



## Review

## Big data in status epilepticus

Steven N. Baldassano<sup>a,b,\*</sup>, Chloé E. Hill<sup>c</sup>, Arjun Shankar<sup>a,b</sup>, John Bernabei<sup>a,b</sup>,  
Pouya Khankhanian<sup>c,d</sup>, Brian Litt<sup>a,b,d</sup>

<sup>a</sup> Department of Bioengineering, University of Pennsylvania, 210 South 33rd Street, Philadelphia, PA 19104, United States

<sup>b</sup> Center for Neuroengineering and Therapeutics, University of Pennsylvania, 240 South 33rd Street, Philadelphia, PA 19104, United States

<sup>c</sup> Department of Neurology, University of Michigan, 1500 East Medical Center Drive, Ann Arbor, MI 48109, United States

<sup>d</sup> Department of Neurology, Penn Epilepsy Center, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104, United States

## ARTICLE INFO

## Article history:

Received 30 June 2019

Accepted 26 July 2019

Available online 21 August 2019

## Keywords:

Epilepsy

Status epilepticus

Big data

Natural language processing

Continuous EEG

Multimodal data

## ABSTRACT

Status epilepticus care and treatment are already being touched by the revolution in data science. New approaches designed to leverage the tremendous potential of “big data” in the clinical sphere are enabling researchers and clinicians to extract information from sources such as administrative claims data, the electronic medical health record, and continuous physiologic monitoring data streams. Algorithmic methods of data extraction also offer potential to fuse multimodal data (including text-based documentation, imaging data, and time-series data) to improve patient assessment and stratification beyond the manual capabilities of individual physicians. Still, the potential of data science to impact the diagnosis, treatment, and minute-to-minute care of patients with status epilepticus is only starting to be appreciated. In this brief review, we discuss how data science is impacting the field and draw examples from the following three main areas: (1) analysis of insurance claims from large administrative datasets to evaluate the impact of continuous electroencephalogram (EEG) monitoring on clinical outcomes; (2) natural language processing of the electronic health record to find, classify, and stratify patients for prognostication and treatment; and (3) real-time systems for data analysis, data reduction, and multimodal data fusion to guide therapy in real time. While early, it is our hope that these examples will stimulate investigators to leverage data science, computer science, and engineering methods to improve the care and outcome of patients with status epilepticus and other neurological disorders.

**This article is part of the Special Issue “Proceedings of the 7th London-Innsbruck Colloquium on Status Epilepticus and Acute Seizures”**

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

Technological advancements in clinical data acquisition and storage are rapidly changing the face of patient care for status epilepticus. In multiple domains, we are now acquiring large-scale data that make it possible to explore previously infeasible clinical and basic science questions [1–3]. For instance, these data come as clinical reports from health-system-wide electronic health records (EHRs) and billing claims from administrative databases, as well as in the form of digital neurophysiologic recordings from continuous electroencephalogram (EEG) or multimodality intensive care unit (ICU) monitoring. While powerful, each dataset has its own challenges: data can be lacking in clinical detail, noisy, uncurated, or poorly annotated. Claims data are collected for billing purposes and therefore provide little clinical context; however, rigorous analysis and careful interpretation allow important inferences

from populations of millions of patients. In the case of clinic notes in the EHR, information is not uniformly documented or organized, making automated extraction of relevant data challenging. These issues, combined with the sheer size of these datasets, necessitate the use of new analysis paradigms and technical approaches, such as algorithmic data mining and machine learning, to perform data reduction and glean clinical insights.

The era of “big data” in clinical medicine offers potential to improve care for status epilepticus through several avenues. One goal is to use clinical records to stratify patients, both for prognostication and to identify patient groups that may benefit from particular interventions such as continuous EEG (cEEG) monitoring, surgery, or neuromodulation. In this way, personalized data mining algorithms can be used to optimize and individualize care. Secondly, extensive intracranial EEG (iEEG) datasets recorded using next-generation implantable devices are capable of driving powerful machine learning algorithms for data interpretation, epileptiform event detection, and per-patient optimization of therapeutic stimulation parameters. This streaming data analysis

\* Corresponding author at: 2520 Waverly Street, Philadelphia, PA 19146, United States.  
E-mail address: [stevennb@pennmedicine.upenn.edu](mailto:stevennb@pennmedicine.upenn.edu) (S.N. Baldassano).

paradigm is now being translated into the neurointensive care unit, where multimodal data inform algorithms for data reduction, automated data monitoring, and notifying caretakers regarding changes in clinical status in near real time. Finally, these data are driving researchers and clinicians to establish data-sharing platforms to propel collaborative research initiatives investigating the mechanisms underlying the pathology of epilepsy and status epilepticus [4–7].

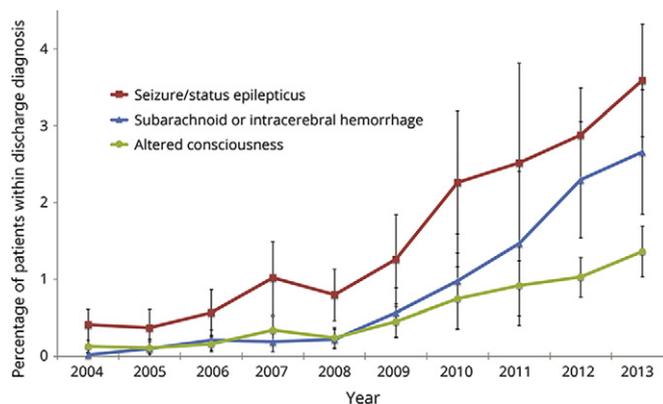
## 2. Continuous EEG and hospitalization outcomes for patients with status epilepticus

Administrative claims datasets offer a unique opportunity to analyze data from millions of patients at once and to assess trends in healthcare utilization over time. These data are collected by insurers (both private and public insurance datasets are available) and typically contain deidentified information regarding patient demographics, measures of comorbidity, diagnoses, procedures, admission and disposition details, payer information, and hospital characteristics. This type of data is particularly advantageous when studying the care of patients with status epilepticus, a relatively rare occurrence with an annual incidence of 10–41 per 100,000 population [8].

