Published in final edited form as:

Epilepsia. 2012 November ; 53(11): e189–e192. doi:10.1111/j.1528-1167.2012.03653.x.

Automated Diagnosis of Epilepsy using EEG Power Spectrum

Wesley T. Kerr, B.A.^{1,2}, Ariana Anderson, PhD², Edward P. Lau, B.S.², Andrew Y. Cho, B.S.², Hongjing Xia, B.S.³, Jennifer Bramen, PhD², Pamela K. Douglas, Ph.D.², Eric S. Braun, B.A.⁴, John M. Stern, M.D.⁵, and Mark S. Cohen, Ph.D.^{2,3,5}

¹UCLA Department of Biomathematics, UCLA Medical Scientist Training Program

²UCLA Neuropsychiatric Institute

³UCLA Department of Biomedical Engineering

⁴UCLA Department of Urban Planning

⁵UCLA Department of Neurology

Abstract

Interictal electroencephalography (EEG) has clinically meaningful limitations in its sensitivity and specificity in the diagnosis of epilepsy because of its dependence on the occurrence of epileptiform discharges. We have developed a computer-aided diagnostic (CAD) tool that operates on the absolute spectral energy of the routine EEG and has both substantially higher sensitivity and negative predictive value than the identification of interictal epileptiform discharges. Our approach used a multilayer perceptron to classify 156 patients admitted for video-EEG monitoring. The patient population was diagnostically diverse with 87 diagnosed with either generalized or focal seizures. The remainder was diagnosed with non-epileptic seizures. The sensitivity was 92% (95% CI: 85–97%) and the negative predictive value was 82% (95% CI: 67%–92%). We discuss how these findings suggest that this CAD can be used to supplement event-based analysis by trained epileptologists.

Keywords

Epilepsy; machine learning; prediction; non-epileptic seizure; computer aided diagnostics

Introduction

Epilepsy is common and has a major impact on the global burden of disease. Though epilepsy is defined as an enduring predisposition for seizures, its diagnostic assessment relies on the clinical and/or electrographic description of transient events. Consequentially, the sensitivity of a single outpatient interictal electroencephalography (EEG) is only 50% (Brigo 2011, Gilbert 2003). If physicians do not observe the hallmark electrical features of interictal epileptiform discharges (IEDs), the assessment is inconclusive. This might help to explain why the average time to the diagnosis of non-epileptic seizures (NES) is 7.2 years (Bodde, et al. 2009). Automated seizure detection algorithms currently help physicians identify these transient events (Saab & Gotman 2005), but they do not detect the stable

Corresponding Author and Address, Wesley T Kerr, 760 Westwood Plaza, Suite 17-369, Los Angeles, CA 90095, wesleytk@ucla.edu, Phone: 310-986-3307, Fax: 310-794-7406.

Disclosure

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines. None of the authors has any conflict of interest to disclose.

pathology underlying each patient's chronic disease. A better understanding of the chronic state of epilepsy has great potential to impact patient care; automated computer methods have the potential to identify this stable abnormality and thereby to increase diagnostic accuracy, saving clinicians valuable time and improving patients' quality of care.

Seizure detection and prediction tools in epilepsy have been proposed frequently, yet efficient and effective computer aided diagnostic(CAD) tools have not yet been established. Only three publications address the question of epilepsy diagnosis using interictal scalp EEG alone (Bao 2009, Isik & Sezer 2010, Sezer, et al. 2010). All three publications report accuracies in excess of 90%. Other publications using the Freiburg dataset compare scalp EEG from normal controls to interictal intracranial EEG patient with epilepsy, which may limit clinical applicability (Tzallas, et al. 2009). Based on their success in the seizure and prediction literature, these tools used largely wavelet-based analysis and time frequency decompositions of short time windows of the signal(Ayala, et al. 2011, Jacobs, et al. 2011, Saab & Gotman 2005). However, longer time windows can capture the stable changes baseline dynamics attributable to epilepsy. The previous literature often compares the EEGs of patients with epilepsy to the EEGs from a healthy control population, a question that does not reflect the actual clinical situation. We consider comparing epilepsy to NES mimics the clinical scenario of a patient that needs to be assessed after experiencing a potential seizure event. As we show below, 30 percent of patients admitted for video-EEG monitoring have NES, including some who previously were diagnosed with intractable epilepsy. To develop directly clinically applicable tools, the diagnosis of each patient in the validation set must be certain therefore a careful discussion of the diagnostic assessment of each patient is critical. Similarly, epilepsy is a heterogeneous syndrome. Generally, the CAD literature either studies temporal lobe epilepsy or does not specify diagnostic subclass.

