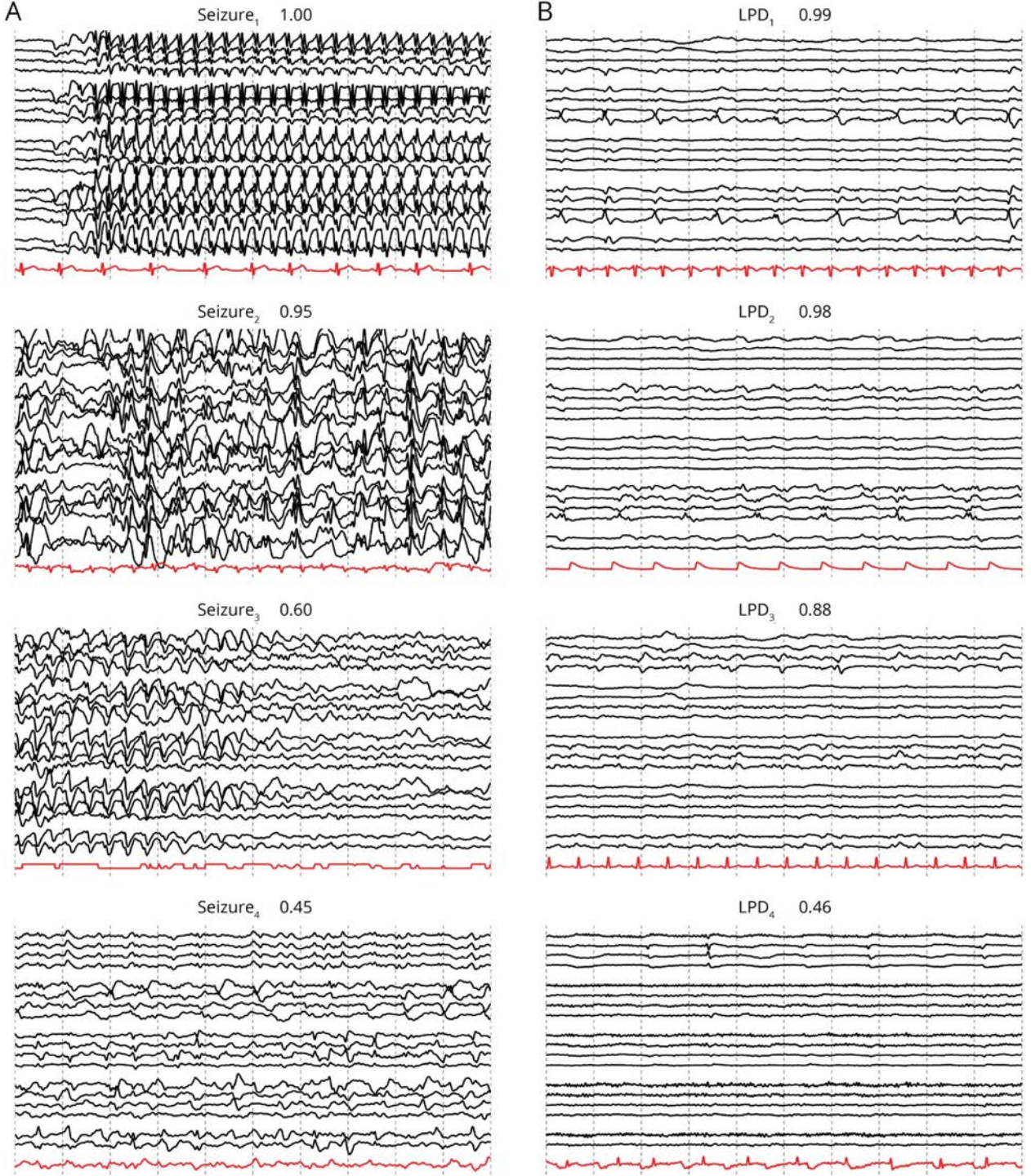
Overall, these results indicate that SPaRCNet can classify SZs and other IIIC events and distinguish them from non-IIIC events at least and human experts, and with calibration better than most individual experts, with performance comparable with the consensus of a committee of experts.

