Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Clin Neurophysiol. Author manuscript; available in PMC 2021 September 01.

Published in final edited form as:

Clin Neurophysiol. 2020 September ; 131(9): 2298–2306. doi:10.1016/j.clinph.2020.05.032.

A standardized nomenclature for spectrogram EEG patterns: Inter-rater agreement and correspondence with common intensive care unit EEG patterns

Sahar F. Zafar[a,*], **Edilberto Amorim**[a,b], **Craig A. Williamsom**[c], **Jin Jing**[a], **Emily J. Gilmore**[d], **Hiba A. Haider**[e], **Christa Swisher**[f], **Aaron Struck**[g], **Eric S. Rosenthal**[a], **Marcus Ng**[h], **Sarah Schmitt**[i], **Jong W. Lee**[j], **M. Brandon Westover**[a]

[a]Massachusetts General Hospital, Department of Neurology, Boston, MA, USA

[b]University of California, Department of Neurology, San Francisco, CA, USA

[c]University of Michigan, Department of Neurosurgery and Neurology, Ann Arbor, MI, USA

[d]Yale School of Medicine, Department of Neurology, New Haven, CT, USA

[e]Emory University School of Medicine, Department of Neurology, Atlanta, GA, USA

[f]Duke University School of Medicine, Department of Neurology, Durham, NC, USA

[g]University of Wisconsin, Department of Neurology, Madison, WI, USA

[h]University of Manitoba, Winnipeg, Canada, USA

[i]University of South Carolina, Department of Neurology, Charleston, SC, USA

[j]Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA

Abstract

Objective: To determine the inter-rater agreement (IRA) of a standardized nomenclature for EEG spectrogram patterns, and to estimate the probability distribution of ictal-interictal continuum (IIC) patterns vs. other EEG patterns within each category in this nomenclature.

Methods: We defined seven spectrogram categories: "Solid Flames", "Irregular Flames", "Broadband-monotonous", "Narrowband-monotonous", "Stripes", "Low power", and "Artifact". Ten electroencephalographers scored 115 spectrograms and the corresponding raw EEG samples. Gwet's agreement coefficient was used to calculate IRA.

Results: Solid Flames represented seizures or IIC patterns 69.4% of the time. Irregular Flames represented seizures or IIC patterns 38.7% of the time. Broadband-monotonous primarily corresponded with seizures or IIC (54.3%) and Narrowband-monotonous with focal or generalized

[*]Corresponding author at: Massachusetts General Hospital, Lunder 6 Neurosciences Intensive Care Unit, 55 Fruit Street, Boston, MA 02114, USA. sfzafar@mgh.harvard.edu (S.F. Zafar).

Declaration of Competing Interest

The authors report no conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinph.2020.05.032.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

slowing (43.8%). Stripes were associated with burst-suppression (37.2%) and generalized suppression (34.4%). Low Power category was associated with generalized suppression (94%). There was "near perfect" agreement for Solid Flames (κ = 94.36), Low power (κ = 92.61), and Artifact (κ = 93.72). There was "substantial agreement" for all other categories (κ = 74.65–79.49).

Conclusions: This EEG spectrogram nomenclature has high IRA among electroencephalographers.

Significance: The nomenclature can be a useful tool for EEG screening. Future studies are needed to determine if using this nomenclature shortens time to IIC identification, and how best to use it in practice to reduce time to intervention.

Keywords

Continuous EEG monitoring; Quantitative EEG nomenclature; Spectrograms; Inter-rater agreement; Seizures; IIC patterns

1. Introduction

Seizures and ictal-interictal continuum (IIC) EEG patterns occur in up to 40% of critically ill patients monitored with continuous EEG (cEEG) (Claassen et al., 2007; Oddo et al., 2009; Kurtz et al., 2014; Sivaraju and Gilmore, 2016). Higher seizure and IIC burden are associated with worse functional outcomes (De Marchis et al., 2016; Zafar et al., 2018), and delay in diagnosis and treatment of non-convulsive seizures has been shown to be associated with higher mortality (Young et al., 1996). However, implementation of real-time screening of seizures and IIC using cEEG has been challenging given the limited availability of experts with training in clinical neurophysiology and the time consuming nature of raw EEG review (Gavvala et al., 2014; Moura et al., 2014).

