HHS Public Access

Author manuscript

J Clin Neurophysiol. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:

J Clin Neurophysiol. 2018 September ; 35(5): 375–380. doi:10.1097/WNP.0000000000000492.

Interictal Epileptiform Discharge Detection in EEG in Different Practice Settings

Jonathan J. Halford, MD[1], **M. Brandon Westover, MD PhD**[2], **Suzette M. LaRoche, MD**[3], **Micheal P. Macken, MD**[4], **Ekrem Kutluay, MD**[1], **Jonathan C. Edwards, MD**[1], **Leonardo Bonilha, MD PhD**[1], **Giridhar P. Kalamangalam, MD PhD**[5], **Kan Ding, MD**[6], **Jennifer L. Hopp, MD**[7], **Amir Arain, MD**[8], **Rachael A. Dawson**[1], **Gabriel U. Martz, MD**[9], **Bethany J. Wolf, PhD**[10], **Chad G. Waters, MS**[11], and **Brian C. Dean, PhD**[11]

[1]Department of Neurology, Medical University of South Carolina, Charleston SC

[2]Department of Neurology, Massachusetts General Hospital, Boston MA

[3]Mission Health, Asheville NC

[4]Department of Neurology, Northwestern University, Chicago IL

[5]Department of Neurology, University of Florida, Gainesville FL

[6]Department of Neurology, University of Texas Southwestern, Dallas TX

[7]Department of Neurology, University of Maryland School of Medicine, Baltimore MD

[8]Amir Arain MD, Department of Neurology, Vanderbilt University, Nashville TN

[9]Norton Neurology Services, Louisville, KY

[10]Department of Public Health Sciences, Medical University of South Carolina, Charleston SC

[11]School of Computing, Clemson University, Clemson SC

Abstract

Objective—The goal of the project is to measure the performance of academic and private practice (PP) neurologists in detecting interictal epileptiform discharges (IEDs) in rEEG recordings.

Methods—35 EEG scorers (EEGers) participated (19 academic and 16 PP), and marked the location of ETs in 200 30-second EEG segments using a web-based EEG annotation system. All participants provided board certification status, years of (EFT), and years in practice. The Persyst P13 automated IED detection algorithm was also run on the EEG segments for comparison.

Corresponding Author: Jonathan J. Halford, Address: 96 Jonathan Lucas St.; Suite 307 CSB; MSC606; Charleston SC 29425, Tel: 843-792-3223; Fax: Fax 843-792-8626; halfordj@musc.edu.

This work has been presented as a poster at the American Epilepsy Society Annual Meeting in December 2016.

Supplementary Digital Content (SDC)

Text explanations of our clustering approach, text explanation for our Brier score and MPD statistical approach, Table 1e, and Figures 1e, 2e, and 3e are provided in the SDC.

Results—Academic EEGers had an average of 1.66 years of EFT versus 0.50 years of EFT for PP EEGers ($p < 0.0001$) and had higher rates of board certification. Inter-rater agreement (IRA) for the 35 EEGers was fair. There was higher performance for EEGers in academics, with at least 1.5 years of EFT, and with ABCN and ABPN-E specialty board certification. The Persyst P13 algorithm at its default setting (perception value = 0.4) did not perform as well at the EEGers, but at substantially higher perception value settings, the algorithm performed almost as well human experts.

Conclusion—IRA among EEGers in both academic and PP settings varies considerably. Practice location, years of EFT, and board certification are associated with significantly higher performance for IED detection in rEEG. Continued medical education of PP neurologists and neurologists without EFT is needed to improve rEEG interpretation skills. The performance of automated detection algorithms is approaching that of human experts.

Keywords

Electroencephalography; EEG; Inter-rater agreement; Epilepsy

The routine scalp electroencephalogram (rEEG) is the most common clinical neurophysiology procedure. Approximately 650,000 outpatient rEEGs were performed in the United States in 2013 with a cost of around $480 million.[1] The EEG of a patient with epilepsy is characterized by the presence of interictal epileptiform discharges (IEDs) consisting of spike or sharp wave discharges. Detecting IEDs in EEG is important because their presence is predictive of seizure recurrence in patients following a first seizure[2, 3] and is useful in supporting the diagnosis of epilepsy.[4] However, due to the wide variety of morphologies of IEDs and their similarity to waves that are part of the normal background activity and to artifacts, the detection of IEDs is far from straightforward.[5]

Accurate EEG interpretation is important because misinterpretation of the rEEG can adversely affect patients, leading to the misdiagnosis of epilepsy, inappropriate use of antiepileptic medications, and delay in appropriate treatment of the true underlying cause of seizure-like events (which can include non-epileptic events and cardiac arrythmias[6]). Fellowship training has long been considered important in developing the skills necessary for accurate EEG interpretation, but data measuring the ability of neurologists to interpret EEG has been lacking.[6-8] We present results from the first study to measure IED detection performance in a group of academic clinical neurophysiologists and private practice (PP) neurologists who have varying levels of training and board certification status.

