Distributed brain co-processor for tracking spikes, seizures and behaviour during electrical brain stimulation


* These authors contributed equally to this work.
† Senior and corresponding authors.

Early implantable epilepsy therapy devices provided open-loop electrical stimulation without brain sensing, computing, or an interface for synchronised behavioural inputs from patients. Recent epilepsy stimulation devices provide brain sensing but have not yet developed analytics for accurately tracking and quantifying behaviour and seizures. Here we describe a distributed brain co-processor providing an intuitive bi-directional interface between patient, implanted neural stimulation and sensing device, and local and distributed computing resources. Automated analysis of continuous streaming electrophysiology is synchronized with patient reports using a handheld device and integrated with distributed cloud computing resources for quantifying seizures, interictal epileptiform spikes and patient symptoms during therapeutic electrical brain stimulation. The classification algorithms for interictal epileptiform spikes and seizures were developed and parameterized using long-term ambulatory data from nine humans and eight canines with epilepsy, and then implemented prospectively in out-of-sample testing in two pet canines and four humans with drug-resistant epilepsy living in their natural environments. Accurate seizure diaries are needed as the primary clinical outcome measure of epilepsy therapy and to guide brain-stimulation optimization. The brain co-processor system described here enables tracking interictal epileptiform spikes, seizures and correlation with patient behavioural reports. In the future, correlation of spikes and seizures with behaviour will allow more detailed investigation of the clinical impact of spikes and seizures on patients.
Correspondence to: Gregory A. Worrell
Department of Neurology, Bioelectronics
Neurophysiology and Engineering Laboratory
Mayo Clinic, 200 First Street SW
Rochester, MN 55905, USA
E-mail: worrell.gregory@mayo.edu

Correspondence may also be addressed to: Vaclav Kremen
E-mail: kremen.vaclav@mayo.edu

Keywords: epilepsy; seizures; electrophysiology; machine learning

Abbreviations:

- ANT: anterior nucleus thalamus
- AUROC: area under receiver operator curve
- CNN: convolutional neural network
- EBS: electrical brain stimulation
- FFT: fast Fourier transform
- HPC: hippocampus
- iEEG: intracranial electroencephalography
- IES: interictal epileptiform spikes
- LSTM: long-short-term memory network
- MH: RC + S™ human subject
- MD: RC + S™ dog subject
- NV: NeuroVista
- NH: NeuroVista human subject
- ND: NeuroVista dog subject
- PRC: precision-recall curve
- RC + S™: investigational Medtronic Summit RC + S™
- ROC: receiver operator curve
- STFT: short-time Fourier transform
- TLE: temporal lobe epilepsy

Graphical Abstract

Introduction

Epilepsy affects nearly 1.0% of the world population and is associated with a high disease burden.1,2 Approximately one-third of people with epilepsy continue to have seizures despite dietary, behavioural and antiseizure medication therapies.3 Electrical brain stimulation has emerged as a reversible and effective palliative therapy for drug-resistant epilepsy, but therapy optimization is slow and long-term seizure freedom rare.4,5 Despite the addition of brain sensing, current electrical stimulation devices lack accurate seizure diaries.6–8 Currently physician rely on patient seizure diaries that are known to be unreliable9,10 coupled with incomplete electrographic data.6,8 The challenge of patient management without accurate seizure counts has remained a persistent technology gap impeding epilepsy management.

Here we describe a distributed brain co-processor that enables wireless streaming of intracranial electroencephalography...
(iEEG), seizure and interictal epileptiform spike (IES) detection, accurate seizure diaries and synchronized patient annotations of seizure symptoms. The system creates a gold-standard seizure diary that can be used to guide electrical brain-stimulation therapy. The distributed brain co-processor provides integration of implantable brain sensing and stimulation devices with off-the-body commercial electronics (smartphone, tablet and watch) for clinical and neuroscience research applications. The integration of implantable devices with commercial electronics via bi-directional wireless connectivity allows algorithm complexity to scale with advances in consumer cloud computer and smartphone hardware. Brain implants providing sensing and bi-directional wireless connectivity enable continuous electrophysiology data streaming, and when coupled with off-the-body computing resources overcome the computational and data storage limitations of current implantable electrical brain-stimulation (EBS) devices. Until recently, there were several obstacles to consolidating the technology layers required for EBS, streaming continuous brain electrophysiology and synchronized behaviour reports. Here, we utilize the investigational Medtronic Summit RC + S™ (RC + S™), a rechargeable sensing and stimulation implantable device with a bi-directional application programming interface, to demonstrate these capabilities in canines and humans living with epilepsy. The system enables continuous streaming of iEEG to a handheld tablet or smartphone for real-time analysis and tracking of IESs, seizures and correlation with synchronized patient reports (Fig. 1). The electrophysiology classifiers (seizure and IES) were validated, tested and then prospectively deployed for out-of-sample testing in pet canines and humans living in their natural environments with epilepsy.

**Materials and methods**

**Study design and data sources**

To develop classification algorithms, we used a large database of iEEG from two different implanted investigational devices that wirelessly stream iEEG data, the NeuroVista (NV) and Medtronic Summit RC + S™ devices. The development data set included 13 humans and 8 canines (Fig. 2). We used two humans and eight dogs for training and validation of the seizure detection algorithm. The seizure detection algorithm was pseudo-prospectively tested in archived data from seven humans implanted with the NV device (NH3–9). True prospective testing was completed using the RC + S™ device in four humans (MH1–4) and two pet dogs (MD2 and MD3) living in their natural environments.