Several studies have shown that quantitative EEG displays can streamline seizure detection by expert electroencephalographers with reasonable sensitivity and specificity (Swisher et al., 2015; Amorim et al., 2017). Nurses and other bedside providers' ability to screen seizures using quantitative EEG displays, in particular spectrograms, has also been tested, and initial reports indicate that sensitivity for seizure screening performance might be comparable to expert electroencephalographer review, despite lower specificity and overall accuracy (Amorim et al., 2017; Kang et al., 2019).

However, implementation and education on seizure screening using spectrograms and other quantitative EEG displays has been limited by lack of a standardized approach to identifying spectrogram signatures associated with seizures or IIC EEG patterns on raw EEG. A recent study evaluating seizure screening performance by expert electroencephalographers and critical care nurses using spectrograms proposed categorizing EEG spectrogram patterns into five categories that appear to be associated with high, intermediate, or low risk of seizures (Amorim et al., 2017). This proposed nomenclature for EEG spectrogram patterns provided preliminary estimates of the probability that each pattern would turn out to be an electrographic seizure upon inspection of the corresponding raw EEG. However, these estimates were considered tentative because they relied on a small number of seizure and IIC cases and only four expert electroencephalographers.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To be clinically useful, a nomenclature needs to be validated as reliable. That is, it must be possible for providers evaluating the same EEG independently to apply the nomenclature consistently. Determining whether this is the case requires a formal study of inter-rater agreement (IRA). We hypothesized that a standardized "easy to use" nomenclature will have high inter-rater agreement. Herein, we extend our prior work by proposing a refined version of the standardized nomenclature for EEG spectrogram patterns. Our primary objective was to assess the inter-rater agreement of our proposed nomenclature. Our secondary objective was to evaluate the performance of the nomenclature by assessing the probabilities that each category on the nomenclature represents IIC patterns vs. other EEG patterns encountered in ICU patients (slowing, suppression and burst suppression). We developed a lookup table or "atlas" of spectrogram patterns to 1) provide both clinical neurophysiologists and bedside providers without EEG expertise the probabilities of seeing IIC patterns vs. other patterns and expediting raw EEG review when needed, 2) serve as an educational resource on quantitative EEG analyses for trainees and other bedside providers, 3) provide a common lexicon for research questions.

2. Methods

2.1. Spectrogram EEG nomenclature

We utilize and refine the standardized nomenclature for spectrograms developed in previous published work from our group (Amorim et al., 2017). The following 7 categories have been defined: "Solid Flames", "Irregular Flames", "Broadband-monotonous", "Narrowband-monotonous", "Stripes", "Low Power" and "Artifact". Representative examples of each category are shown in Fig. 1.

- **"Solid flames":** This pattern is characterized by an abrupt appearance of spectrogram segments with high-power and bandwidth, rising "up" from the delta range into theta and often into the alpha frequency range, similar in appearance to candle flames. Solid flames have smooth edges. When they recur, they are stereotyped and each instance appears similar to the others.
- **"Irregular flames":** These events have a flame-like shape, but are characterized by edge "choppiness" or "irregularity". They tend not to be stereotyped.
- **"Broadband-monotonous":** Spectrograms with this pattern are characterized by minimal variation over time or gradual waxing and waning of sustained high power (yellow/red/white) extending across a broad band of frequencies, extending up to high (>5 Hz) frequencies.
- **"Narrowband-monotonous":** Spectrograms with this pattern have a sustained band of high-power (yellow/red/white) restricted to low (<5 Hz) frequencies. There is minimal variation within the high-power band.
- **"Stripes":** This pattern is essentially synonymous with burst suppression, and is characterized by alternation between diffuse low power (suppression), and high power "stripes" (bursts). In other words, sequential events with high power across a range of frequencies (bursts) give the spectrogram a striped appearance.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

- **"Low power":** These spectrograms are characterized by diffuse low power, and appear monotonous. There is often a narrow band of dark blue/cyan/green (low power), which should be distinguished from the higher power yellow/red band seen in narrowband-monotonous.
- **"Artifact":** This pattern is characterized by irregular high-power signal saturating the spectrogram colormap across a broad range of frequencies.