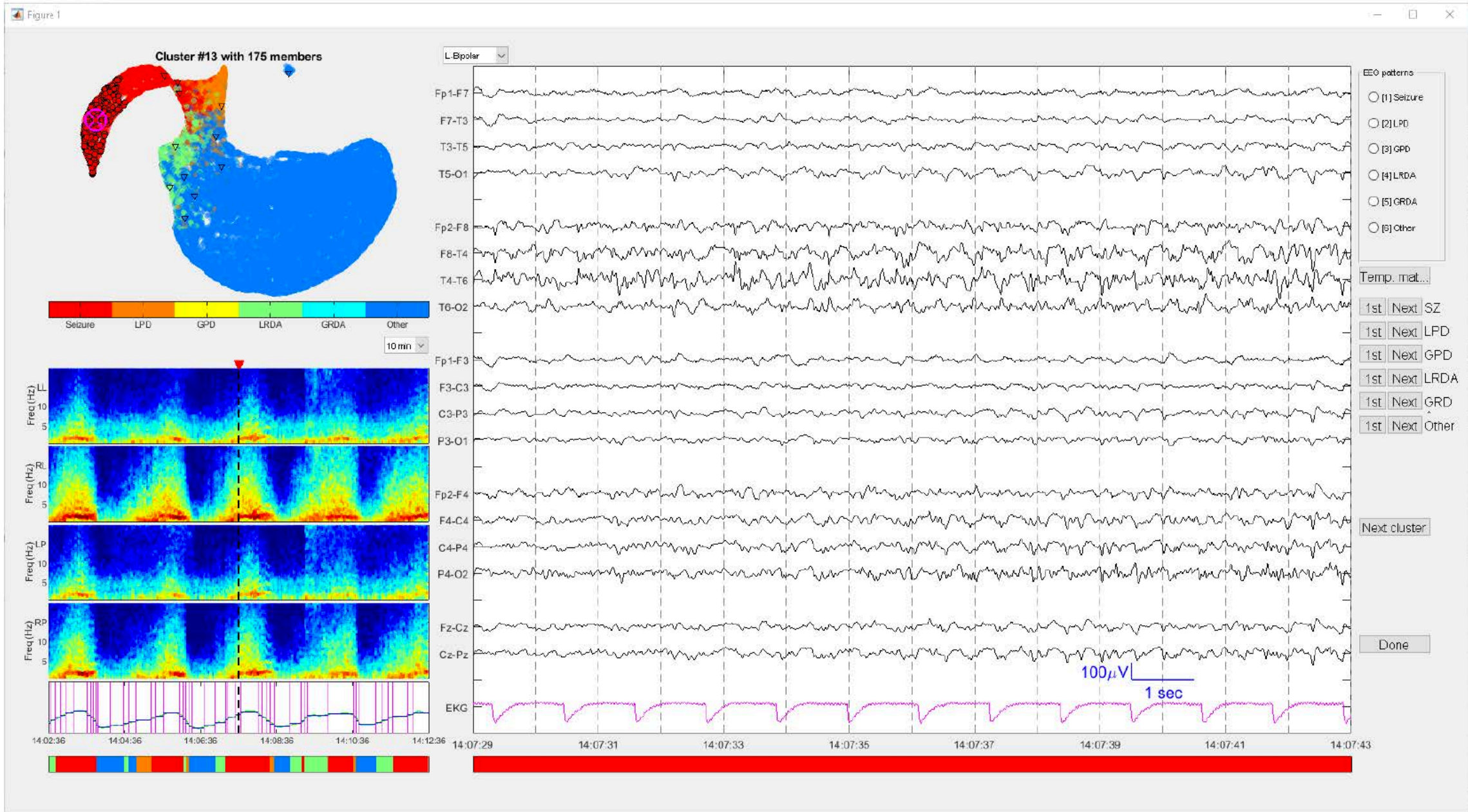
Based on these results, for EUROC, SpaRCNet achieves performance superior to experts on GPD and LRDA and noninferior on SZ, LPD, GRDA, and “other.”



Neurology 2023;100:e1750-e1762

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.