In this report, we outline the success of a novel CAD tool applied to a larger population of patients who have either focal or generalized epilepsies. By comparing to patients with NES and also inspecting time-frequency features of longer time windows of the EEG signal, we harness the stable interictal changes in the EEG that can be used to diagnose epilepsy. Further, we provide a detailed discussion of how such a tool can be used to supplement, not replace, manual analysis.

Patients and Methods

We studied the diagnostic test results from 514 patients admitted between 2008 and 2011 to the UCLA Seizure Disorder Center video-EEG monitoring unit. A subset of 156 patients was identified for further study because their diagnoses were definitive and they had not experienced previous penetrating head trauma. Within this subset, 87 were diagnosed with epilepsy and 69 were diagnosed with NES (full breakdown in Supplementary Information). Patients with NES and those with epilepsy underwent an identical evaluation. All methods were approved by the UCLA IRB and complied with the Helsinki Declaration.

All scalp EEG recordings were collected in accordance with standardized clinical procedures with a 200 Hz sampling rate using 26 electrodes placed according to the International 10–20 system. During acquisition, an analog 0.5 Hz high pass filter was applied to all recordings. Reviewed data consisted of between 1.5 and 25 hours (mean 9 hours, S.D. 4.5 hours) of archived EEG from either the first or second night of video-EEG monitoring. To assess the diagnostic yield of long term monitoring, we also inspected the records of all 514 reviewed patients admitted to UCLA for video-EEG monitoring.

The mean, standard deviation, minimum and maximum absolute spectral energy for non-overlapping 1 sec, 5sec, 60sec, 30 min windows of EEG recordings from all electrodes

relative to reference electrode 1, located between Fz and Cz, were calculated using the fast Fourier transform in MATLAB. The absolute value of spectral energy from 1–100 Hz was averaged over 1 Hz spectral bands. Short window lengths measure phenomena analogous to event related spectral perturbations (ERSPs) whereas longer windows capture baseline activity and connectivity. Each input feature corresponds to a separate electrode location, frequency band, statistical parameter and window length. The spectral energy from 58–62 Hz was excluded from all analysis to avoid AC line noise contamination. No other artifacts were removed. Ictal activity and muscle artifact were included in analysis.

Using a cyclical leave-one-out cross validation technique, a subset of the power spectrum was identified as potentially diagnostic by a highly-efficient minimum redundancy, maximum relevancy (mRMR) feature selection algorithm (Ding 2005, Peng 2005). This subset was then used as input for the Multilayer Perceptron neural network algorithm as implemented in Weka (Bouckaert 2010). For algorithmic details please refer to the supplementary material and Kerr *et al.* (Kerr, et al. 2012).

Results

The Multilayer Perceptron performance was comparable to manual event-based EEG analysis. Both manual and automated analyses were substantially and significantly better than a chance classifier based on clinical trial statistics (Figure 1A). The diagnostic accuracy of the CAD tool was 71% (64%–76%, $p < 10^{-4}$) significantly higher than chance: 56%. The risk ratio (the probability that a positive finding occurred in a patient with epilepsy compared to a patient with NES) was 3.68 (1.92–8.19, $p < 10^{-6}$). The odds ratio was 9.32 (3.51–25.73, $p < 10^{-5}$). In the study population, the results of a single outpatient non-video-EEG are not significantly different and have a relative risk ratio, odds ratio and accuracy of 2.52 (2.05–2.64, $p < 10^{-10}$), 99 (8.90–1100, $p < 10^{-3}$) and 72% (66–73%, $p < 10^{-4}$), respectively (Brigo 2011, Gilbert 2003). All intervals reflect 95% confidence bounds and all p values reflect comparisons to a naïve classifier.