Automated systems to help neurologists detect IEDs have been under development for many years.[9] Although most commercial automated IED detection systems have not been adequately tested by independent investigators, the general opinion among clinicians has been that their performance is suboptimal, mostly due to unacceptably low specificity. Recently, the performance of a new automated IED detector developed and marketed by Persyst Development Corporation was reported to perform as well as of three experienced EEG technologists.[10] In this report we also compare the performance of this new Persyst P13 detector to a large group of neurologists.

1. Methods

EEG Dataset

A library of rEEG segments was created by retrospective review of approximately 1000 rEEGs performed in the Medical University of South Carolina (MUSC) Neurophysiology Laboratory for clinical purposes during the years 2010-2011. Two hundred 30-second rEEG segments from 200 different rEEG studies performed on 200 different patients were selected by two experts (JJH and GUM). Fifty of the segments were chosen from within 50 EEGs which were randomly selected from EEGs performed at MUSC within the previous two years which were interpreted as containing IEDs in the clinical report. Fifty of the segments were from randomly selected EEGs read as normal in the clinical report. Fifty segments were selected because they contained IEDs from patients with known epilepsy with difficult-to-interpret IEDs. Finally, fifty were selected because they contained benign paroxysmal activity (e.g. exaggerated alpha activity, wicket spikes, and small sharp spikes) which could easily be misinterpreted by an inexperienced reviewer as being epileptiform. The two investigators (JJH and GUM) who selected the EEG data for this study were not included as EEG scorers (EEGers).

The rEEG studies were recorded referentially using digital Natus/XLTEK EEG recording equipment with a sampling frequency of 256 Hz with the standard 10-20 electrode placement. Institutional Review Board approval was granted from the Medical University of South Carolina (MUSC) and Clemson University. The rEEG data was exported from the MUSC clinical XLTEK server as European Data Format (EDF)[11] files. All segments were randomly ordered and concatenated into one long 100 minute recording.

EEGnet Annotation System

EEGnet is a software system that allows users to view and annotate EEG data using a web browser with no local software installation required.[12] EEGnet software leverages recent improvements in web standards (e.g., HTML5), in order to run smoothly in most modern web browsers and was designed to create a user interface very similar to commercially available software used in clinical EEG interpretation. The web-based “front end” of the system, written in JavaScript, interacts with a secure backend database (MySQL running on CentOS Linux) hosted at the Clemson University School of Computing. The ECG channel is also displayed. Channel gain for all channels can be adjusted together or individually, and the user can scroll forward and backward through the EEG data one second or ten seconds at a time. The user can adjust digital high pass, low pass, and notch filter settings. Scoring is performed using any one of a list of conventional EEG montages including anterior-posterior bipolar, transverse bipolar, hatband bipolar, average reference, Cz reference, and ipsilateral ear reference. For this study, display time window was fixed at 10 s. See Figure 1e in Supplemental Digital Content (SDC).

EEG Scorer Participants

Invitations to take part in the study were sent by email to the 36 members of the Critical Care EEG Monitoring Research Consortium (CCEMRC). CCEMRC participants were paid $200 for participation. The offices of all neurologists in the states of SC, NC, GA and parts

of FL were contacted and a list of 192 neurologists who interpreted rEEGs (per their office staff) was compiled. A letter of invitation was sent to these neurologists providing a website to login to the EEGnet system and offering reimbursement of $400 for participation in the study. PP neurologists who participated choose their own 5-digit numerical identifier when they logged into the EEGnet system so that their performance would not be linked to their name or any other identifying information. Information was collected from each EEGer: board certification, years of EFT (years of fellowship training dedicated specifically to EEG training), and years in practice. EEGers were asked if they possessed board certification by the American Board of Clinical Neurophysiology (ABCN), the American Board of Psychiatry and Neurology (ABPN) Neurophysiology Subspeciality (ABPN-NP), and the ABPN Epilepsy Subspecialty (ABPN-E). It was not asked whether the physician was trained as an adult or pediatric neurologist out of concern that the answer would be considered identifying.

EEG Scoring Paradigm

Participating EEGers were instructed to label all IEDs in the recording on the channel on which they thought it was best represented. An algorithm used in previous studies[12, 13] clustered the EEGer annotations producing a number of “events clusters”. (See text and Figure 2e in SDC for further explanation.)