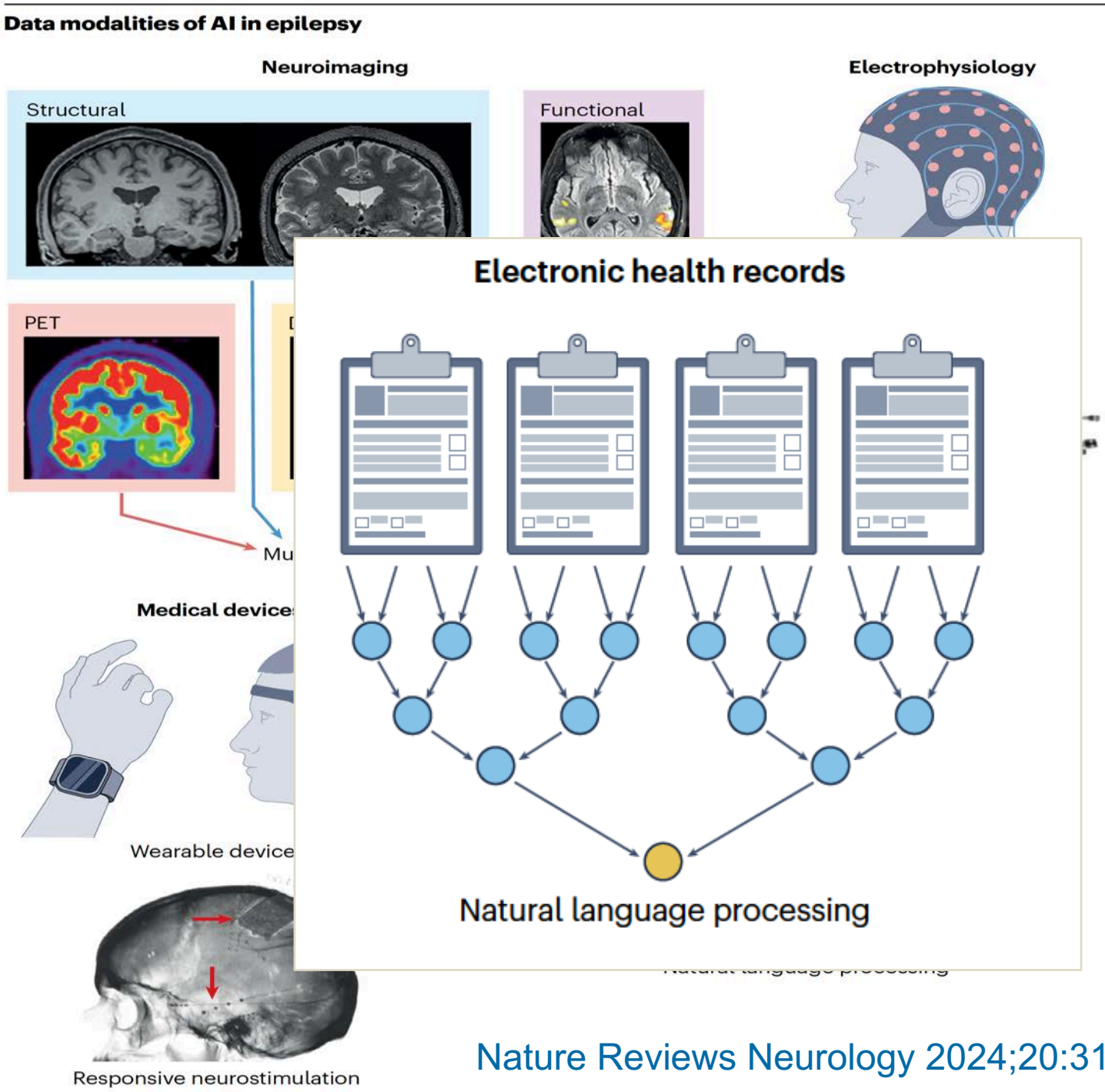
eFigure 11. The user graphical interface of "hybrid" method for expert to review model annotations.

eTable 3. Experiment to show how *SPaRCNet* can assist rapid EEG interpretation.

EEG	Seizure %	LPD %	Burden GPD %	LRDA %	GRDA %	EEG duration (hour)	Total time cost (minute)
case01	2.74	0	89.79	0	1.01	12.81	3.13
case02	3.16	83.17	0	10.28	0	12.90	2.28
case03	0	0	94.57	0	0	12.82	1.61
case04	0	76.03	21.80	0	0	12.47	2.00
case05	0.29	0	0	90.58	0	13.27	2.21
case06	1.33	53.26	11.04	0.11	0.25	12.75	2.14
case07	0	0	0	69.18	3.92	12.67	2.00
case08	16.11	0	0	24.53	2.46	12.96	2.92
case09	0	0	0	27.50	47.64	13.48	1.82
case10	0	11.28	0	13.57	17.07	12.60	2.01
case11	6.38	23.70	0	20.07	0	12.00	2.35
case12	14.46	2.73	0	3.63	0	12.00	2.18
case13	34.75	0	11.48	0	19.22	12.00	1.72
case14	1.73	95.30	0	2.97	0	12.00	1.59
case15	1.75	40.56	3.20	9.14	8.30	12.00	1.90

Important limitations

- SPaRCNet does not identify all EEG patterns of clinical relevance. Examples of other key patterns include burst suppression, nonrhythmic slowing, and nonperiodic epileptiform discharges
- 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
- 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 artifact



Nature Reviews Neurology 2024;20:319-36

Research

JAMA Neurology | Original Investigation

Development and Validation of a Deep Learning Model for Predicting Treatment Response in Patients With Newly Diagnosed Epilepsy

Haris Hakeem, MD; Wei Feng, MS; Zhibin Chen, PhD, CStat; Jiun Choong, BEng; Martin J. Brodie, MD, PhD; Si-Lei Fong, MBBS; Kheng-Seang Lim, MBBS, PhD; Junhong Wu, MD; Xuefeng Wang, MD; Nicholas Lawn, MBChB; Guanzhong Ni, MD; Xiang Gao, MSc; Mijuan Luo, MD; Ziyi Chen, MD; Zongyuan Ge, PhD; Patrick Kwan, MD, PhD

OBJECTIVE To develop and validate a deep learning model using readily available clinical information to predict treatment success with the first ASM for individual patients.