Continuous EEG, a tool that provides noninvasive real-time information about brain function, plays an important role in the diagnosis and management of status epilepticus [9,10]. The American Clinical Neurophysiology Society (ACNS) recommends the use of cEEG in seizure diagnosis in patients with critical illness as there is a very high incidence of status epilepticus in this population [11] – 8–68% of intensive care patients who are undergoing EEG monitoring are found to have seizures [10]. Continuous EEG is critical to identifying subclinical seizures, which have been associated with neuronal injury [12,13] and neurological decline [14,15]. However, the impact of cEEG utilization on outcomes for patients with critical illness has not been well established. Furthermore, there is great variability in application of cEEG because of resource availability and variation in provider practices [16,17]. While a randomized controlled trial of cEEG is underway in Europe (“Continuous EEG Randomized Trial in Adults”) [18], this trial will be limited by logistical and ethical considerations, challenges that administrative data analysis does not face.

Trends in the utilization of cEEG can be readily characterized with claims data, as many administrative datasets offer a large nationwide sample of patients and serial years of data. Evaluation of the National Inpatient Sample (NIS), the largest all-payer inpatient national database through the Agency for Healthcare Research and Quality (AHRQ) [19], found an increase in the proportion of patients with critical illness receiving cEEG from 0.06% in 2004 to 0.8% in 2013. This increase in utilization was observed across all considered indications: status epilepticus, subarachnoid or intracerebral hemorrhage, and altered consciousness (Fig. 1). Yet, cEEG monitoring was only administered for a very small minority of patients with critical illness (0.3%) and only employed by 6% of the nationally-sampled hospitals [20].

The impact of cEEG monitoring on patient outcomes can also be considered with claims data analysis. Two studies have assessed the association of cEEG monitoring with inpatient mortality within the population with critical illness. In a study of 40,945 ventilated patients from the NIS dataset from 2005 to 2009, undergoing cEEG was associated with a lower risk of inhospital mortality as compared to patients who received routine EEG only (OR [odds ratio]: 0.63; 95% CI [confidence interval]: 0.51–0.76;  $p < 0.001$ ); including epilepsy/convulsion diagnoses as covariates did not alter these findings [21]. A second study of 2004–2013 NIS data identified 7,102,399 ventilated patients, and found a similar association of cEEG with decreased inpatient mortality as compared to patients who did not undergo cEEG (OR: 0.83; 95% CI: 0.75–0.93;  $p < 0.001$ ); however, this relationship was only significant for patients with subarachnoid/intracerebral hemorrhage or altered mental status, and was not significant for the subgroup with a diagnosis of seizure/status epilepticus [20]. The most plausible explanation for this unexpected



**Fig. 1.** Use of cEEG in critically ill discharges from 2004 to 2013 by diagnostic subcategory. Data derived from National Inpatient Sample database. Numerator is the number of patients who underwent cEEG; denominator is the total discharges with a particular diagnosis.

Figure reproduced with permission from Hill et al., “Continuous EEG is associated with favorable hospitalization outcomes for critically ill patients,” *Neurology*, Jan 2019, 92 (1) e9–e18 [20].

finding is that this subanalysis was confounded by disease severity, as the diagnosis of seizure or status epilepticus identifies a highly heterogeneous population with a range of seizure etiologies and related prognoses.

While limited by the absence of detailed clinical information, these administrative claims analyses illustrate the current utilization of cEEG, which is growing at a rapid pace, and should inform the appropriate expansion of cEEG moving forward. Continuous EEG appears beneficial to patients with critical illness, however, targeting is critical as the observed benefit varies across subgroups of patients.

Claims data offer unique opportunities to find associations between diagnostic approaches, therapies, and outcomes in patients with status epilepticus, although it is important to note that such studies are not a substitute for prospective, randomized clinical trials. Rather, studies of this nature can inform the objectives and design of clinical trials such as the Established Status Epilepticus Treatment Trial (ESETT) [22,23]. As health system infrastructures continue to advance across the world, particularly with consolidation of digital medical records [24] and more centralized care delivery, administrative claims research is likely to become increasingly powerful and informative.

## 3. Natural language processing for interpreting epilepsy clinical notes

### 3.1. Natural language processing for clinical data

Large EHR datasets can reveal patterns in patient populations that are useful for both clinical and research applications. These datasets include clinical notes (such as admission notes, progress notes, discharge summaries, operative notes, and outpatient clinic notes) and diagnostic reports (such as radiology and pathology reports). However, much of EHR data are recorded as unstructured free text [25]. As human review of unstructured text is time intensive and inefficient, natural language processing (NLP) can greatly facilitate analysis of large EHR datasets [26]. Natural language processing is a range of computational techniques used to analyze and process human language [27], critical for information extraction (IE) from unstructured text [28]. In clinical domains, researchers have used NLP IE to aid clinical decision support systems [29,30], identify patient cohorts [31,32], and identify risk factors from clinical notes [33]. While current applications of NLP in epilepsy have so far been limited by algorithm efficacy and generalizability, increasing study sizes in genetic and precision medicine trials are making NLP essential for epilepsy phenotyping and stratification [26].

### 3.2. Epilepsy specific ontologies

Epilepsy datasets are particularly challenging as they are generated in a wide range of settings using differing terminologies. A seizure with alteration of consciousness may be called a complex partial seizure, a focal dyscognitive seizure, or a dialeptic seizure by different physicians [34]. One way to standardize a vocabulary is to create an ontology, a formal representation of knowledge of a given domain that allows both humans and computers to consistently and accurately interpret terms. Ontologies are now commonly used across biomedical domains to simplify data management, as they enable NLP tools to better extract heterogeneous information from clinical free text [35].