Persyst Automated Detection

Each of the 200 EEG EDF files were loaded into the Persyst Development Corporation software system and the P13 detector was run on each 30 sec EEG file individually. Automated spike detections were outputted as a SD4 file which contained the timepoints for each detection and a Persyst ‘perception value‘ for each detection. Data in these SD4 files were imported into MATLAB for analysis. A Persyst detection was considered to occur at the same time as a cluster of EEGer annotations if the Persyst detection occurred after the start time of the first temporally occurring EEGer event in a cluster or before the end time of the last temporally occuring EEGer event in a cluster.

Statistical Analysis

Data analysis was performed in MATLAB (MATLAB 8.6 with the Statistical Analysis Toolbox) with the exception that EEGer events were clustered within the EEGnet system. Statistical analysis was performed within MATLAB, using the Statistical Analysis Toolbox. Inter-rater agreement (IRA) was measured using linearly weighted Gwet’s agreement coefficient (AC2)[14].

We employed two methods for evaluating EEGer performance, the Brier score and the mean pair-wise distance (MPD). The Brier score (also called the mean squared error), is a well-known measure of calibration for the accuracy of probabilistic classification decisions.[15, 16] A lower Brier score indicates better performance. The MPD was derived by using the mean pair-wise sensitivity and mean pair-wise false positive rates. See the SDC for further explanation of our analysis using the Brier score and MPD.

All continuous variables were found to be non-normally distributed and simple statistical comparisons between groups of continuous variables were performed using the non-parametric Wilcoxon rank sum test. 95% confidence intervals (CI) are reported. Comparisons were made between the Brier scores of groups of EEGers based on board certification status for three board exams (ABCN certification, ABPN-NP certification, ABPN-E certification), years of EFT, and years in practice. We designate differences associated with p-values <0.05 as statistically significant. No correction was made for multiple comparisons because most EEGer characteristics were not independent, since academic and PP neurologists had very different rates of board certification and years of EFT (see Table 1).

3. Results

A total of 35 EEGers participated in the project, 19 from academic and 16 from PP settings. The response rate for invitations to participate was 53% for academic neurologists and 8.3% for PP neurologists. Academic EEGers had an average of 1.66 years of EFT (95% CI 1.43 – 1.89) versus 0.50 years of EFT for PP EEGers (95% CI 0.18 – 0.82; $p < 0.00002$). There was no statistically significant difference in number of years in practice between the two groups. See Table 1 and Table 1e (in SDC) for further information about participant EEGer and invited neurologist characteristics. Although a complete survey was not made of all neurologists in the regions we studied, we contacted most neurology practices in the states of SC, NC, GA and parts of FL, and office staff reported that more than 50% of PP neurologists in these offices regularly interpreted rEEGs.

Participants used the EEGnet system to label a total of 2368 single-channel EEG segments as IEDs. The median number of IEDs labeled by each EEGer was 54 (mean 67.7, range 6 - 212, standard deviation 52.3). Each EEGer spent approximately 1-2 hours completing the task. The annotations were clustered by temporal occurrence, which produced a list of 573 event clusters containing an annotation from at least one EEGer. Figure 3e (in SDC) displays the morphology of the many of the events which were annotated by the academic neurologists[17]. Most event clusters contain only a few annotations and there are a large number of clusters with only one or two members (Figure 1A). A small group of EEGers produced most of the annotation clusters with few members: 66% of the one-annotation clusters and 59% of the two-annotation clusters were from just six EEGers. Many of the EEGers did not contribute many annotations to these event clusters with few members: 15 EEGers contributed only between 0 – 2 annotations to the 364 one or two annotation clusters.

Overall, the AC2 IRA for all 35 EEGers across all 573 events was good at 0.81, although the IRA drops rapidly if it is calculated using only event clusters which were labeled by more than one EEGer. Subgroups of EEGers have higher AC2 IRA (Figure 1B). Performance based on Brier score was calculated for each EEGer and it correlated highly (rho = 0.77) with the MPD score. Because performance results using the Brier score and the MPD score were so similar, statistical comparisons were performed with only the Brier score.

Significant differences in performance were evident between academic and PP EEGers across multiple EEGer characteristics. Academic EEGers had a median Brier score performance of 0.0082 versus 0.022 for PP EEGers ($p < 0.04$); EEGers with ABCN board certification had a median Brier score performance of 0.0074 versus 0.012 for those without ($p < 0.04$); and with ABPN-E certification had a median Brier score of 0.0064 versus 0.012 for those without ($p < 0.02$). EEGers with at least 1.5 years of EFT also showed a trend toward better performance (0.0078 versus 0.012 for EEGers with less than 1.5 years training, $p < 0.08$) and years of EFT correlated to a small and non-significant degree with Brier score ($\rho = -0.28$, $p < 0.11$). Further Brier score results, including confidence intervals, are given in Table 2. Mean Brier score values are plotted in Figure 2. Lower Brier score indicates better performance. There was no significant difference in Brier score between EEGers with and without ABPN-CN certification and Brier score was uncorrelated with years of practice.