JAMA Neurol. 2022;79(10):986-996.

INPUT DATA → FEATURE SELECTION → Analyze using **MAPPING FUNCTION** → **OUTPUT PREDICTION**

Data pipeline

A total of 2404 adults with epilepsy newly treated at specialist clinics in Scotland, Malaysia, Australia, and China between 1982 and 2020 were considered, of whom 606 (25.2%) were excluded due to missing information

→ Feature selection
16 clinical factors and ASM information

→ Attention-based deep learning model "the transformer model" to predict the probability of treatment success with the first prescribed ASM

→ Prediction

JAMA Neurol. 2022;79(10):986-996.

Table 1. Input Variables for the Machine Learning Models

Input variable	Categorization
Sex	Male or female
Age at treatment initiation	Age groups (tertiles), y[a]
History	
Febrile convulsions	Yes or no
Central nervous system infection in childhood	Yes or no
Significant head trauma	Yes or no
Cerebral hypoxic injury	Yes or no
Substance abuse	Yes or no
Alcohol abuse	Yes or no
Epilepsy in first-degree relatives	Yes or no
Presence of	
Cerebrovascular disease	Yes or no
Intellectual disability	Yes or no
Psychiatric disorder	Yes or no
No. of pretreatment seizures	≤5 or >5
Type of epilepsy	Focal, generalized, or unclassified
Electroencephalography findings	Normal, abnormal epileptiform, or abnormal nonepileptiform
Brain imaging findings[b]	Normal, abnormal epileptogenic, or abnormal nonepileptogenic
Drug used	Carbamazepine, lamotrigine, levetiracetam, oxcarbazepine, phenytoin, topiramate, or valproate

[a] Tertiles are 18 to 29 years, older than 29 to 46 years, and older than 46 years.

[b] Computed tomography or magnetic resonance imaging.

JAMA Neurol. 2022;79(10):986-996.

Table 3. Comparison of Model Performance on the Test Set of the Pooled Cohort

Model parameter	Transformer	Multilayered perceptron	Logistic regression	Support vector machine	XGBoost	Random forest
Mean AUROC (95% CI)	0.65 (0.63-0.67)	0.63 (0.60-0.66)	0.61 (0.58-0.64)	0.61 (0.59-0.63)	0.60 (0.58-0.62)	0.58 (0.56-0.60)
Weighted balanced accuracy (95% CI)	0.62 (0.60-0.64)	0.59 (0.57-0.61)	0.60 (0.58-0.62)	0.57 (0.55-0.59)	0.59 (0.57-0.61)	0.59 (0.57-0.61)
Sensitivity (95% CI)	0.69 (0.66-0.72)	0.59 (0.55-0.63)	0.54 (0.52-0.56)	0.65 (0.62-0.68)	0.54 (0.52-0.56)	0.47 (0.44-0.50)
Specificity (95% CI)	0.55 (0.52-0.58)	0.60 (0.57-0.63)	0.63 (0.60-0.66)	0.52 (0.49-0.55)	0.61 (0.58-0.64)	0.62 (0.59-0.65)

Abbreviation: AUROC, area under the receiver operating characteristic curve.

The transformer model that was trained using the pooled cohort had an AUROC of 0.65 (95%CI, 0.63-0.67) and a weighted balanced accuracy of 0.62 (95%CI, 0.60-0.64) on the test set.

JAMA Neurol. 2022;79(10):986-996.

Table 4. Model Performance After Training Exclusively on the Glasgow Cohort (N = 1065)