A major milestone in epilepsy NLP research was the creation of a comprehensive epilepsy specific ontology by Sahoo et al. in 2012 [36]. Epilepsy and Seizure Ontology (EpSO) integrates recommendations from the International League Against Epilepsy (ILAE) Classification and Terminology Commission with elements from existing ontologies like the Neural ElectroMagnetic Ontologies. Epilepsy and Seizure Ontology includes very specific terms for epilepsy types, seizure types, EEG patterns, symptoms, and brain locations. It uses the World Wide Web Consortium recommended Web Ontology Language (OWL2), a formal knowledge representation language based on description logic that supports software applications [34]. While EpSO has been utilized in several NLP systems, it is not the only ontology currently used in epilepsy NLP research. Several studies from Portugal using text mining of medical records to classify epilepsy in children utilized an ontology based on the Unified Medical Language System and the International Classification of Diseases (ICD9) for epilepsy types. This ontology had additional Portuguese words to aid computer analysis of the medical notes [30,37,38]. Other sources of standardized epilepsy terms include the SNOMED-CT (Systematized Nomenclature of Medicine – Clinical Terms) [39], and HPO (Human Phenotype Ontology) [40]. In 2017, the ILAE released an updated classification of seizure types to allow more flexibility in diagnosis [41]. The ILAE also released an updated classification of epilepsy types with new terminology such as developmental epilepsy and epileptic encephalopathy [42]. These 2017 ILAE classification systems are expected to remain stable for at least a decade to come. As EpSO is now outdated, the field of epilepsy NLP would greatly benefit from a new OWL2 encoded ontology that reflects the updated ILAE recommendations. It is also important to note that these approaches, while efficient for analysis once data are entered, likely sacrifice some of the richness of unstructured data and suffer from the need to revise data entries as terminology, classification, and their scientific underpinnings evolve over time.

### 3.3. Integrated tools for NLP-based epilepsy data management and cohort identification

Epilepsy and Seizure Ontology provided the backbone to a suite of epilepsy specific informatics tools which were developed for a

multicenter study funded by the National Institute of Neurological Disorders [31,32,36,43]. The OPIC system, Ontology-driven Patient Information Capturing system for epilepsy, was developed in 2012 to facilitate structured entry of multimodal epilepsy data according to the EpSO format [36]. Another software tool based on EpSO is the Epilepsy Data Extraction and Annotation (EpiDEA) system. The EpiDEA system utilizes Mayo Clinic's open source NLP software tool cTAKES (clinical Text Analysis and Knowledge Extraction System) to identify and analyze EpSO terminology in Epilepsy Monitoring Unit (EMU) discharge summaries for patient cohort identification [31]. In 2014, the OPIC and EpiDEA systems were combined into the Multi-Modality Epilepsy Data Capture and Integration System, MEDCIS, developed to facilitate data integration for large-scale multicenter EMU studies [43–45]. The robustness of these tools rests in the use of EpSO as the common reference terminology across all software components, with the goal of increasing generalizability across health centers. A timeline of milestones for development of NLP tools for epilepsy is shown in Fig. 2.

### 3.4. Comparison of recent epilepsy information extraction studies and future directions

Despite the abundance of open source software tools, there are only a few published studies on epilepsy IE from the EHR. Natural language processing performance is often measured by the following three metrics: precision, the fraction of retrieved items relevant to the query; recall, the fraction of relevant items successfully retrieved; and F1 score, the harmonic mean of precision and recall. The EpiDEA system developed by Cui et al. achieved overall precision, recall, and F1 scores of 93.59%, 84.01%, and 88.53%, respectively, for extraction of EEG patterns and medication information from 104 EMU discharge summaries [31]. Cui et al. also developed the PEEP (Phenotype Extraction in Epilepsy) system for extracting the epileptogenic zone, seizure semiology, lateralizing sign, interictal and ictal EEG pattern from discharge summaries. Phenotype Extraction in Epilepsy achieved overall precision, recall, and F1 scores of 92.4%, 93.1%, and 92.7%, respectively, when tested over 262 discharge summaries [32]. Most recently, Fonferko-Shadrach et al. developed the ExECT (Extraction of Epilepsy Clinical Text) system which retrieves epilepsy diagnosis, epilepsy type, focal seizures, generalized seizures, seizure frequency, medications, CT (computed tomography), MRI (magnetic resonance imaging), and EEG data from clinic letters. Extraction of Epilepsy Clinical Text achieved overall precision, recall, and F1 scores of 91.4%, 81.4%, and 86.5%, respectively [46]. The gold standard in all studies consisted of manual annotations created by clinicians. Though the EpiDEA and PEEP systems have higher scores on the performance metrics than ExECT, this could be because EpiDEA and PEEP utilize EMU discharge summaries while ExECT utilizes clinic letters. Discharge summaries tend to be much more organized than clinic notes, making them easier for an NLP system to classify.

The improvement in classifier performance on prestructured data compared to free-form notes is clear; however, there are several

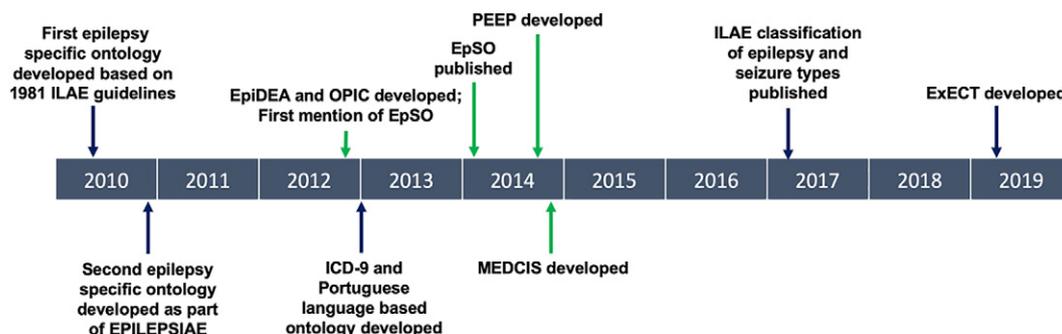


Fig. 2. Timeline of milestones for NLP in epilepsy. Integrated EpSO-based tools are highlighted with green arrows. EPILEPSIAE = Evolving Platform for Improving the Living Expectations of Patients Suffering from Ictal Events.

drawbacks of mandating a structured note format. Predefined categories are not always sufficient to describe every patient and lack the detail and customizability of free-form prose. Clinicians pressed for time may also be less careful with cumbersome, structured data entry systems and may autopopulate forms with default options. Continued work is needed to identify the optimal combination of structured and unstructured data entry to maximize interpretability while offering convenience in clinician workflow. To this end, the field of epilepsy research could benefit greatly from updated data entry tools to balance clinician and researcher interests, an updated ontology to reflect ILAE changes, and increased use of the existing software tools for data organization and NLP.