Mean pairwise sensitivity and mean pairwise false positives per minute for all EEGers and for the Persyst P13 detection algorithm are plotted in Figure 3. The performance of the Persyst algorithm is plotted using a range of 'perception value' thresholds. According to Persyst Development Corporation, the "out of the box" default perception value threshold setting for the P13 detector is 0.4. At this setting, the P13 detector has a lower specificity (higher mean pairwise false positive rate) than all of the EEGers. At higher perception value settings, the P13 detector performs considerably better, but overall the P13 algorithm generates more false positives at a given sensitivity than the EEGers.

4. Discussion

There are multiple anecdotal reports to suggest that EEGs are sometimes misinterpreted.[5, 6, 18, 19] This study provides the first objective evidence that there is considerable variability in the performance of neurologists for IED identification in rEEG recordings. Academic neurologists and neurologists with ABCN and ABPN-E certification perform significantly better at IED identification than PP neurologists and those without these board certifications. There was also a trend towards improved performance of EEGers with 1.5 years of EFT. The effect of EEG misinterpretation on patient care is probably considerable, leading to misdiagnosis of epilepsy and initiation of treatment with antiepileptic medications in patients who do not need these medications. Further research is needed to confirm these findings using larger studies.

As reported in previous studies, the IRA among all EEGers was low to fair and considerably better among EEGers with the best IRA.[9, 12, 20] The finding of only moderate IRA among EEGers is probably at least partially due to half of the EEG segments having been chosen to be difficult-to-interpret IEDs and difficult-to-interpret benign EEG transients. As shown in Figure 3, certain EEGers produced considerably more false positive detections, in comparison to the group of all EEGers. This lower specificity is concerning since the mislabeling of benign EEG events at IEDs has potentially greater adverse effects on patient care than failing to label an EEG event as an IED.[5]

In a recent paper by Scheuer et al.[10], it is reported that the Persyst P13 algorithm, set at a perception threshold setting of 0.1, performs as well as three experienced EEG technologists

in labeling the location of IEDs in 40 prolonged EEG recordings. In our assessment, at a setting of 0.1, the P13 algorithm has a good sensitivity performance, based on mean pairwise sensitivity in comparison to most of the EEGers. But at this setting it, has an unacceptably high mean pairwise false positive rate. At higher perception value settings, the P13 algorithm produces less false positive detections, but at the default setting of 0.4, it still produces more false positive detections than any EEGer. If the perception value setting is increased to the 0.95 – 0.99 range, the false positive rate drops down into the range of many EEGers, but it is unclear whether a user would ever know to increase the perception value threshold setting or how the user would know that the perception value may need to be set this high. Based on this study, it appears that further work needs to be done to find the optimal perception value threshold setting for the P13 detector.

There are several weaknesses in this study. First, our IRA method for measuring IED detection accuracy, while the only method we have, has limitations since there is no gold-standard measurement and training probably increases homogeneity of EEGer opinion. Second, EEGers were asked to annotate short 30 second EEG segments, which could have lowered IRA since it is known that agreement between experts for interpreting rEEGs is higher for 20 minute rEEGs than for shorter EEG recordings.[21] Third, some of the EEG segments used for this study were selected by only two investigators, which could have led to selection bias. Steps were taken to mitigate this by randomly selecting some rEEG segments and by not including the investigators who selected the segments as EEG scorers. Fourth, despite soliciting participation from nearly two hundred PP neurologists, the participation rate of PP neurologists was low with only 16 PP neurologists participating. As a result, we may not have collected data from a representative sample of PP neurologists. It should be noted that PP participants had a significantly higher rate of neurophysiology board certification than those invited, so the performance we measured possibly overestimates the performance of PP neurologists. Further investigation is needed to better define the inter-rater agreement of neurologists who interpret rEEG.