Model parameter	Type of model[a]					
	Transformer	Multilayered perceptron	Logistic regression	Support vector machine	XGBoost	Random forest
Kuala Lumpur cohort (n = 242)						
Mean AUROC	0.58 (0.57-0.59)	0.55 (0.53-0.57)	0.57 (0.55-0.59)	0.57 (0.55-0.59)	0.57 (0.55-0.59)	0.46 (0.44-0.48)
Weighted balanced accuracy	0.58 (0.56-0.60)	0.52 (0.50-0.54)	0.56 (0.54-0.58)	0.54 (0.52-0.56)	0.55 (0.53-0.57)	0.49 (0.47-0.51)
Sensitivity	0.46 (0.44-0.48)	0.55 (0.51-0.59)	0.56 (0.52-0.60)	0.59 (0.55-0.63)	0.50 (0.47-0.53)	0.38 (0.35-0.41)
Specificity	0.65 (0.61-0.69)	0.50 (0.46-0.54)	0.56 (0.53-0.59)	0.50 (0.47-0.53)	0.59 (0.56-0.62)	0.55 (0.53-0.57)
Chongqing cohort (n = 191)						
Sensitivity	0.41 (0.39-0.43)	0.50 (0.47-0.53)	0.50 (0.47-0.53)	0.50 (0.47-0.53)	0.59 (0.56-0.62)	0.42 (0.39-0.45)
Guang… (n = 1…)						
Mea…						
Weighted balanced accuracy	0.51 (0.49-0.53)	0.48 (0.46-0.50)	0.52 (0.50-0.54)	0.49 (0.47-0.51)	0.49 (0.47-0.52)	0.45 (0.43-0.47)
Sensitivity	0.47 (0.44-0.50)	0.47 (0.44-0.50)	0.53 (0.50-0.56)	0.49 (0.46-0.52)	0.49 (0.46-0.52)	0.44 (0.40-0.48)
Specificity	0.55 (0.52-0.58)	0.50 (0.46-0.54)	0.50 (0.47-0.53)	0.47 (0.44-0.50)	0.51 (0.49-0.53)	0.46 (0.43-0.49)

Abbreviations: AUROC, area under the receiver operating characteristic curve; XGBoost, extreme gradient boosting.

[a] The numbers in parentheses are 95% CIs.

CONCLUSIONS AND RELEVANCE In this cohort study, a deep learning model showed the feasibility of personalized prediction of response to ASMs based on clinical information. With improvement of performance, such as by incorporating genetic and imaging data, this model may potentially assist clinicians in selecting the right drug at the first trial.

The model that was trained using the largest cohort only had AUROCs ranging from 0.52 to 0.60 and a weighted balanced accuracy ranging from 0.51 to 0.62 in the external validation cohorts.

JAMA Neurol. 2022;79(10):986-996.

Received: 13 February 2023 | Revised: 30 April 2023 | Accepted:
DOI: 10.1111/epi.17637

RESEARCH ARTICLE

Predicting seizure outco
need more complex mod

Maria H. Eriksson[1,2,3,4] | Mathilde
Friederike Moeller[6] | Krishna B.
John Booth[8] | Kirstie J. Whitaker[4]
Ana Perez Caballero[9] | Lara Menzi
J. Helen Cross[1,3,5,11] | Torsten Bald

Results: Our logistic regression achieved an accuracy of 72% (95% confidence interval [CI] = 68%–75%, area under the curve [AUC] = .72), whereas our multilayer perceptron and XGBoost both achieved accuracies of 71% (95% CI_{MLP} = 67%–74%, AUC_{MLP} = .70; 95% $CI_{XGBoost\ own}$ = 68%–75%, $AUC_{XGBoost\ own}$ = .70). There was no significant difference in performance between our three models (all $p > .4$) and they all performed better than the external XGBoost, which achieved an accuracy of 63% (95% CI = 59%–67%, AUC = .62; $p_{LR} = .005$, $p_{MLP} = .01$, $p_{XGBoost\ own} = .01$) on our data. All models showed improved performance with increasing sample size, but limited improvements beyond our current sample. The best model performance was achieved with data-driven feature selection.

Significance: We show that neither the deployment of complex machine learning models nor the assembly of thousands of patients alone is likely to generate significant improvements in our ability to predict postoperative seizure freedom. We instead propose that improved feature selection alongside collaboration, data standardization, and model sharing is required to advance the field.

Epilepsia. 2023;64:2014–2026.

Epilepsia. 2020;61:421–432.