#### 4. Distributed systems for real-time analysis of intracranial EEG

##### 4.1. Next generation implantable devices for iEEG telemetry

The area with likely the most immediate potential impact of “big data” on the care and science of status epilepticus is the real-time analysis of neurophysiologic data. The next generation of implantable monitoring and therapeutic devices are producing vast amounts of iEEG data capable of being transmitted to analysis systems in real time, and will soon be deployed in the care and prevention of status epilepticus. Traditionally, the computational capabilities of neuroimplantable devices [47,48], such as the NeuroPace RNS (Responsive Neurostimulation) system [49], have been limited by device processing power and battery life [50]. Furthermore, such devices typically store only brief segments of iEEG for intermittent review by a physician, precluding evaluation of the continuous record or trending of iEEG features over time. A promising approach for improving the utility of implanted devices is to transmit data from the device to a remote analysis platform [51,52], such as a hand-held personal assist device (PAD) or cloud-based server [53] (Fig. 3A). With continuous or intermittent data telemetry, patient data can be archived and annotated in centralized repositories accessible to physicians [54], and can also be exposed to offboard computing resources and custom algorithms.

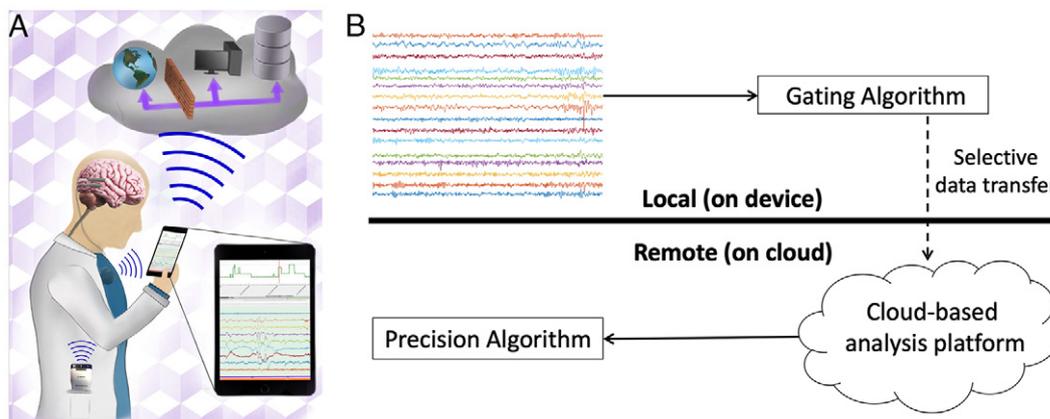
A critical benefit of telemetry-capable devices has been the production of high-quality continuous iEEG datasets for both research and clinical use. Automated seizure detection, seizure diary generation, seizure prediction, and behavioral state classification have been demonstrated using the NeuroVista Seizure Advisory System (SAS), an implanted device providing telemetry of iEEG data coupled to off-body analytics, in both humans [55–57] and canines [58–61]. These novel datasets have

also been featured in [kaggle.com](https://www.kaggle.com) competitions to crowdsource development of seizure detection [56] and prediction [61,62] algorithms, which have produced validated, generalizable, open-source methods at a fraction of the cost of traditional research studies. These datasets have made it possible to investigate questions that require continuous interictal data, such as long-term trends in recording metrics [63], temporal relationships among seizures [64,65], and identification of iEEG biomarkers [66]. Clinically, continuous iEEG data streams offer a means of circumventing reliance on self-reported seizure diaries, which suffer from poor accuracy and hamper assessment of therapy efficacy [67,68].

##### 4.2. Continuous iEEG datasets for optimizing neuromodulation

Neuromodulation (in the form of vagal nerve stimulation (VNS), responsive nerve stimulation (RNS), or deep brain stimulation (DBS)) is a viable treatment option for patients with medication-refractory seizures who are not candidates for surgical intervention [69,70]. While neurostimulation has been shown to reduce seizure frequency by 30–50% on average [71–73], optimizing of stimulation parameters remains a significant challenge. For instance, in the case of the NeuroPace RNS device, the first FDA (Food and Drug Administration)-approved closed-loop stimulator for epilepsy, stimulation parameters such as frequency, pulse width, and burst duration are manually tuned during clinic visits in an effort to decrease seizure frequency [74]. There is also a lack of convincing evidence in the literature as to the optimal location for cortical or subcortical stimulation [75], or whether stimulation should be directed in response to seizure activity [71,76,77] or chronic and subthreshold [78,79].

The next generation of implanted devices with continuous streaming capabilities offers exciting potential to generate comprehensive datasets for algorithmic optimization of neurostimulation. While novel strategies for adjustment of stimulation parameters, such as gradient descent learning [80] and physiologic adaptation [81], have been applied in animal models, they have not yet been translated to human studies [82]. Ongoing work using an investigational implantable neurostimulator (Summit System RC + S, Medtronic Inc.) in a canine model aims to leverage off-body algorithms for optimizing stimulation, in combination with two-way telemetry, to actively probe seizure networks [52,53]. Given the wealth of data produced by continuous iEEG records, it is now possible to dynamically adjust therapeutic stimulation parameters, guided by surrogate interictal markers of seizures such as interictal epileptiform discharges (IEDs), evoked responses, and brain state [52]. Advances in hardware, software, and real-time analysis



**Fig. 3.** Implanted devices with cloud connectivity. (A) Schematic of next-generation epilepsy management system using Medtronic investigational Summit RC + S. The device interfaces with a local PAD and the cloud to create a flexible platform with local and distributed computing, analytics, and data storage. Figure reproduced with permission from Kremen et al., “Integrating Brain Implants with Local and Distributed Computing Devices: A Next Generation Epilepsy Management System,” *IEEE/JTEHM*, Sep 2018, Vol 6 [53]. (B) Seizure detection paradigm using cloud computing. Potential seizure clips are identified using a hypersensitive, onboard seizure detector. These candidate clips are transmitted to a cloud platform for analysis using a highly accurate, computationally intensive algorithm. Figure reproduced with permission from Baldassano et al., “Cloud computing for seizure detection in implanted neural devices,” *Journal of Neural Engineering*, Feb 2019, 16 (2) [88].

systems are likely to bring these techniques into the management of patients with status epilepticus.