The results of this study raise the question of whether rEEGs should be read only by or always “over-read” by neurologists with EFT. The Joint Commission on Accreditation of HealthCare Organizations (JCAHO) mandates that hospitals under their jurisdiction provide a mechanism for quality assurance interpretation of electrocardiograms (ECGs) by qualified physicians with interpretation privileges.[22] For this reason, ECGs performed at JCAHO-accredited hospitals are often read or over-read by cardiologists, who receive extensive education in ECG interpretation during their fellowship training. The question of whether EFT should be required of neurologists who interpret rEEGs was recently put forward by Benbadis.[18] Since, by our assessment, a large fraction of PP neurologists interpret rEEGs, this is an important clinical question. But because so many rEEGs are performed in the PP setting and so few PP neurologists have EFT, it is not possible for all rEEGs to be interpreted only by neurologists with EFT. It is also unreasonable to expect PP neurologists who are already well-established in practice to get EFT. Therefore, the only practical options to address the problem of rEEG misinterpretation are to improve the skills of neurologists interpreting rEEGs and to improve automated detection systems to provide assistance. Continuing educational modules could be created using web-based EEG rendering systems. Using its default setting, the Persyst P13 detector assessed in this study shows inferior

performance to most participating neurologists. Hopefully, over time the performance of the Persyst automated IED detector and other automated IED detectors will continue to improve. With progress in education and the continued improvement of automated detection systems, rEEG interpretation will hopefully become more uniform, which will improve patient care.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The project is unfunded. The authors would like to acknowledge of support of the Critical Care EEG Monitoring Research Consortium. Statistical analysis for this project was supported by the South Carolina Clinical & Translational Research Institute, Medical University of South Carolina's CTSA, NIH/NCRR Grant Number UL1RR029882. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH or NCRR.

Study funding: NIH/NCRR Grant Number UL1RR029882

References

1. Simpson KN. Unpublished data from Medicare 5% sample and Truven MarketScan 2013 data set. 2016
2. Van Donselaar CA, Schimsheimer R, Geerts AT, Declerck AC. Value of the Electroencephalogram in Adult Patients With Untreated Idiopathic First Seizure. Archives of Neurology. 1992; 49:231–237. [PubMed: 1536624]
3. Seidel S, Pablik E, Aull-Watschinger S, Seidl B, Pataraia E. Incidental epileptiform discharges in patients of a tertiary centre. Clinical Neurophysiology. 2016; 127:102–107. [PubMed: 25802204]
4. Fountain NB, Freeman JM. EEG is an essential clinical tool: pro and con. Epilepsia. 2006; 47:23–25. [PubMed: 17044821]
5. Tatum WO. How not to read an EEG: introductory statements. Neurology. 2013; 80:S1–3.
6. Benbadis SR. Errors in EEGs and the misdiagnosis of epilepsy: importance, causes, consequences, and proposed remedies. Epilepsy & Behavior. 2007 Nov.11:257–262. [PubMed: 17719853]
7. Benbadis SR, Tatum WO. Overinterpretation of EEGs and Misdiagnosis of Epilepsy. Journal of Clinical Neurophysiology. 2003; 20:42–44. [PubMed: 12684557]
8. Benbadis SR. The EEG in nonepileptic seizures. Journal of Clinical Neurophysiology. 2006 Aug. 23:340–52. [PubMed: 16885708]
9. Halford JJ. Computerized Epileptiform Transient Detection in Scalp EEG: Obstacles to Progress and the Example of Computerized ECG Interpretation. Clinical Neurophysiology. 2009; 120:1909–1915. [PubMed: 19836303]
10. Scheuer ML, Bagic A, Wilson SB. Spike detection: Inter-reader agreement and a statistical Turing test on a large data set. Clinical Neurophysiology. 2017; 128:243–250. [PubMed: 27913148]
11. Kemp B, Värri A, Rosa AC, Nielsen KD, Gade J. A simple format for exchange of digitized polygraphic recordings. Electroencephalography and Clinical Neurophysiology. 1992; 82:391–393. [PubMed: 1374708]
12. Halford JJ, Schalkoff RJ, Zhou J, et al. Standardized Database Development for EEG Epileptiform Transient Detection: EEGnet Scoring System and Machine Learning Analysis. Journal of Neurosciences Methods. 2013; 212:308–316.
13. Halford JJ, Arain A, Kalamangalam GP, et al. Characteristics of EEG Interpreters Associated With Higher Interrater Agreement. Journal of Clinical Neurophysiology. 2017; 34:168–173. [PubMed: 27662336]

14. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol. 2008; 61:29–48. [PubMed: 18482474]
15. Brier GW. Verification of forecasts expressed in terms of probabilities. Monthly Weather Review. 1950; 78:1–3.
16. Ferri C, Hernández-Orallo J, Modroiu R. An experimental comparison of performance measures for classification. Pattern Recognition Letters. 2009; 30:27–38.
17. Bagheri E, Dauwels J, Dean BC, Waters CG, Westover MB, Halford JJ. Interictal epileptiform discharge characteristics underlying expert interrater agreement. Clinical Neurophysiology. 2017; 128:1994–2005. [PubMed: 28837905]
18. Benbadis SR. "Just like EKGs!" Should EEGs undergo a confirmatory interpretation by a clinical neurophysiologist? Neurology. 2013; 80:S47–S51. [PubMed: 23267045]
19. Miller JW, Henry JC. Solving the dilemma of EEG misinterpretation. Neurology. 2013; 80:13–14. [PubMed: 23267028]
20. Halford JJ, Pressly WB, Benbadis SR, et al. Web-based Collection of Expert Opinion on Routine Scalp EEG: Software Development and Inter-rater Reliability. Journal of Clinical Neurophysiology. 2011; 28:174–178.
21. Black MA, Jones RD, Carroll GJ, Dingle AA, Donaldson IM, Parkin PJ. Real-time Detection of Epileptiform Activity in the EEG: A Blinded Clinical Trial. Clinical Electroencephalography. 2000; 31:122–130. [PubMed: 10923198]
22. Medical electrical equipment – part 12-51: particular requirements for safety, including essential performance, of recording and analyzing single channel and multichannel eletrocardiographs. Geneva: International Electrotechnical Commission; 2003. IEC 60601-2-51

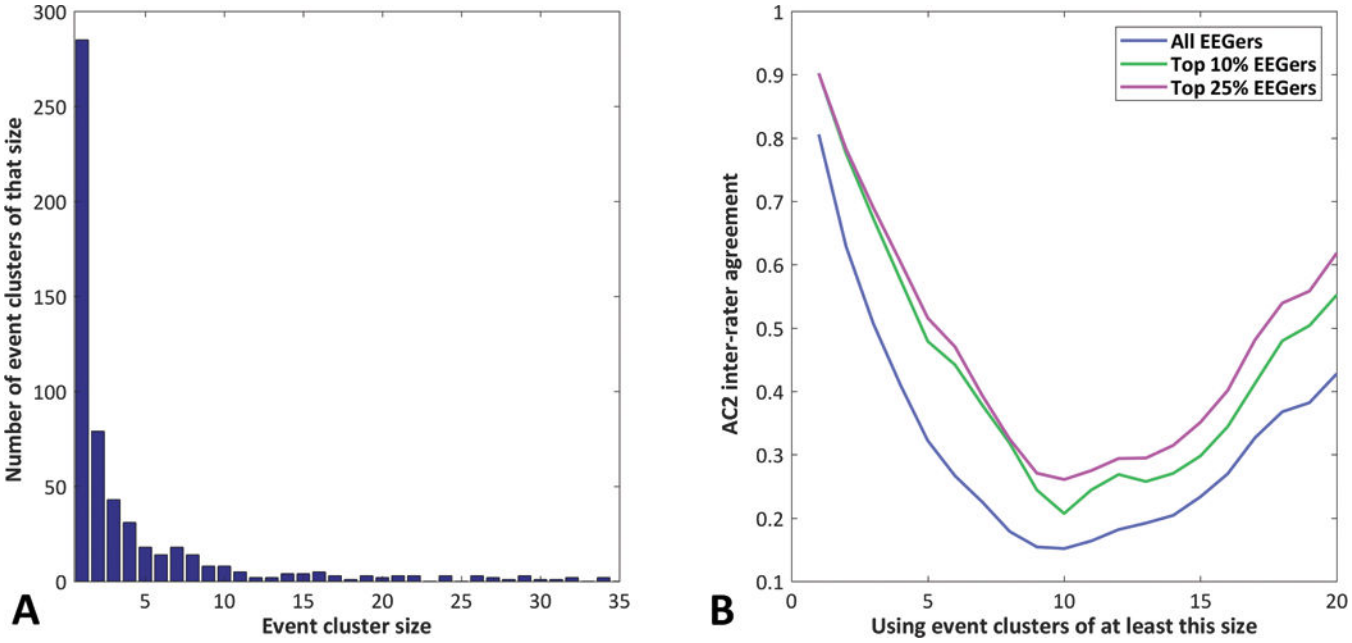

Figure 1. Size of Event Clusters and Inter-rater Agreement

Subplot A shows the size of each event cluster (the number of EEGers that annotated an event in that cluster) versus the number of event clusters of that size. Subplot B shows the AC2 inter-rater agreement calculated using events which were labeled by a minimum number of EEGers.