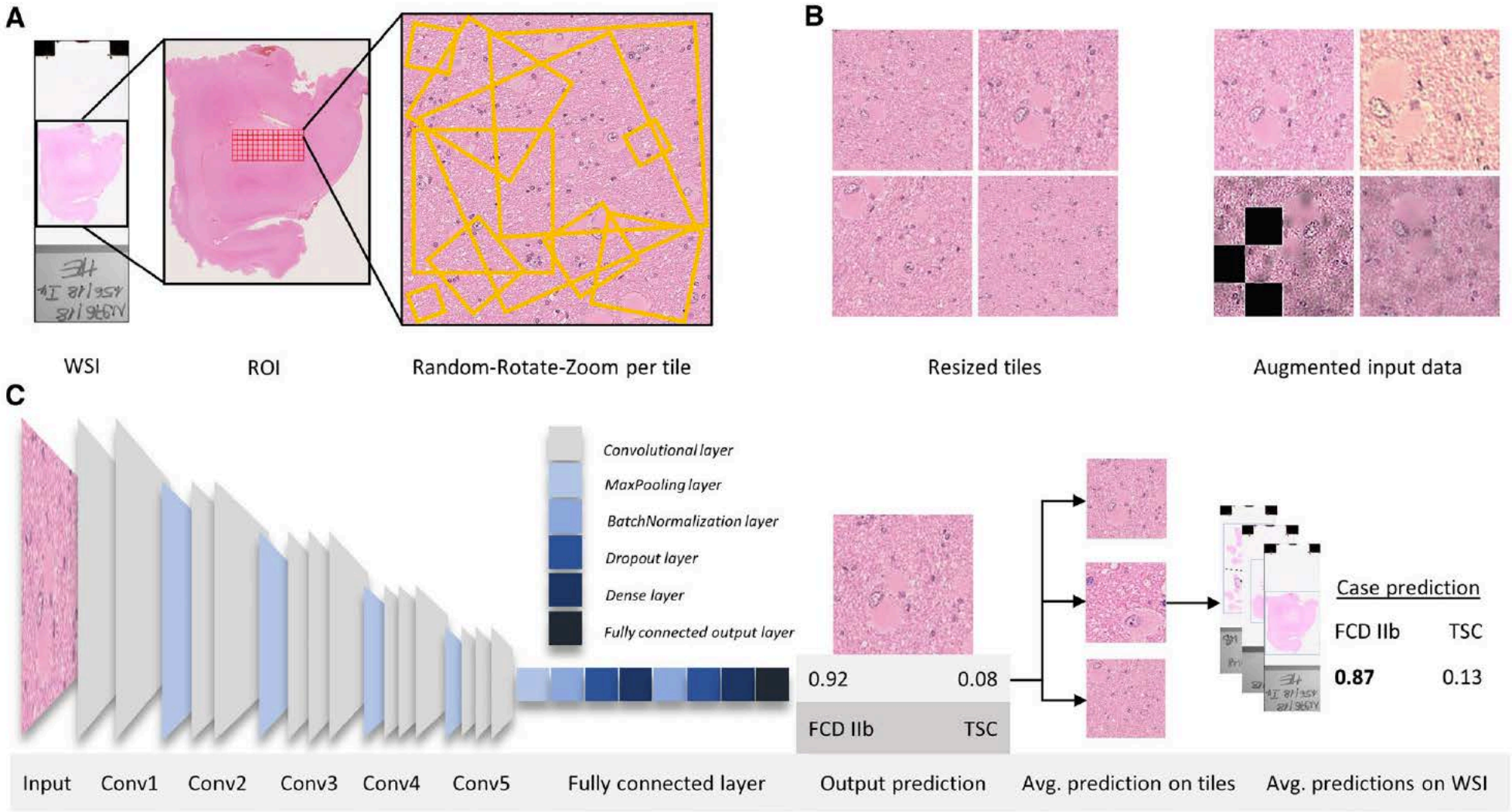
Received: 12 November 2019 | Revised: 23 January 2020 | Accepted: 23 January 2020