**Devices, training, validation and testing data**

Data sets collected from two implantable devices were utilized for system training, validation and testing (Fig. 2). The investigational NV system is a 16-channel brain sensing (0.1–100 Hz bandwidth; 400 Hz sampling) implantable device providing continuous iEEG wireless streaming to an off-the-body data storage and analytics device carried by the patients and dogs. The RC + S™ is a 16-channel electrical stimulation and sensing implantable device capable of selective sensing from any 4 of the 16 channels (1–70, 125, 250 Hz bandwidth; programmable sampling 250, 500 or 1000 Hz) and wireless streaming to a handheld tablet computer with cellular and internet connectivity to a central cloud based data and analytics platform. The investigational NV and RC + S™ devices have yielded massive data sets of ambulatory iEEG from dogs and humans with epilepsy in naturalistic settings and are ideal for development of robust automated algorithms for brain behavioural state classification, IES and seizure detection. We have previously used the NV device data from humans and canines for developing seizure detection and forecasting algorithms.

**Canine device implants**

The animal research took place at Mayo Clinic, Rochester, MN and University of California Davis, Davis, CA under IACUC Protocol A00002635 Chronic Wireless Electrophysiology and Modulation in Epileptic Dogs. Epilepsy occurs naturally in dogs with prevalence, age of onset and clinical presentation similar to human epilepsy. Naturally occurring canine epilepsy is often drug resistant and new therapies are needed. In addition, the canines provide a platform for preclinical testing, since dogs are large enough to accommodate devices designed for humans. All canines were implanted with either the NV or RC + S™ devices at either Mayo Clinic or University of California, Davis. The pet dogs with epilepsy were implanted and clinically managed at University of California, Davis, CA.

**NeuroVista seizure advisory system**

Five dogs [NeuroVista dog subject (ND) 1–5] were implanted with the investigational NV device. All canines were implanted with four subdural, four contact strip electrodes placed through small keyhole craniotomies. The electrode tails were tunnelled to the NV device in a pocket behind the canine’s right scapula.

**Investigational Medtronic Summit RC + S™**

Three dogs (MD1–3) where implanted with the RC + S™. Deep brain-stimulation electrodes were implanted intracranially into bilateral anterior nucleus of the thalamus (ANT), hippocampus or neocortex in canines under anaesthesia using a custom-made stereotactic frame. Canines underwent a 3.0T MRI using a stereotactic T1-weighted sequence (Fig. 3). Targets and trajectories were planned using stereotactic software (Compass™ Stereotactic Systems) adapted for a large animal head frame. Burr holes were drilled into the skull for each of the four electrode leads (Medtronic lead models 3391 and 3387) that were inserted to the target depth and secured with metal anchors and bone screws. The electrode lead tails were tunnelled to the
RC+STM in a pocket behind the canine’s right scapula. The canine underwent a post-op X-ray CT scan, which was then co-registered to the stereotactic MRI (Analyze 12.0; BIR, Mayo Foundation) in order to verify targeting accuracy. We have previously described the similar procedure for the NV device implants.16,17

Figure 1 Distributed brain co-processor. Integrating implanted sensing and stimulation devices with off-the-body and cloud computing resources. The system was developed and prospectively tested in canines and humans with drug-resistant epilepsy living in their natural environments. (Top) Schematic for bi-directional data transmission between implanted brain sensing and stimulation device integrated with local handheld computer (epilepsy patient assist device) and cloud environment. Deep brain-stimulation (DBS) electrodes were implanted in anterior nucleus of thalamus (ANT) and hippocampus (HPC). The integrated system provides a platform for chronic ambulatory monitoring of patient reported behaviour, device data (battery, telemetry and electrode impedance), seizures and interictal epileptiform spikes (IESs). (Bottom) The cloud co-processor enables connection to distributed devices, review of electrophysiology data and analytics from a battery of algorithms running on the patient’s local handheld or in the cloud environment. The physician can quickly review and confirm or reject automatically detected and patient reported candidate seizure events. The panel shows 7 days of continuous hippocampal IES rates and seizure detection probability. Triangles show patient reported seizure events. Circles denote automated seizure detections either confirmed as seizures (blue dots) or false positive (red) by expert visual review. Monitor inset shows example of raw data from hippocampus with automated IES detections (red dots). The patient was aware and reported (triangle) one out of the six seizures detected in the continuous intracranial EEG (iEEG) and confirmed by the physician.
Human subjects

NeuroVista seizure advisory system

The data from nine (seven males) human subjects implanted with the investigational NV device (NH 1–9) were from the human NV device trial carried out in Melbourne, Australia, between 24 March 2010 and 21 June 2011. All subjects were implanted with four subdural, four contact strip electrodes.

Investigational Medtronic Summit RC + S™

The four human subjects with RC + S™ (MH1–4) were implanted at Mayo Clinic under an FDA IDE: G180224 and Mayo Clinic IRB: 18-005483 ‘Human Safety and
Feasibility Study of Neurophysiologically Based Brain State Tracking and Modulation in Focal Epilepsy. The study is registered at https://clinicaltrials.gov/ct2/show/NCT03946618. The patients provided written consent in accordance with the IRB and FDA requirements.

We consented six patients and implanted four female patients with drug-resistant temporal lobe epilepsy (TLE) as part of the NIH Brain Initiative UH3-NS95495 Neurophysiologically-Based Brain State Tracking & Modulation in Focal Epilepsy. The details of the approach for implantation have been previously described.22 MRI was performed after Leksell (Elekta Inc.) frame fixation for stereotactic targeting. Medtronic 3387s electrodes were then implanted in the ANT by direct targeting of the mammillothalamic tract on MRI (FGATIR sequence). 23 Medtronic 3391 electrodes were implanted into the amygdala and hippocampus through direct targeting (Fig. 4). After confirmation of the electrode location with intraoperative CT, the leads were connected to bifurcated extensions and tunelled to the RC + S™™ in an infraclavicular pocket. The FDA IDE protocol investigates EBS paradigms, including low frequency (2 and 7 Hz) and high frequency (100 and 145 Hz) stimulation, IES and seizure detection, forecasting, behavioural state tracking and adaptive EBS control.