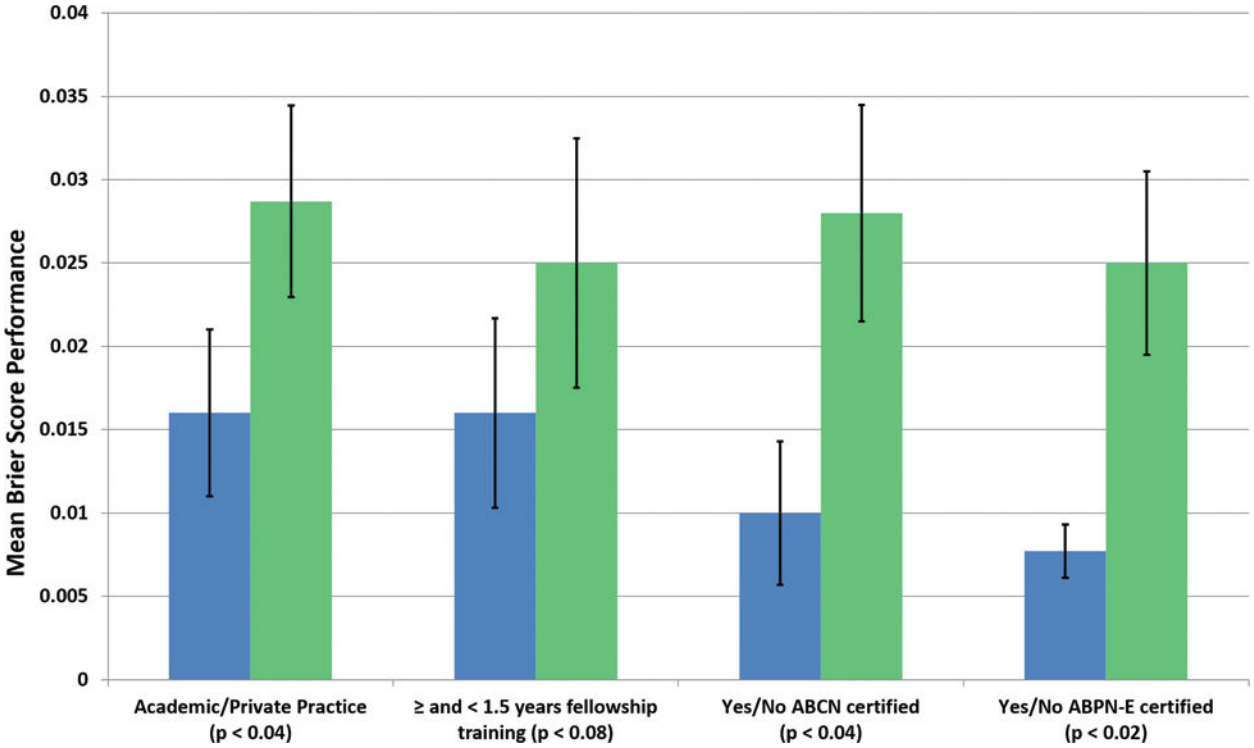

Figure 2. Mean Brier Score Performance for Labeling IEDs

Mean Brier score performance with standard error bars for labeling IEDs based on practice setting, years of neurophysiology fellowship training, ABCN board certification, and ABPN-E board certification. Blue color represents performance for academics, ≥ 1.5 years EFT, ABCN board certified and ABPN-E certified.