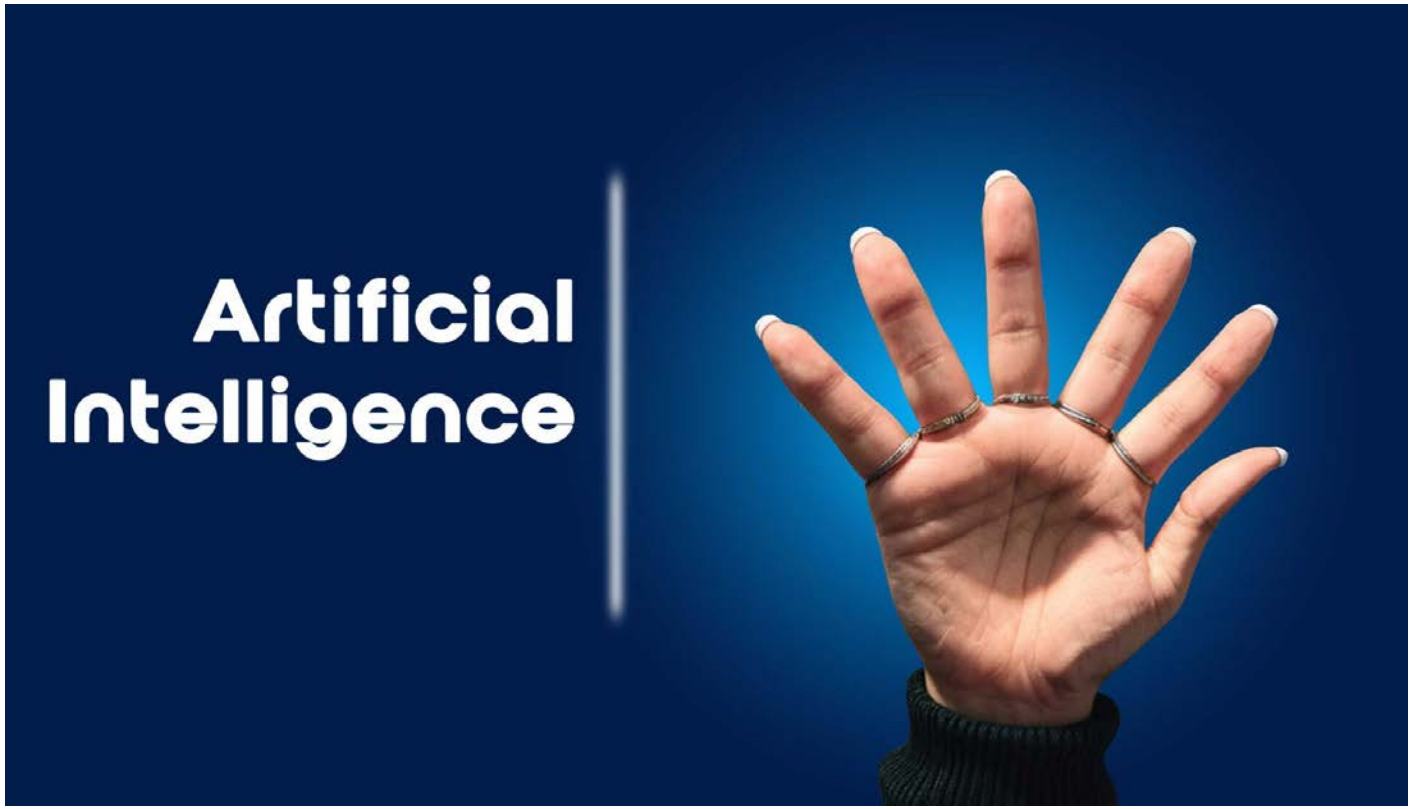
DOI: 10.1111/epi.16447

FULL-LENGTH ORIGINAL RESEARCH

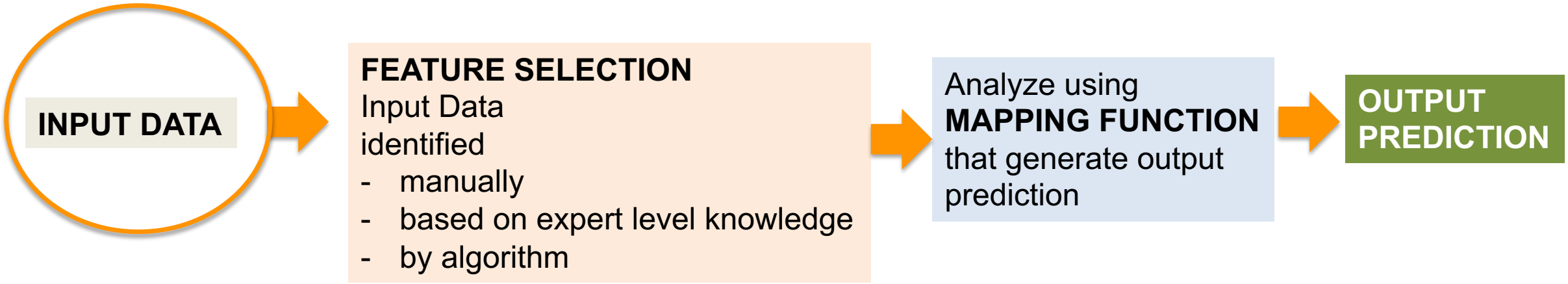
Epilepsia®

Same same b revealed class cortical malfo




Artificial
Intelligence

Data pipeline



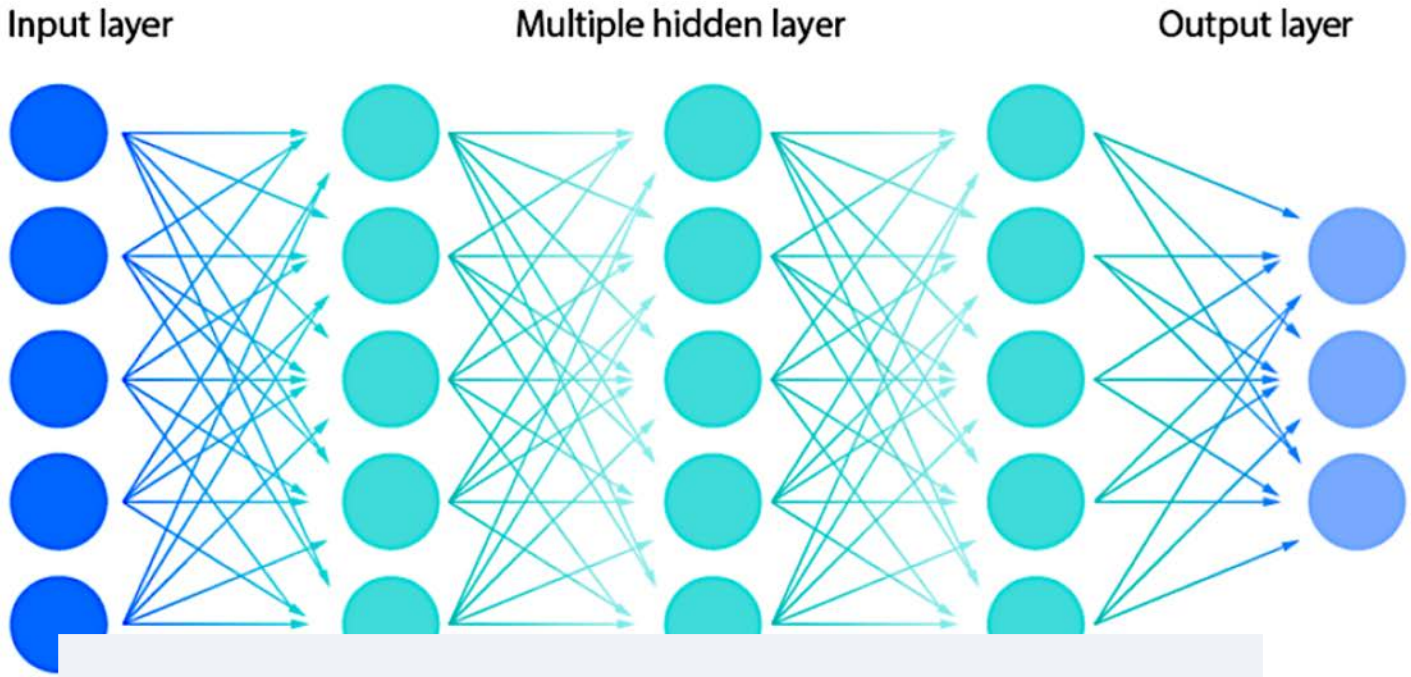
1. AI needs “gold standard labels” for evaluation

- Garbage in→ a lot of garbage out
- Example
- EEG: expert to expert agreement of seizure is low
- Electronic medical record: incomplete
- ICD codes: limited codes for epilepsy

2. Training data reflects where we can apply the particular AI program

Data pipeline

Hallucinations



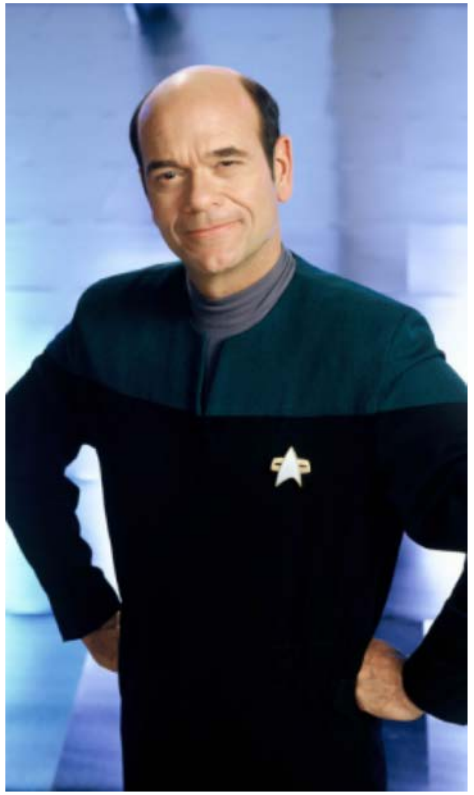
Multilayers artificial neural networks process data through layers of nodes, passed through a nonlinear activation these may in turn proceed through add tput nodes

“Black box problem”

- Advantages
  - Do more work in less time
  - Improve clinical decision in challenging situations
- Limitations
  - Need large number of “good” data
  - Machines only knows what it has seen in training
  - Require supervision
  - Hallucinations

Current Neurology and Neuroscience Reports 2023; 23:869–879

AI will not replace clinicians, but clinician assisted by AI will replace clinician without AI.



EMH: Emergency Medical Hologram (Star Trek: Voyager)

Wesly T. Kerr, MD., PhD
University of Pittsburgh

Questions?