**Patient MH1**
A 57-year-old ambidextrous woman with drug-resistant mesial temporal lobe epilepsy (mTLE). History of head trauma with loss of consciousness followed by generalized tonic-clonic seizure beginning at age 9. She did well until age 21 years, when her seizures became drug resistant. She has comorbid depression and anxiety.

**Patient MH2**
A 20-year-old right-handed woman with diabetes mellitus type 1 and drug-resistant mTLE. No epilepsy risk factors. Epilepsy onset at age 7 years and a prior left temporal lobectomy at age 9 years. She was seizure free until age 17 years when seizures recurred while off all medications. Thereafter, she has been drug resistant. She has comorbid depression and anxiety.

**Patient MH3**
A 35-year-old right-handed woman history of diabetes mellitus and drug resistant mTLE. She has no epilepsy risk factors. Epilepsy onset at age 4 years old. Significant comorbid depression. She had elevated glutamic acid decarboxylase 65-kilodalton isoform antibody (GAD65) that did not respond to trials of immunotherapy.

**Patient MH4**
41-year-old right-handed woman with drug resistant mTLE. No clear risk factors for epilepsy. Epilepsy diagnosis was at age 31 years. Despite vagal nerve stimulator (VNS) and medications she had continued seizures. She has comorbid depression and anxiety.

**Detection of interictal epileptiform spikes**
The IES is an electrographic marker of pathological brain tissue capable of generating unprovoked seizures. In recent years, there has been rapid development of reliable techniques for automated IES detection. To train and evaluate the IES detector, we used continuous hippocampal recordings from the RC + S™™.12 We used a previously validated algorithm that models and adapts based on statistical distributions of signal envelopes from background (normal) iEEG activity. This enables differentiating signals containing IESs from signals with background activity even in long-term data recordings with changing background electrophysiological activity. The IES detector also identified low-amplitude IES in cases where the background activity power is low and IES are often missed by expert visual review.

We benchmarked the IES detector using data acquired with a chronically implanted brain stimulator. We deployed the detector in a cloud system that received the continuously streaming hippocampal data over 1 year. We compared the detector performance with the manual visual review (G.A.W. and N.M.G. electroencephalographers) scoring selected epochs (see Data for IES Detector). The IES detector ran during different ANT stimulation paradigms (no stimulation, 2, 7 and 145 Hz stimulation) with changing stimulation current amplitudes (2, 3 and 5 mA) and pulse widths of 90 and 200 µs.

To investigate how IESs characteristics change in periods of different seizure frequency, we selected epochs of the data in periods of frequent (cluster) and less frequent seizure activity (non-cluster). The seizure cluster period was defined as more than two seizures in a day. For each of the two...
(cluster, non-cluster), we selected 5-min-long epochs for left and right-hippocampal channels. Each selected epoch was taken at distinct times to assess differences between sleep and wake cycles. In total, we selected 24 5-min-long epochs reviewed independently by 2 electroencephalographers. All IESs were marked in both hippocampal channels and used subsequently to calculate congruence score between experts and to validate the automated IES detector. Subsequently, we used the 2-month period of continuously streaming human data from the implanted RC S™ to analyse IES rates and IES characteristics.

**Statistical analysis**

We statistically analysed automatically detected IES. We grouped peak-to-peak (P2P) amplitudes of IES by location (left or right) in each patient (MH1–4). The number of samples per group varies with all groups containing more than 5000 samples. Due to the non-normal data distribution the two tailed Mann–Whitney U test was used to determine statistical significance between P2P amplitudes during day/night periods of time in each patient. Since the number of samples in each group varies, we bootstrapped the distributions, and the test was repeated with random sampling of the data. Average P-values were calculated from the repeated tests.

**Generic seizure detector**

The training data set consists of long-term NV recordings from five canines (ND1–5) and two human patients (NH1–2; Fig. 2B). In canines, all seizures were included in training (340 in total). Another 628 interictal segments with various electrophysiological activity patterns were manually selected. The human data set consists of 1049 seizures and 846 interictal non-seizure segments. Half of the seizures (524) and half of the interictal segments (423) were bootstrapped and used as training data and the other half of data used in the validation data set. The validation data set included two sets of data. The first data set includes RC S™ recordings from three canines. Each recording spans at least 210 days. In total, 133 electrographic seizures and 833 interictal segments were selected from the continuous

<table>
<thead>
<tr>
<th>Pseudo-prospective NV human data set</th>
<th>Number of seizures</th>
<th>Tested interval (days)</th>
<th>AUPRC</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH3</td>
<td>39</td>
<td>728</td>
<td>0.21</td>
<td>0.90</td>
</tr>
<tr>
<td>NH4</td>
<td>43</td>
<td>726</td>
<td>0.99</td>
<td>0.96</td>
</tr>
<tr>
<td>NH5</td>
<td>731</td>
<td>558</td>
<td>0.88</td>
<td>1.00</td>
</tr>
<tr>
<td>NH6</td>
<td>684</td>
<td>183</td>
<td>0.70</td>
<td>0.99</td>
</tr>
<tr>
<td>NH7</td>
<td>173</td>
<td>766</td>
<td>0.86</td>
<td>0.99</td>
</tr>
<tr>
<td>NH8</td>
<td>277</td>
<td>394</td>
<td>0.83</td>
<td>0.98</td>
</tr>
<tr>
<td>NH9</td>
<td>99</td>
<td>465</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td>Total</td>
<td>2046</td>
<td>3820</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Avg. ± Std.</td>
<td>292 ± 273</td>
<td>545 ± 198</td>
<td>0.78 ± 0.24</td>
<td>0.97 ± 0.03</td>
</tr>
</tbody>
</table>

Performance of the generic seizure detection model for human seizures deployed on out-of-sample human NV data set in pseudo-prospective testing. Pseudo-prospective data were previously collected but analysed while maintaining the temporal relationship of all seizures. Machine learning performance metrics are shown together with the number of seizures and number of recording days in the data sets.