#### 4.3. Cloud computing for seizure detection on intracranial EEG

Effective closed-loop neurostimulation requires highly accurate algorithms to classify brain signals and detect seizures [83–85]. While many machine learning algorithms to detect seizure activity have been published [56,86], such methods are typically far too complex to be implemented in an implanted device. As a result, current responsive stimulation devices, such as the NeuroPace RNS and Medtronic Activa PC + S, are limited to simple feature sets and classifiers [47,48,87] and are consequently plagued by high rates of false detections [72].

Next-generation implantable devices capable of intermittent data telemetry may offer substantial improvements in seizure detection accuracy through offboarding of computationally-intensive algorithms. Our group recently proposed a novel paradigm for implementing real-time complex seizure detection algorithms on implanted device iEEG data [88]. This method (Fig. 3B) relies on two sequential seizure detection algorithms. First, a simple “gating” algorithm running onboard the device is used to control selective transmission of data from the device to a cloud storage and analysis platform. The gating algorithm is designed to be compatible with the computational limitations of an implanted device, and is tuned to be hypersensitive in order to identify every seizure event for transmission. Transmitted data are then analyzed by a “precision” algorithm, a much more accurate and computationally intensive algorithm deployed using cloud computing resources. Evaluated using continuous, human iEEG data collected from the NeuroVista SAS [55], this system achieved seizure detection accuracy far outpacing that of existing neuroimplantable devices with significantly lower false positive rates [72]. This early work is encouraging, as by using telemetry to access cloud computational resources, the next generation of implanted devices will be able to employ much more sophisticated algorithms to improve device performance and patient outcomes.

### 5. Real-time analysis of multimodal ICU data

#### 5.1. Quantitative approaches to continuous EEG

Despite its benefits for detection of nonconvulsive status epilepticus (NCSE) [89], a condition with a mortality rate greater than 50% [90], cEEG is underutilized in part because of the significant labor required to initiate recordings and interpret the vast amount of data generated. In many cases, around-the-clock staffing of physicians and EEG technologists to handle the unpredictable needs of cEEG monitoring presents a prohibitive level of cost [91]. Furthermore, cEEG is often read at infrequent intervals which could lead to hours of delay between the onset of a seizure and notification of the event to the patient's physician. In instances that require rapid clinical action, the therapeutic window may pass before cEEG is interpreted or a patient's condition may deteriorate and result in an adverse outcome. Automated methods of cEEG interpretation and detection of NCSE hold the potential to solve some of these problems and increase the utilization of cEEG at both large centers and small hospitals.

While there are several approved algorithms for seizure detection in scalp EEG, they suffer from a variety of drawbacks which preclude their widespread clinical use. The most prominent EEG monitoring algorithm is the FDA-approved Reveal algorithm implemented as part of Persyst's clinical EEG system. Reveal was initially reported to have a sensitivity of 76% with a false positive rate of 0.11/h [92], though subsequent studies have found a much higher false positive rate [93]. This level of performance leads to a significant proportion of seizures going undetected, as well as a substantial number of false alarms. One accepted method for providing semiautomated ICU EEG analysis is known as quantitative EEG (qEEG) in which signals are evaluated in real time, displaying

metrics such as amplitude, rhythmicity, and spectral features in a visual interface [94]. Quantitative EEG can be presented in a temporally compressed format allowing long trends to be quickly visualized, such as in the NeuroTrend algorithm [95]. However, while qEEG may significantly reduce review time by the clinician, sensitivity for seizure identification may be as low as 51–67% [96], limiting its clinical reliability.

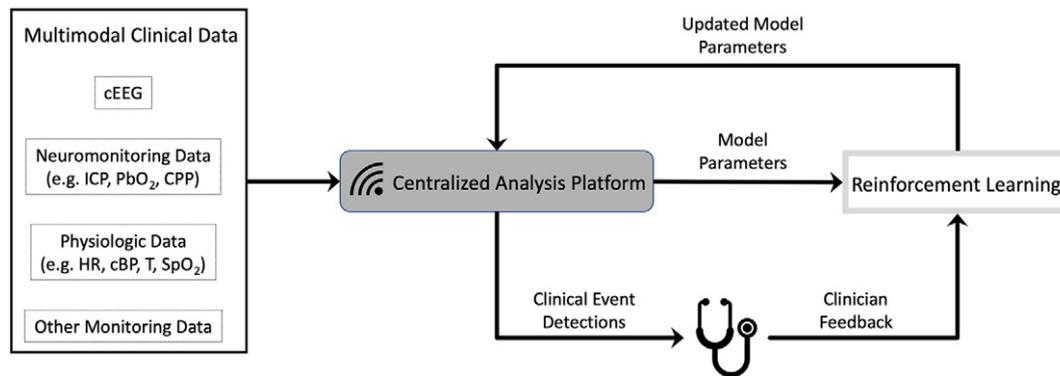
Approaches that leverage innovative methods in data science and machine learning may allow future algorithms for automated and semiautomated cEEG analysis to be more impactful than previous attempts. One such method uses quantitative features and a simple classifier to identify patients as belonging to the following groups: normal, isoelectric, slowing, low-voltage, burst suppression, or containing discharges or seizures [97]. While classification accuracy was 85% in the test set, this algorithm is not designed to perform in real time and would not necessarily improve early detection of seizure activity [97]. ICU-ASDA (ICU Automated Seizure Detection Algorithm) is an algorithm which uses signal amplitude variation and a regularity statistic to detect changes in rhythmicity [98]. This algorithm detected seizures with a 90.4% sensitivity and a false positive rate of 1.6 per 24 h in a dataset with 24 patients of which 11 had nonconvulsive seizures [98]. However, the authors only reported training accuracy, and given their small sample size, the generalizability of this algorithm is unknown. There remains a clear need for high-performance algorithms designed for real-time deployment, supported by studies which use large, high-quality, ICU cEEG datasets.