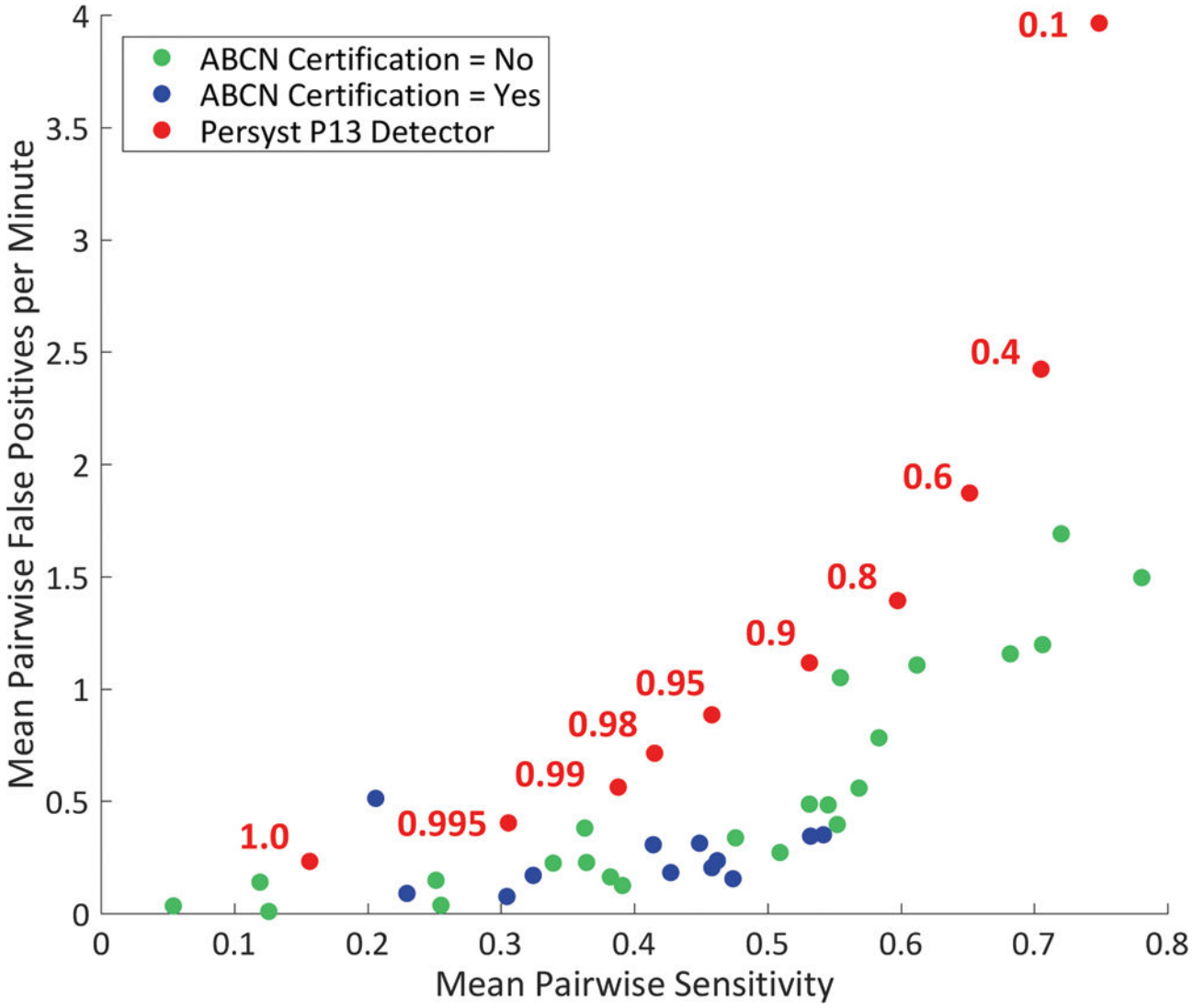

Figure 3. Mean Pair-wise Sensitivity and False Positive Rate for Labeling IEDs

Mean pair-wise sensitivity and false positive rate for labeling the 200 EEG segments for all EEGers along with Persyst P13 detector performance at 10 different perception value thresholds. ABCN board certification status of EEGers is indicated.

Table 1

Characteristics of academic participants, PP participants and invited PP neurologists

	Academic participants	PP participants	All invited PP neurologists
Number	19	16	192
ABCN certified %	57.9 *	6.3 *	3.6 *
ABPN-CN certified %	57.9 **	31.3	14.1 **
ABPN-E certified %	36.9 *	0 *	2.1 *
At least one neurophysiology board certification %	94.7 *	31.3 *	16.7 *
ABPN-Neurology certified %	100 **	**81.2** **	82.8 **
Average years of fellowship training	1.66 ***	0.50 ***	unknown
Average years in practice	10.2	14.6	unknown
Median years in practice	9.0	13.5	unknown

* Significant difference between all three groups, $p < 0.001$

** Significant difference between academic participants and all invited PP neurologists, $p < 0.0001$

*** Significant difference between academic participants and PP participants, $p < 0.0001$

Table 2

Median Brier score performance based on EEGer characteristics including board certification from the American Board of Clinical Neurophysiology (ABCN), The American Board of Psychiatry and Neurology (ABPN) Clinical Neurophysiology (CN) subspecialty and Epilepsy (E) subspecialty, and years of epilepsy fellowship training (EFT).

EEGer Characteristic		Median	95% Confidence Intervals
Practice Location	Academic	0.0081	0.0039 - 0.028
	Private Practice	0.022	0.014 - 0.044
ABCN Certification	Yes	0.0075	0.0051 - 0.015
	No	0.012	0.014 - 0.041
ABPN – CN Certification	Yes	0.011	0.0075- 0.026
	No	0.0105	0.010- 0.042
ABPN – E Certification	Yes	0.0064	0.0036 - 0.011
	No	0.012	0.014 - 0.037
Years of EFT	EFT ≥ 1.5	0.0078	0.000097 - 0.032
	EFT < 1.5	0.012	0.014 - 0.037
Years of EFT	EFT ≥ 1.0	0.0097	0.010- 0.031
	EFT < 1.0	0.011	0.0025 - 0.047