AUPRC, area under precision-recall curve; AUROC, area under receiver operating characteristic curve.

![Figure 4](image-url) **Figure 4 Human subject MH1.** (A) Papez circuit and implanted electrodes targeting bilateral anterior nucleus thalamus (ANT), Hippocampus (HPC) and amygdala (AMG). (B) MRI—the ANT and HPC electrodes from co-registration of MRI and post-implant CT are highlighted.
recordings upon visual review by an expert reviewer. The second data set contains the other half of the data (two NV human recordings) generated by bootstrapping the training data set.

The testing data sets include previously collected NV data sets that were used for pseudo-prospective testing and RC+S™ data sets for prospective testing. The pseudo-prospective testing was done with NV human data (NH3–9) and retains the original temporal order of the data. The NV human data set spans ~10.5 years and includes 2046 seizures (Table 1). True prospective testing of seizure detection ran over 723 days and contains 204 seizures that were recorded in the four humans (MH1–4) and two pet canines (MD2 and MD3) implanted with RC+S™ devices (Table 3) and living in their natural environments.

**Detector design—convolutional neural network with long-short-term memory**

To design a generalizable seizure detection algorithm for a generic implantable system, we required the algorithm to operate independently of the recording system, spatial electrode position and species tested. We used two of the few fully implantable devices capable of continuous iEEG wireless streaming. This allows long-term, real-time monitoring since the collected data are continually transferred from the implantable device to the brain co-processor system (tablet or smartphone and cloud computational resource). Previous seizure detectors usually utilize a combination of computationally expensive features extracted from multiple channels, or features extracted from shorter segments without adaptation to a long-term baseline. Another common limitation is that the testing is done on isolated ictal and interictal segments and not on long-term continuous recordings spanning multiple months. Lastly, deployment of seizure detectors on out-of-sample unbalanced data in subjects in their natural environments is rarely provided.

In order to address these limitations, we developed a convolutional neural network (CNN) combined with long-short-term memory (LSTM) neural network utilizing short-time Fourier transform (STFT) calculated from single lead iEEG as an input. We previously used a CNN with LSTM for automated classification of iEEG transients. The STFT converts the single lead time series data into a time-frequency representation (spectrogram). The spectrogram hyperparameters were selected based on the spectral content of electrographic seizures. Invariance to sampling frequency is achieved by using a constant time window of 1 sec with 0.5 sec overlap, and subsequently selecting only frequencies lower than 100 Hz. A raw data segment is transformed into a spectrogram image with dimensions 100 × T, where T is the number fast Fourier transform (FFT) calculations, not depending on sampling frequency (frequency domain resolution is always 1 Hz per sample). Time series data of 5 min length were empirically chosen to provide long enough EEG baseline temporal context for the LSTM, so the relative power of seizure stands out of the background activity. The final classification is made for every 0.5 s of the 5-min input raw data signal using a many-to-many LSTM architecture. Raw data are z-score normalized prior to STFT calculation and each frequency band of the resulting spectrogram is z-score normalized prior to the neural network inference. Preprocessing for all data sets in training, validation and testing is the same (Fig. 2). Dropout layers in neural networks are used for regularization during training to prevent overfitting. Similarly, we drop random segments prior to the spectrogram computation. This enables the network to handle the data from the wireless system with possible short data gaps.

The CNN LSTM model consists of two convolutional blocks (convolution and ReLU) with kernels (5, 5) and (96, 3), respectively. Subsequently, time distributed feature representation is processed with two layers of bi-directional LSTM recurrent neural network. Lastly, a fully connected layer with a softmax activation function transforms the LSTM output into probability output. The proposed architecture is trained with Adam optimizer (learning rate = 10⁻³, weight regularization = 10⁻⁶) in a many-to-many training scheme, where every input FFT window has a multiclass label. We implemented four types of labels: normal activity, IES together with artefacts, dropout segments and seizures. Adding additional labels might improve learning because the model is forced to not only distinguish interictal activity from continuous seizure activity but also interictal discharges which are not considered as electrographic seizures in different behavioural states, and thus could lower the number of false positives. The temporal resolution of the detector is defined by the FFT window step (0.5 sec). In order to train the network, we use a special purpose deep-learning computer Lambda Labs Inc. (8x GTX2080TI GPU, 64 CPU cores and 512 GB RAM). The data-parallel training method runs on all GPUs and average model gradients are used to reduce training time. The model is built in the PyTorch deep-learning library for Python.

**Training and validation of seizure detection model**

The model was trained on NV data (five canines, two human patients; Fig. 2A). All training segments were 10 min long. Random 5 min intervals were sampled from the full segments during the training every time the segment was used in training. Because the human training data set had a higher number of data segments than the canine training data set, we randomly sampled the human data epochs during the training in a way that the number of data segments from both classes was balanced.