#### 5.2. Multimodal ICU monitoring: an opportunity for translational clinical data platforms

In addition to cEEG, other modalities of continuous physiologic monitoring are increasingly relied upon for neurointensive care of patients in or recovering from status epilepticus [99,100]. Such modalities include advanced monitoring techniques such as intracranial pressure probes [101], tissue oxygen probes [102], thermal diffusion probes [103], invasive and noninvasive monitors of cardiac output [104], and near-infrared spectroscopy [105], as well as standard physiologic signals such as heart rate, continuous blood pressure, and oxygen saturation. Just as in the case of cEEG, multimodal monitoring data in the ICU must be manually reviewed by physicians, imposing a significant burden to caretakers and increasing cost, latency of response to clinical events, and risk of missing events altogether. In response, development of clinical informatics to fuse outputs from multiple clinical monitors in a meaningful way for monitoring or prediction is flourishing [106]. The ability to detect nonconvulsive seizures after subarachnoid hemorrhage based on cEEG can be improved by incorporating other physiological data such as heart rate, blood pressure, respiratory rate, and cerebral perfusion pressure [107]. Methods incorporating cEEG in combination with metrics such as intracranial pressure, brain tissue PO<sub>2</sub>, and cerebral microdialysis [108], or demographic and admissions data [109], have been used to detect and monitor cerebral hypoperfusion. Algorithms interpreting multimodal ICU data have also shown efficacy for stratifying patient outcomes after cardiac arrest [110] or severe traumatic brain injury [111].

While such methods for inferring patient status and predicting outcomes are increasingly common, there is nevertheless a dearth of tools for practical data reduction and interpretation that integrate with caretaker workflow. Currently, tremendous amounts of data are generated from clinical monitors but are not centralized or archived to allow for remote viewing, annotation, sharing among providers, or retrospective retrieval. In parallel with the advancement of monitoring, technology must come an advancement in clinical informatics infrastructure [112]. Specifically, manufacturer-specific data systems must be augmented with customizable, modular data platforms that facilitate data storage, visualization, and analytics (Fig. 4).

The need for next-generation data management and analysis platforms is beginning to be addressed by companies such as Blackfyynn,



**Fig. 4.** ICU data analysis pipeline. The next generation of real-time data analysis platforms will incorporate multimodality data and employ machine learning algorithms to provide custom clinical event detections. By incorporating clinician feedback through reinforcement learning, these algorithms can intelligently adapt over time. cEEG = continuous EEG, ICP = intracranial pressure, PbO<sub>2</sub> = brain tissue oxygen, CPP = cerebral perfusion pressure, HR = heart rate, cBP = continuous blood pressure, T = continuous temperature, SpO<sub>2</sub> = blood oxygen saturation.

Inc. [113,114], a secure cloud-based platform for sharing and analyzing medical data, cofounded by members of our group. Ongoing work in our lab also seeks to meet this need through development of a modular ICU data monitoring platform called IRIS (ICU Real-Time Informatics System) (unpublished). This automated system is capable of streaming ICU data from multimodal clinical monitors in real time, archiving the data on a central server, and carrying out custom analysis modules. Early proof-of-concept trials have included modules to detect faulty cEEG electrodes, automatically trend burst-suppression ratios, and detect sustained elevations in intracranial pressure. The system is linked to a custom, HIPAA (Health Insurance Portability and Accountability Act)-compliant API (Application Programming Interface) and smartphone application to securely deliver notifications of detected clinical events to caretakers including physicians, nurses, and EEG technicians. Such platforms have great potential to benefit patient care by improving ICU workflow, reducing the cost and caretaker burden associated with continuous monitoring, and exposing existing data streams to state-of-the-art machine learning algorithms. Furthermore, by designing platforms with a modular structure, it is possible to add custom modules based on clinical needs such as methods of seizure detection, spike detection, or vasospasm monitoring. Future work may also integrate therapy in a closed-loop system, for applications including titration of medications for treatment of status epilepticus or maintenance of general anesthesia. We hope that our work will encourage the continued implementation of translational medical data systems to optimize care of patients with epilepsy undergoing neurointensive care.

### 5.3. Economic impact of automated ICU data analysis

Automated analysis of ICU data may have significant ramifications for healthcare costs in the ICU. As ICU datasets become larger and more descriptive, tools leveraging these data for prognostication will become more effective. In the case of patients with critical illness in status epilepticus, improved outcome prediction has potential to significantly influence goals of care and resource allocation at the end of life [115–117]. While automated scoring systems may not yet outperform ICU physicians in predicting mortality [118], their ability to fuse large amounts of multimodal data promises to reveal clinical insights that may affect management decisions. Furthermore, automated analysis of continuous monitoring data will, in itself, change revenue models in critical care. Continuous monitoring studies, such as 24-h EEG, are costly to patients [119,120], in part because of the need for manual physician review. The implications of this economic impact remain to be seen. In the current, private-insurance driven model, recommendations for continuous monitoring may be monetarily disincentivized by hospitals due to lower billing potential; however, if healthcare payments shift toward

single-payer or bundled payments, automated monitoring studies may be more prevalent due to increased cost effectiveness.

## 6. Future directions and conclusion

The revolution of “big data” in status epilepticus has already begun, and tools for leveraging these data are flourishing. Yet, for extracting data from clinical sources such as claims data and the EHR, we have merely scratched the surface. Part of future progress in these areas is predicated on developing new methods for NLP, integrating clinical notes with imaging and laboratory data, and mining health systems databases. An equally important part will be driven by ambitious pushes to share data and algorithms through collaborative platforms, and to encourage investigators to work together to pose powerful questions. In the realm of physiologic monitoring, implanted devices and ICU systems currently in development are capable of streaming EEG and multimodality data to local computing systems or the cloud. By implementing distributed data reduction and interpretation algorithms, these systems provide real-time analytics for treatment of status epilepticus. As we continue to improve technologies for acquiring, storing, and accessing “big data”, we will be able to rapidly design, test, and employ data-driven methods and classifiers. “Big data” are being established as an agent for change in healthcare that will impact clinical practice, decision-making, and care guidelines. In addition, population-based datasets offer opportunities to compare medical management decisions, costs, and outcomes across different health systems, potentially providing new perspectives on healthcare among countries with a range of payment models. Such information may play an important role in legislative decisions reshaping the US healthcare system. Through data science and engineering, we have potential to improve the quality and efficiency of care for status epilepticus and other neurologic disorders.