In contrast with manual analysis, the performance of our CAD was driven by exceptionally high sensitivity (85%–97%, $p < 10^{-82}$) in comparison to only modest specificity (37%–51%, $p > 0.20$). Consequentially, the negative predictive value (67%–92%, $p < 10^{-24}$) is high compared to the positive predictive value (62%–71%, $p < 10^{-5}$). There was no significant difference in performance for focal and generalized epilepsies (see Suppl. Materials).

These results can be expressed in combination with the results of outpatient non-video EEGs as likelihood ratios, assuming the two tests are independent (Figure 1B) based on the

formula: $LR_{-M+CAD} = \frac{P(-_M|Ep) P(+_{CAD}|Ep)}{P(-_M|NES) P(+_{CAD}|NES)}$ We assume that 99% of neurologically normal patients have negative EEGs and that 50% and 90% of patients with epilepsy have abnormal outpatient EEGs after 1 and 4+ recordings, respectively (Brigo 2011, Gilbert 2003).

To illustrate the clinical problem further, we addressed the diagnostic yield of long term video-EEG monitoring specifically. As summarized in Figure 2, 9 percent of the 514 patients in our sample had inconclusive results upon the completion of monitoring (6%–12%). Six percent of patients admitted for pre-surgical assessment or intractable epilepsy were diagnosed with NES (2–10%).

Discussion

Inconclusive EEG results are a significant challenge to the effective treatment of epilepsy. For patients diagnosed with epilepsy, our finding that 6% of patients are later found to have

NES is concerning. Further, the most reliable diagnostic test, conventional long term video-EEG monitoring, is inconclusive for roughly 9% of epilepsy patients due to lack of relevant electrophysiological events. To reduce this rate, admission duration must increase. Our technology, however, avoids this problem altogether by focusing on baseline diagnostic features. Successful validation and then implementation of our CAD tool could therefore provide additional information to that could, in time, substantially reduce both of these values. Validation would require a prospective assessment of patients who are later admitted for video-EEG monitoring or retrospective analysis of records from other institution(s).

We hypothesize that our results capitalize both on low frequency trends used by previous literature and, potentially, also on high frequency oscillations up to 100 Hz. Most ictal activity is within the 3–25 Hz range (Saab & Gotman 2005). Seizure detection algorithms have achieved impressive results operating on frequency bands less than 40 Hz using much more complex machine learning methods (Tzallas, et al. 2009). However, recent evidence in intracranial EEG suggests that patients with epilepsy have increased high frequency oscillations in the 40+ Hz range (Ayala, et al. 2011, Jacobs, et al. 2011). Due to the nature of our algorithm, the contribution of each window length, spectral band and electrode location is unclear.

Our entirely automated tool diagnosed patients with performance similar to epileptologists manually reading outpatient EEGs. Our performance was quantitatively less than previous methods. However, ours was designed and tested in the real-world context of an inpatient unit, with its heterogeneous mixture of medications, ages and patient histories. The statistics reveal that our approach has a high negative predictive value whereas manual analysis has, instead, a high positive predictive value. These improvements are based on information not observable without CAD and are independent of reader expertise, suggesting that our methods can be used in combination with manual analysis to improve the diagnostic yield of EEG. This synergistic combination could more efficiently and quickly identify those patients who may require further diagnostic or pre-surgical assessment. Given the broad and growing evidence that early epilepsy surgery — when supported by accurate diagnostics — may be more effective than treatment with AEDs alone (Engel, et al. 2012), we believe that this application offers the potential to meaningfully impact the care of patients with epilepsy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The author would like to give special thanks to Matthew Spector and Kirk Shattuck for data management of the clinical EEG records. This work was supported by the UCLA-Caltech Medical Scientist Training Program (NIH T32 GM08042), the Systems and Integrative Biology Training Program at UCLA (NIH T32 GM008185), NIH R33 DA026109 (to M.S.C.) and the UCLA Department of Biomathematics.