The validation data set was used to select the best CNN LSTM model which was then subsequently deployed in testing. The model hyperparameters were selected heuristically
using the established features of seizure spectrograms and were the same for all analysis and not optimized further. Performance of the model during the training was evaluated by area under the precision-recall curve (AUPRC), where all seizure targets were set to one, and all the other classes were set to zero. The AUPRC was used as a validation score because it is independent of the probability threshold of the classifier and is not dependent on the true negative samples in these highly imbalanced data sets. Validation examples were fixed 5-min intervals and were not randomly sampled. Validation scores (AUPRC) were calculated on two different data sets (three canines with RC + SPM, two human patients with NV device) independently. The two validation scores were averaged after each training epoch and the model with the best score achieved during training was deployed on the test data set in order to obtain results (Fig. 2).

**Model deployment**

We arbitrarily chose 10 continuous seconds of ictal activity as an electrographic event that we want to detect. The model iterates over the data with 5-min windows with 100 s of overlap. The model gives a probability of seizure for every 0.5 s (higher probability is used in the overlap region) in every channel. Seizures in the test data set are marked across all channels without specification; therefore, we combine probabilities from all channels in the following way. The three highest probabilities from all channels are averaged, and from this averaged probability, the final performance measures are calculated. For a given probability threshold, the system identified continuous detection whenever the probability was above a threshold (see the transition from raw iEEG to a detection in Fig. 5). Next, every detection

**Figure 5 From raw iEEG data to probability and seizure detection.** (A) One minute of iEEG data recorded with NeuroVista device, 16 neocortical electrode contacts, containing a spontaneous seizure of subject NH7. The seizure is present on a few channels with a good signal to background ratio suitable for automated detection. (B) Time-frequency analysis of iEEG signals shows the different signatures of seizure electrophysiology (shaded area) in different channels: channel 1, where seizure is notable and channel 4 where it is hard to identify. (C) Plots of classifier probabilities for each electrode (channels 1–4 in colour CH1–4) below actual raw data showing that for some electrodes, the seizure is very prominent and for some not differentiable from the background signal. (D) The classifier output probabilities for top three probabilities together with the mean (bold line) and threshold (horizontal line) showing when the detection is raised (time 0).
interval above a threshold was automatically extended if in the next 10 s from the current detection was another detection. Subsequently, the two detections were merged into one interval. Thus, for every probability threshold, we detected intervals of various lengths which the model marks as seizures. Intervals shorter than 10 s were dropped from detected events. For detected events longer than 10 s AUPRC and area under the receiver operator curve (AUROC), scores were calculated based on the region overlap with gold-standard seizures marked by an expert reviewer. The window hyperparameter (10 s) was heuristically selected based on clinical seizures domain knowledge and not optimized.

The model was deployed to continuously process incoming data from RC + S™ animal and human subjects. Due to a different electrode configuration in the RC + S™ system in comparison with the NV system, we could not use an average of the three highest probabilities. Instead, a maximal probability given by two hippocampal channels was taken as an output of the model. Subsequently, the detected intervals were calculated from the probabilities in the same manner as for the data from the NV data set. Review of the raw data by an expert created gold-standard seizure marks for evaluating classifier performance. Thus, with all detected events and true seizure marks, the AUPRC and AUROC scores were calculated.

**Data and materials availability**

All results associated with this study are present in the paper. The data are available upon reasonable request (https://www.mayo.edu/research/labs/bioelectronics-neurophysiology-engineering/overview). The analysis code is available on GitHub (https://github.com/mselair/best_toolbox).

**Results**

**Tracking behaviour and epilepsy biomarkers in humans and canines**

We used analysis of intracranially recorded electroencephalography to detect seizures and IES in ambulatory humans and canines with drug-resistant epilepsy living in their natural environments. Continuous streaming iEEG was analysed in a cloud environment and on a tablet computer carried by subjects, which also enabled synchronized patient inputs. Physicians and engineers remain in the loop using a web-based Epilepsy Dashboard to review biomarker trends (IES rates and seizures), patient annotations (seizures, auras and medication logs) and implanted device data (battery status, telemetry and EBS parameters). The system provides an integrated machine learning platform for algorithm development, data viewing, biomarker tracking and expert annotation of events, e.g. confirmation that a detected electrophysiological event or patient reported event was a true positive seizure (Fig. 1).

**Detection of interictal epileptiform spikes**

The IES is an established biomarker of epileptogenic brain, and associated with risk for spontaneous, unprovoked seizures. For long iEEG data sets, it is labour intensive and impractical to use visual analysis to calculate IES rates. Here, we trained, validated and tested an automated IES detector on long-term continuous ambulatory iEEG recordings. We implemented a previously published automated IES detection algorithm, where the data are continuously accumulated by streaming iEEG from the RC + S™ device to a cloud database. We compared the automated IES detections to expert visual scoring from two epileptologists (N.M.G. and G.A.W.). These data included periods during day, night, seizure clusters (two or more seizures in 12 hours) and non-seizure cluster periods. There was good concordance for the IES labelling by expert visual review (Cohen’s kappa score 0.87) and between the algorithm and experts (F1 score 0.82 ± 0.08 with sensitivity 91 ± 0.6% and positive predictive value 77 ± 1.6%).