## Acknowledgments

Funding for studies described in this review provided by the Mirowski Family Fund, Neil and Barbara Smit, the Family of Jonathan Rothberg, the American Epilepsy Society/Epilepsy Foundation Research and Training Fellowship for Clinicians, the Ashton Fellowship at the University of Pennsylvania, and the National Institutes of Health (NIH) (UH2-NS095495-01, R01NS092882, 1K01ES025436-01).

## Declaration of competing interest

Brian Litt is a cofounder of Blackfynn, Inc., an advisor for EpilepsyCo, and has licensed intellectual property to NeuroPace, Inc., through the University of Pennsylvania. No other authors have disclosures to report.



- [61] Brinkmann BH, Wagenaar J, Abbot D, Adkins P, Bosshard SC, Chen M, et al. Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. *Brain* 2016;139:1713–22.
- [62] Kuhlmann L, Karoly P, Freestone DR, Brinkmann BH, Temko A, Barachant A, et al. Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain* 2018;141:2619–30.
- [63] Ung H, Baldassano SN, Bink H, Krieger AM, Williams S, Vitale F, et al. Intracranial EEG fluctuates over months after implanting electrodes in human brain. *J Neural Eng* 2017;14:056011.
- [64] Karoly PJ, Nurse ES, Freestone DR, Ung H, Cook MJ, Boston R. Bursts of seizures in long-term recordings of human focal epilepsy. *Epilepsia* 2017;58:363–72.
- [65] Karoly PJ, Ung H, Grayden DB, Kuhlmann L, Leyde K, Cook MJ, et al. The circadian profile of epilepsy improves seizure forecasting. *Brain* 2017;140:2169–82.
- [66] Roehri N, Bartolomei F. Are high-frequency oscillations better biomarkers of the epileptogenic zone than spikes? *Curr Opin Neurol* 2019;32:213–9.
- [67] Elger CE, Mormann F. Seizure prediction and documentation—two important problems. *Lancet Neurol* 2013;12:531–2.
- [68] Hoppe C, Poepel A, Elger CE. *Epilepsy*. Arch Neurol 2007;64:1595.
- [69] Nune G, DeGiorgio C, Heck C. Neuromodulation in the treatment of epilepsy. *Curr Treat Options Neurol* 2015;17:375.
- [70] Dalkic EB. Neurostimulation devices used in treatment of epilepsy. *Curr Treat Options Neurol* 2017;19:7.
- [71] Heck CN, King-Stephens D, Massey AD, Nair DR, Jobst BC, Barkley GL, et al. Two-year seizure reduction in adults with medically intractable partial onset epilepsy treated with responsive neurostimulation: final results of the RNS System Pivotal trial. *Epilepsia* 2014;55:432–41.
- [72] Sun FT, Morrell MJ. Closed-loop neurostimulation: the clinical experience. *Neurotherapeutics* 2014;11:553–63.
- [73] Fisher R, Salanova V, Witt T, Worth R, Henry T, Gross R, et al. Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy. *Epilepsia* 2010;51:899–908.
- [74] Sun FT, Morrell MJ. The RNS system: responsive cortical stimulation for the treatment of refractory partial epilepsy. *Expert Rev Med Devices* 2014;11:563–72.
- [75] Schulze-Bonhage A. Brain stimulation as a neuromodulatory epilepsy therapy. *Seizure* 2017;44:169–75.
- [76] Krook-Magnuson E, Armstrong C, Oijala M, Soltesz I. On-demand optogenetic control of spontaneous seizures in temporal lobe epilepsy. *Nat Commun* 2013;4:1376.
- [77] Armstrong C, Krook-Magnuson E, Oijala M, Soltesz I. Closed-loop optogenetic intervention in mice. *Nat Protoc* 2013;8:1475–93.
- [78] Lundstrom BN, Worrell GA, Stead M, Van Gompel JJ. Chronic subthreshold cortical stimulation: a therapeutic and potentially restorative therapy for focal epilepsy. *Expert Rev Neurother* 2017;17:661–6.
- [79] Lundstrom BN, Van Gompel J, Britton J, Nickels K, Wetjen N, Worrell G, et al. Chronic subthreshold cortical stimulation to treat focal epilepsy. *JAMA Neurol* 2016;73:1370.
- [80] Panuccio G, Guez A, Vincent R, Avoli M, Pineau J. Adaptive control of epileptiform excitability in an in vitro model of limbic seizures. *Exp Neurol* 2013;241:179–83.
- [81] Ward MP, Qing KY, Otto KJ, Worth RM, John SWM, Irazoqui PP. A flexible platform for biofeedback-driven control and personalization of electrical nerve stimulation therapy. *IEEE Trans Neural Syst Rehabil Eng* 2015;23:475–84.
- [82] Nagaraj V, Lee ST, Krook-Magnuson E, Soltesz I, Benquet P, Irazoqui PP, et al. Future of seizure prediction and intervention: closing the loop. *J Clin Neurophysiol* 2015;32:194–206.
- [83] Ramgopal S, Thome-Souza S, Jackson M, Kadish NE, Sánchez Fernández I, Klehm J, et al. Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. *Epilepsy Behav* 2014;37:291–307.
- [84] Fountas KN, Smith JR, Murro AM, Politsky J, Park YD, Jenkins PD. Implantation of a closed-loop stimulation in the management of medically refractory focal epilepsy: a technical note. *Stereotact Funct Neurosurg* 2005;83:153–8.
- [85] Sun FT, Morrell MJ, Wharen RE. Responsive cortical stimulation for the treatment of epilepsy. *Neurotherapeutics* 2008;5:68–74.
- [86] Tzallas AT, Tsikalakis DG, Karvounis EC, Astrakas L, Tzaphlidou M, Tsipouras MG, et al. Automated epileptic seizure detection methods: a review study. INTECH Open Access Publisher; 2012.
- [87] Echaz J, Esteller R, Tcheng T, Pless B, Gibb B, Kishawi E, et al. Long term validation of detection algorithms suitable for an implantable device. *Epilepsia* 2001;42 (Suppl. 7).
- [88] Baldassano S, Zhao X, Brinkmann B, Kremen V, Bernabei J, Cook M, et al. Cloud computing for seizure detection in implanted neural devices. *J Neural Eng* 2019;16:026016.