References

- Ayala M, Cabrerizo M, Jayakar P, Adjouadi M. Subdural EEG classification into seizure and nonseizure files using neural networks in the gamma frequency band. *J Clin Neurophysiol*. 2011; 28:20–29. [PubMed: 21221013]
- Bao, FS.; Gao, J-M.; Lie, DYC.; Zhang, Y.; Oommen, KJ. 31st Annual International Conference of the IEEE EMBS. Minnesota: Minneapolis; 2009. Automated Epilepsy Diagnosis Using Interictal Scalp EEG; p. 6603-6607.

- Bodde NM, Brooks JL, Baker GA, Boon PA, Hendriksen JG, Aldenkamp AP. Psychogenic non-epileptic seizures--diagnostic issues: a critical review. *Clin Neurol Neurosurg.* 2009; 111:1–9. [PubMed: 19019531]
- Bouckaert RR, Frank E, Hall MA, Holmes G, Pfahringer B, Ruetemann P, Witten IH. Weka-experiences with a java open-source project. *J Mach Learn Res.* 2010; 11:2533–2541.
- Brigo F. An evidence-based approach to proper diagnostic use of the electroencephalogram for suspected seizures. *Epilepsy Behav.* 2011; 21:219–222. [PubMed: 21624850]
- Ding C, Peng HC. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol.* 2005; 3:185–205. [PubMed: 15852500]
- Engel J Jr, McDermott MP, Wiebe S, Langfitt JT, Stern JM, Dewar S, Sperling MR, Gardiner I, Erba G, Fried I, Jacobs M, Vinters HV, Mintzer S, Kieburtz K. Early surgical therapy for drug-resistant temporal lobe epilepsy: a randomized trial. *JAMA.* 2012; 307:922–930. [PubMed: 22396514]
- Gilbert DL, Sethuraman G, Kotagal U, Buncher R. Meta-analysis of EEG test performance shows wide variation among studies. *Neurology.* 2003; 60:564–570. [PubMed: 12601093]
- Isik H, Sezer E. Diagnosis of epilepsy from electroencephalography signals using multilayer perceptron and elman artificial neural networks and wavelet transform. *J Med Syst.* 2010; 36:1–13. [PubMed: 20703754]
- Jacobs J, Kobayashi K, Gotman J. High-frequency changes during interictal spikes detected by time-frequency analysis. *Clin Neurophysiol.* 2011; 122:32–42. [PubMed: 20599418]
- Kerr, WT.; Anderson, A.; Xia, H.; Braun, ES.; Lau, EP.; Cho, AY.; Cohen, MS. 2nd International Workshop Pattern Recognition in Neuroimaging. London: Conference Publishing Services; 2012. Parameter Selection in Mutual Information-Based Feature Selection in Automated Diagnosis of Multiple Epilepsies Using Scalp EEG. [accepted].
- Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Analysis and Machine Intelligence.* 2005; 27:1226–1238.
- Saab ME, Gotman J. A system to detect the onset of epileptic seizures in scalp EEG. *Clin Neurophysiol.* 2005; 116:427–442. [PubMed: 15661120]
- Sezer E, Isik H, Saracoglu E. Employment and comparison of different artificial neural networks for epilepsy diagnosis from EEG signals. *J Med Syst.* 2010; 36:347–362. [PubMed: 20703714]
- Tzallas AT, Tsipouras MG, Fotiadis DI. Epileptic seizure detection in EEGs using time-frequency analysis. *IEEE Transactions on Information Technology in Biomedicine.* 2009; 13:703–710. [PubMed: 19304486]

Figure 1.

Panel A directly compares the summary statistics of our computer aided diagnostic tool (CAD) compared to the same statistics regarding conventional analysis of EEG. Panel B assesses the likelihood ratios that can be achieved when our CAD is combined with conventional analysis. Error bars denote 95% CIs and are calculated without normal assumptions. Dashed lines indicate chance or 95% CIs of chance. All effects are significantly different from chance ($p < 0.001$) except when CAD is positive and manual analysis is negative. No comparative effects are significantly different.

Figure 2.

Diagnostic yield of long term video-EEG monitoring. Numbers indicate how many patients are in each class and the size of the bar denotes percent of total, listed right, that belong to each class. When the presurgical and intractable classes are combined, six percent of the patients have inconclusive results. NES: Non-epileptic seizures.