The algorithm performs well during night, day, high and low seizure periods (Table 2). The IES rates are higher during seizure clusters periods (two or more seizures in 12 hour period), but performance of the automated detector is similar during periods with high and low IES rates (F1 score was 0.84 in seizure cluster and 0.80 in non-cluster seizure periods). Despite the difference in IES rates between day (approximately 25% lower IES rates) and night the algorithm performed similarly (day F1 score was 0.81 and 0.82 at night). Visual examples of IES and comparison of automated detections with expert visual review are shown in Fig. 6 for day (A) and night (B) and illustrate the concordance between expert visual review and the automated classifier. The hippocampus IES rate variations during day and night over a 2-month period show circadian and multi-day fluctuations (Fig. 6C). We analysed IES characteristics to explore how the hippocampal IES properties differ in various behavioural states (Fig. 6D) and find higher peak-to-trough IES amplitudes during night compared with wake for all four human subjects (P < 0.01).

**Automated seizure detection**

Accurate seizure catalogues are critical for optimal epilepsy management and assessment of EBS outcomes, but remain a basic technology gap for the field. We created an accurate seizure diary based on a generic seizure detector using an LSTM artificial recurrent neural network combined with convolutional neural network (CNN) applied to continuous iEEG to reliably detect seizures in ambulatory canine and human subjects with epilepsy.

The large testing, validation and training data set from multiple brain structures in humans and canines was collected over multiple years with two different fully implantable recording devices (NV or Medtronic PLC, see Methods). The CNN LSTM model was trained on a data set from five
Table 2 Interictal epileptiform spike rates (IES)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cluster</th>
<th>Non-cluster</th>
<th>Day</th>
<th>Night</th>
</tr>
</thead>
<tbody>
<tr>
<td>IES rate per minute</td>
<td>37.4 ± 29.1</td>
<td>17.9 ± 7.2</td>
<td>13.8 ± 8.7</td>
<td>41.6 ± 23.7</td>
</tr>
<tr>
<td>F1 score</td>
<td>0.84 ± 0.09</td>
<td>0.8 ± 0.05</td>
<td>0.81 ± 0.08</td>
<td>0.82 ± 0.08</td>
</tr>
<tr>
<td>PPV</td>
<td>0.81 ± 0.2</td>
<td>0.74 ± 0.14</td>
<td>0.71 ± 0.12</td>
<td>0.8 ± 0.18</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.9 ± 0.08</td>
<td>0.9 ± 0.08</td>
<td>0.94 ± 0.001</td>
<td>0.89 ± 0.09</td>
</tr>
</tbody>
</table>

Results from prospective testing of the automated IES classifier at different time periods (day versus night) and seizure counts (seizure clusters/non-clusters) compared with expert visual review. Periods of seizure clusters were defined by two, or more, seizures in a 12-h period. The F1 score comparing the automated detector and expert visual review for labelling IES was similar for each condition studied.

Table 3 Prospective RC+S™ seizure detection results

<table>
<thead>
<tr>
<th>Prospective RC+S™ data set</th>
<th>Number of seizures</th>
<th>Tested interval (days)</th>
<th>AUPRC</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MH1</td>
<td>134</td>
<td>147</td>
<td>0.93</td>
<td>0.99</td>
</tr>
<tr>
<td>MH2</td>
<td>8</td>
<td>149</td>
<td>0.89</td>
<td>0.99</td>
</tr>
<tr>
<td>MH3</td>
<td>20</td>
<td>44</td>
<td>0.82</td>
<td>0.99</td>
</tr>
<tr>
<td>MH4</td>
<td>19</td>
<td>156</td>
<td>0.75</td>
<td>0.99</td>
</tr>
<tr>
<td>MD2</td>
<td>17</td>
<td>107</td>
<td>0.47</td>
<td>0.96</td>
</tr>
<tr>
<td>MD3</td>
<td>6</td>
<td>120</td>
<td>0.88</td>
<td>0.99</td>
</tr>
<tr>
<td>Total</td>
<td>204</td>
<td>723</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Performance of automated seizure detection in canine and human seizures deployed prospectively in pet canines and humans living with epilepsy in their home environments (MH1–4 are four human subjects and MD2 and MD3 are the two pet dogs). Machine learning performance metrics are shown together with the number of seizures and number of recording days in the data sets.

AUPRC, area under precision-recall curve; AUROC, area under receiver operating characteristic curve.

dogs with naturally occurring epilepsy (340 seizures) implanted with NV devices and one half of the data from two, randomly selected human subjects with epilepsy implanted with the NV device (524 seizures). The model was then validated on the other half of the data from two NV patients (524 seizures) and three canines (133 seizures in dogs: MD1–3) implanted with RC+S™ devices (Fig. 2).

Automated detection of spontaneous seizures recorded with iEEG is possible because of the characteristic spectral patterns that are readily identified visually and by machine learning approaches. Fig. 7 shows an example of a typical seizure with its time-frequency (spectrogram) characteristics, raw data and the CNN LSTM model seizure probability for MH1 from the out-of-sample data. The model probability for seizure classification changes in context of raw iEEG and spectral content showing high probability within the seizure activity and low probability outside the seizure (before and after the seizure). The example highlights the importance of the LSTM network in the model. While feature-based machine learning models would detect the bursts of IES at the beginning and during the seizure, the LSTM network raises the seizure probability prior and during the seizure time.

The precision-recall curves (PRCs) and receiver operator curves (ROCs) are calculated by sequentially changing the model probability detection threshold and evaluating the results for all seizures from each subject in the testing data sets (Fig. 8).

The performance of the generalized automated seizure detector using out-of-sample data from seven human patients implanted with the NV device (total seizures 2046 over 3820 days) was AUPRC 0.78 ± 0.24 and AUROC 0.97 ± 0.03 (Table 1). The performance of the model on out-of-sample data in subjects with the NV device is shown in Table 1. The performance of the generalized classifier is visualized using standard machine learning graphs of PRC and ROC for each individual human (Fig. 8). The model outperforms recently published state of the art detectors and runs in near real-time.