- [89] Kubota Y, Nakamoto H, Egawa S, Kawamata T. Continuous EEG monitoring in ICU. *J Intensive Care Med* 2018;6:39.
- [90] Young GB, Jordan KG, Doig GS. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. *Neurology* 1996;47:83–9.
- [91] Young GB. Continuous EEG monitoring in the ICU: challenges and opportunities. *Can J Neurol Sci* 2009;36(Suppl. 2):S89–91.
- [92] Wilson SB, Scheuer ML, Emerson RG, Gabor AJ. Seizure detection: evaluation of the Reveal algorithm. *Clin Neurophysiol* 2004;115:2280–91.
- [93] Kelly KM, Shiau DS, Kern RT, Chien JH, Yang MCK, Yandora KA, et al. Assessment of a scalp EEG-based automated seizure detection system. *Clin Neurophysiol* 2010;121:1832–43.
- [94] Foreman B, Claassen J. Quantitative EEG for the detection of brain ischemia. *Crit Care* 2012;16:216.
- [95] Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. *Epilepsy Behav* 2015;49:273–9.
- [96] Haider HA, Esteller R, Hahn CD, Westover MB, Halford JJ, Lee JW, et al. Sensitivity of quantitative EEG for seizure identification in the intensive care unit. *Neurology* 2016;87:935–44.
- [97] Cloostermans MC, de Vos CC, van Putten MJAM. A novel approach for computer assisted EEG monitoring in the adult ICU. *Clin Neurophysiol* 2011;122:2100–9.
- [98] Sackellares JC, Shiau D-S, Halford JJ, LaRoche SM, Kelly KM. Quantitative EEG analysis for automated detection of nonconvulsive seizures in intensive care units. *Epilepsy Behav* 2011;22(Suppl. 1):S69–73.
- [99] Suarez JL. Outcome in neurocritical care: advances in monitoring and treatment and effect of a specialized neurocritical care team. *Crit Care Med* 2006;34:S232–8.
- [100] Busl KM, Bleck TP, Varelas PN. Neurocritical care outcomes, research, and technology. *JAMA Neurol* 2019. <https://doi.org/10.1001/jamaneurol.2018.4407>.
- [101] Czosnyka M, Pickard JD. Monitoring and interpretation of intracranial pressure. *J Neurol Neurosurg Psychiatry* 2004;75:813–21.
- [102] Okonkwo DO, Shutter LA, Moore C, Temkin NR, Puccio AM, Madden CJ, et al. Brain oxygen optimization in severe traumatic brain injury phase-II. *Crit Care Med* 2017;45:1907–14.
- [103] Korbakis G, Vespa PM. Multimodal neurologic monitoring. *Handbook of clinical neurology*; 2017. p. 91–105.
- [104] Saugel B, Cecconi M, Wagner JY, Reuter DA. Noninvasive continuous cardiac output monitoring in perioperative and intensive care medicine. *Br J Anaesth* 2015;114:562–75.
- [105] Green MS, Sehgal S, Tariq R. Near-infrared spectroscopy. *Semin Cardiothorac Vasc Anesth* 2016;20:213–24.
- [106] Hemphill JC, Andrews P, De Georgia M. Multimodal monitoring and neurocritical care bioinformatics. *Nat Rev Neurol* 2011;7:451–60.
- [107] Claassen J, Perotte A, Albers D, Kleinberg S, Schmidt JM, Tu B, et al. Nonconvulsive seizures after subarachnoid hemorrhage: multimodal detection and outcomes. *Ann Neurol* 2013;74:53–64.
- [108] Bouzat P, Marques-Vidal P, Zerlauth J-B, Sala N, Suys T, Schoettker P, et al. Accuracy of brain multimodal monitoring to detect cerebral hypoperfusion after traumatic brain injury\*. *Crit Care Med* 2015;43:445–52.
- [109] Amorim E, Rittenberger JC, Zheng JJ, Westover MB, Baldwin ME, Callaway CW, et al. Continuous EEG monitoring enhances multimodal outcome prediction in hypoxic-ischemic brain injury. *Resuscitation* 2016;109:121–6.
- [110] Oddo M, Rossetti AO. Early multimodal outcome prediction after cardiac arrest in patients treated with hypothermia. *Crit Care Med* 2014;42:1340–7.
- [111] Sykora M, Czosnyka M, Liu X, Donnelly J, Nasr N, Diedler J, et al. Autonomic impairment in severe traumatic brain injury: a multimodal neuromonitoring study. *Crit Care Med* 2016;44:1173–81.
- [112] Schmidt JM, De Georgia M. Participants in the international multidisciplinary consensus conference on multimodality monitoring, multimodality monitoring: informatics, integration data display and analysis. *Neurocrit Care* 2014;21(Suppl. 2):S229–38.
- [113] Johnson C, Christini A, Felmeister A, Baglieri C, Kambhampati M, Waanders A, et al. DIPG-51. BLACKFYNN: a secure, cloud-based platform for sharing and analyzing research ready data for pediatric CNS cancers. *Neuro Oncol* 2018;20:i59–i59.
- [114] Christini A, Waanders AJ, Wagenaar JB, Felmeister AS, Santi M, Wadhvani NR, et al. Abstract 2593: accelerating pediatric brain tumor research through team science solutions. *Cancer Res* 2017;77:2593.
- [115] Barnato AE, Angus DC. Value and role of intensive care unit outcome prediction models in end-of-life decision making. *Crit Care Clin* 2004;20:345–62 [vii–viii].
- [116] Campbell ML, Guzman JA. Impact of a proactive approach to improve end-of-life care in a medical ICU. *Chest* 2003;123:266–71.
- [117] Rocker G, Cook D, Sjøkvist P, Weaver B, Finfer S, McDonald E, et al. Clinician predictions of intensive care unit mortality. *Crit Care Med* 2004;32:1149–54.
- [118] Sinuff T, Adhikari NKJ, Cook DJ, Schünemann HJ, Griffith LE, Rocker G, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. *Crit Care Med* 2006;34:878–85.
- [119] Vespa PM, Nenov V, Nuwer MR. Continuous EEG monitoring in the intensive care unit: early findings and clinical efficacy. *J Clin Neurophysiol* 1999;16:1–13.
- [120] Jordan KG. Continuous EEG monitoring in the neuroscience intensive care unit and emergency department. *J Clin Neurophysiol* 1999;16:14–39.