Prospective long-term ambulatory monitoring and algorithm testing

After training, validation and retrospective testing using previously collected data, we then deployed the automated IES and seizure classifiers prospectively in four humans (subjects MH1–4) and two pet dogs (MD2 and MD3) with epilepsy living in their home environments. In total, the system was able to record an average of 66 ± 0.17% of the data across all human subjects.

The performance of the IES detection in the ambulatory prospective data compared with gold-standard expert visual reviewed events was 0.90 sensitivity and F1 score of 0.81 (Table 2).

Prospective testing of the seizure detector in ambulatory subjects in real-world environments showed excellent performance, with an AUROC of 0.99 ± 0.01 and PRC of 0.76 ± 0.25 using the expert visual review of the continuously acquired iEEG as the gold standard for the humans (MH1–4) and the two pet canines (MD2 and MD3). The AUPRC more accurately describes the performance results of this highly imbalanced data with over 99% of the time spent in a non-seizure (interictal) state, the AUPRC was
0.93, 0.89, 0.82, 0.75 and 0.47, 0.88 for the four humans and two pet canines (MD2 and MD3), respectively (Table 3).

The human subjects reported a total of 555 seizures using the epilepsy patient assist device over the course of 945 days of monitoring, but only 39.71 ± 29.20% of the seizures reported were actually associated with an electrographic correlate (verified seizures; Table 4). Interestingly, of the 407 detected iEEG seizures, 43.86 ± 30.77% were not identified by the patient (Table 4). These results reflect the challenge of patient seizure reporting and the established unreliability of seizure diaries.

**Discussion**

There has been significant progress in EBS devices for drug-resistant epilepsy, but the time to achieve optimal individualized stimulation parameters is long and seizure-free outcomes remain relatively rare. We suspect that the ability to continuously track electrophysiology, seizure counts and patient behaviour will accelerate the optimization of individualized EBS therapy. To address the technology gaps in currently available EBS systems, we developed and deployed a distributed brain co-processor to investigate and
continuously track patient reported symptoms, IES biomarkers and seizures during EBS.

While seizure detection using iEEG is a well-established field, it has not been applied for creating accurate seizure diaries in ambulatory humans (or canines) with epilepsy. Despite advances in iEEG-sensing capability, the current devices approved for human epilepsy (NeuroPace RNS and Medtronic Percept) do not provide accurate seizure diaries. This is a notable gap given that a primary outcome measure for epilepsy is seizure counts. This highlights the significance of the current paper where for the first time in ambulatory humans and canines, accurate seizure diaries are demonstrated during concurrent anterior nucleus of the thalamus brain stimulation over multiple months in naturalistic settings. The primary device innovation that makes accurate seizure diaries possible is the ability to stream continuous
iEEG off the embedded device and onto a local computing device capable of running a sophisticated generalizable CNN LSTM detector that was demonstrated to work without subject specific tuning in two different species (dogs and humans) in out of sample, prospective ambulatory subjects.

We show that seizures and hippocampal IES rates and characteristics are dynamically changing in a circadian pattern with IES rates highest at night. The hippocampus IES rate variations showed a circadian fluctuations and higher peak-to-trough IES amplitudes during the night. The accurate automated quantification of IES is potentially of fundamental importance in epilepsy. Interestingly, seizures preferentially occurred during wakefulness in the human subjects despite increased IES rates during sleep. The reason for this phenomenon remains unclear, but future research using accurate seizure and IES rate data streams in ambulatory subjects will enable further investigation into long-term temporal dynamics of IES and seizures and enable investigations exploring the IES rates, changing IES morphology and circadian rhythms in association with seizures occurrence. The use of IES as a biomarker for seizure forecasting in the setting of EBS is an important direction for future investigation.

Regarding seizure reporting, there are two important observations. Similar to previous studies, we found that patients and pet owners often do not create reliable seizure diaries when compared with gold-standard seizure catalogues created from automated seizure detection algorithms applied to continuous iEEG. This is not surprising given that seizures can be subtle, can go unnoticed by caregivers and patients are often amnestic for their seizures. This result highlights the potential challenge of optimizing EBS and medical therapy, if arguably the most critical measure of epilepsy therapy outcome, seizure rates, is inaccurate. This may play a role in the long time required for therapy optimization with current FDA-approved devices. Furthermore, we determined that only 56.13 ± 30.77% of iEEG captured seizures are reported by patients, thus many electrographic seizures would not be available for informing EBS therapy adjustments. Whether the unreported iEEG electrographic seizures reflect amnestic episodes or are truly subclinical seizures without clinical symptoms is unclear and raises an interesting future avenue of investigation where automated seizure detections could trigger an automated patient assessment to probe mood, cognition, memory and motor impairments during and around seizures.

The current study has several limitations. The technology layers deployed in the system described here are associated with additional patient burden given that rechargeable devices must be periodically charged (Fig. 1; implantable device and tablet computer). Given the fact that seizures can be relatively rare events, the accumulation of adequate statistics remains a fundamental challenge for epilepsy research.

Here we will also discuss some of the challenges in ambulatory subjects. The iEEG data transmission from the embedded device to the handheld device is the most severe system case). In the future, performing computing on the embedded device, we can likely achieve more efficient data sampling and decrease the battery load on the implanted device. Similarly, the volume of iEEG transferred from the handheld to the cloud can be reduced given most analysis can be accomplished on a modern handheld (iPhone and Tablet) and significantly reduce the need to transfer iEEG data to the cloud.

There are also significant privacy issues with streaming brain data and patient reports, but this is largely managed by de-identifying and end-to-end encrypting all data streams. Furthermore, the EBS parameters can only be changed within an established safe parameter space set by the physician working directly with the patient. These applications likely reflect the future of implanted device management. Patients will no longer have to travel to the clinic for device management. These systems will strengthen the connection of patients and their care teams as they live in their home environments.

Table 4 Analysis of patient seizure diaries and electrographic seizures

<table>
<thead>
<tr>
<th>Subject</th>
<th>Days monitored</th>
<th>iEEG seizures</th>
<th>Reported seizures</th>
<th>Patient report with iEEG seizure (%)</th>
<th>iEEG seizures not reported by patient (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MH1</td>
<td>269</td>
<td>273</td>
<td>80</td>
<td>38.75</td>
<td>88.64</td>
</tr>
<tr>
<td>MH2*</td>
<td>258</td>
<td>13</td>
<td>279</td>
<td>2.87</td>
<td>38.46</td>
</tr>
<tr>
<td>MH3</td>
<td>153</td>
<td>89</td>
<td>165</td>
<td>43.03</td>
<td>20.22</td>
</tr>
<tr>
<td>MH4**</td>
<td>265</td>
<td>32</td>
<td>31</td>
<td>74.19</td>
<td>28.12</td>
</tr>
<tr>
<td>Total</td>
<td>945</td>
<td>407</td>
<td>555</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Avg. ± Std.</td>
<td>378 ± 287</td>
<td>163 ± 153</td>
<td>222 ± 187</td>
<td>39.71 ± 29.20</td>
<td>43.86 ± 30.77</td>
</tr>
</tbody>
</table>

Continuous intracranial EEG (iEEG) from bilateral amygdala, hippocampus and anterior nucleus of thalamus enabled direct assessment of patient seizure diary reports and iEEG recorded electrographic seizure activity. On average 39.7% of iEEG electrographic seizures had a patient seizure diary report within ± 30 min of the iEEG event and 56.14% of the patient reported seizures had associated iEEG confirmed events. These results from continuous ambulatory iEEG in natural environments demonstrate the complex and inaccurate relationship between patient diary reports and gold-standard verified seizures from continuous iEEG. Note: All four patients had seizures independently from left and right hippocampus. The table includes all seizures, except for MH2* where the results are for right-hippocampal seizures only because of very frequent electrographic seizures (average 30/day) originating from the left hippocampal remnant of a prior left anterior temporal lobectomy. The excellent reporting for patient MH4** included her caregiver reports. The accuracy of patient reporting of seizures is currently confounded by fact that iEEG seizures may be truly subclinical seizures or amnestic seizures that the patient does not recall.
In summary, we present results from a powerful system integrating a new investigational neural sensing and stimulation device with local and distributed computing that should prove useful for future investigation and optimization of EBS in drug-resistant epilepsy. This research identifies areas for future research including bi-directional interfaces to enable iEEG event triggered behavioural assessments,47 continuous behavioural state tracking,48 seizure forecasting and adaptive EBS therapy. Future implantable systems with greater device computational power and data storage capacity will enable smart sampling paradigms to buffer data, run embedded algorithms and trigger alarms for therapy change, behavioural queries and data transfer that should enhance understanding of behaviour and brain activity.

Acknowledgements
The authors thank the people with epilepsy for participating in this research. We appreciate the partnership with pet owners who made the canine research possible. We appreciate the technical and patient-focused support provided by Cindy Nelson and Karla Crockett. They thank Certicon a.s. for use of CyberPSG tool for visual review of EEG and Medtronic Plc for providing the investigational Medtronic Summit RC+STM devices used in this research. The authors thank the research and clinical partners who provided the NeuroVista Inc. data used for developing the seizure detector. This research benefited from the community expertise and resources made available by the NIH Open Mind Consortium NIH U24-NS113637 (https://openmind-consortium.github.io/).

Funding
This research was supported by the National Institutes of Health: UH2/UH3-NS95495 and R01-NS09288203. European Regional Development Fund-Project ENOCH (No.CZ.02.1.01/0.0/0.0/16_019/0000868) Ministry of Education, Youth and Sports of the Czech Republic project no. LTAUSA18056. Additional support was provided by Epilepsy Foundation of America Innovation Institute and Mayo Clinic Benefactors, Mayo Clinic Graduate School of Biomedical Sciences (IB, VM and LW), and Czech Technical University, Prague, Czech Republic (VK), and grant FEKT-K-22: 7649 realized within the project Quality Internal Grants of Brno University of Technology (KInG BUT), Reg. No. CZ.02.2.69/0.0/0.0/19_073/0016948 (FM), and The International Clinical Research Centre at St. Anne’s University Hospital (FNUSA-ICRC), Brno Czech Republic.

Competing interests
G.A.W., V.S., V.K. and B.H.B. declare intellectual property disclosures related to behavioural state and seizure classification algorithms. G.A.W., B.H.B., J.J.V.G. and B.N.L. declare intellectual property licensed to Cadence Neuroscience Inc. G.A.W. has licensed intellectual property to NeuroOne, Inc. G.A.W. and B.N.L. are investigators for the Medtronic Deep Brain Stimulation Therapy for Epilepsy Post-Approval Study. T.D. is an advisor for Synchron and Cortec, and has business agreements for research tools with Bioinduction Ltd and Magstim Ltd; all unrelated to this content. V.S.M. is a paid summer intern at Medtronic. V.K. consults for Certicon a.s. The remaining authors declare that they have no competing interests. The investigational Medtronic Summit RC+STM devices used in this research were provided free of charge as part of NIH Brain Initiative UH2/UH3-NS95495.

References