2.2. EEG acquisition and samples

Under a protocol approved by the Institutional Review Board, EEG samples were taken from a database of adult patients who underwent cEEG monitoring at Massachusetts General Hospital (MGH) between 2012 and 2017. All cEEG recordings were obtained using 21 scalp electrodes and the conventional International 10–20 system. Raw EEG data was reviewed and reported in the medical record by two clinical neurophysiologists per institutional protocol. Three clinical neurophysiologists independently reviewed raw EEG data to classify findings based on the standardized American Clinical Neurophysiology Society (ACNS) Intensive Care Unit (ICU) EEG terminology (Hirsch et al., 2013).

We selected 115 raw EEG samples that comprised 10–12 examples each of the commonly encountered cEEG patterns: seizures, lateralized and generalized periodic discharges (LPDs, GPDs), lateralized and generalized rhythmic delta activity (LRDA, GRDA), focal or generalized slowing, burst suppression, generalized suppression and artifact. Each raw EEG sample was 10 seconds in duration.

For each 10-s raw EEG sample selected, we extracted a one-hour segment of cEEG data that contained the 10-s sample. The initial sampling rate was 512 Hz. However, for our purposes 200 Hz was adequate (spectrograms were computed only up to 20 Hz), and a lower sampling rate was slightly more computationally convenient. Therefore segments were resampled to 200 Hz and used to compute spectrograms. We divided the scalp into 4 different regions for spectrogram construction: Left lateral (Fp1, F7, T3, T5, O10; Right lateral (Fp2, F8, T4, T6, O2); Left Parasagittal (Fp1, F3, C3, P3, O1); Right Parasagittal (Fp2, F4, C4, P4, O2). Spectrogram images were computed using multitaper spectral estimation (Thomson, 1982; Bokil et al., 2010), with a window size of 2 seconds and a 1 second overlap. All signal processing was performed using MATLAB (Natick, MA). The spectrograms generated by MATLAB were similar in appearance and color scheme to the quantitative EEG and spectrogram displays provided by standard bedside EEG software typically used in real time clinical practice.

2.3. Inter-rater agreement: survey and raters

We used a web-based testing system (SurveyMonkey, Inc; San Mateo, CA) to evaluate inter-rater agreement of the proposed spectrogram nomenclature and to quantify associations between spectrogram and raw EEG patterns. The test was comprised of 230 images (115 raw EEG samples and the corresponding 115 spectrogram images), each with multiple choice answer options. The spectrogram and raw images were presented in random order, such that consecutive spectrogram and raw EEG images were not related. Each raw EEG image included a 10 second clip provided with multiple choice options: 1) Seizure, 2)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Lateralized Periodic Discharges (LPDs), 3) Generalized Periodic Discharges (GPDs), 4) Lateralized Rhythmic Delta Activity (LRDA), 5) Generalized Rhythmic Delta Activity (GRDA), 6) Focal Slowing, 7) Generalized Slowing, 8) Burst-Suppression, 9) Generalized Suppression, 10) Artifact. Raters were provided the option "other" if in their opinion the raw EEG clip did not fall into any of the pattern categories listed in the choices. Each 1-hour spectrogram image was associated with the following multiple-choice options: 1) Solid flames, 2) Irregular flames, 3) Broadband-monotonous, 4) Narrowband-monotonous, 5) Stripes, 6) Low power, 7) Artifact. A green arrow in each spectrogram identified the segment from which the corresponding raw EEG clip was obtained (Fig. 2). Raters were asked to choose the spectrogram pattern that in their opinion was representative of the pattern immediately below the green marker (Fig. 2). A reference guide with a representative spectrogram image for each nomenclature category was provided and could be used while completing the survey.

Additionally, the survey had questions pertaining to EEG training and experience for each rater. Ten fellowship trained electroencephalographers from 6 centers completed the survey. Raters were blinded to the original EEG reports and findings.

2.4. Statistical analysis

We calculated inter-rater agreement using Gwet's multi-rater agreement coefficient AC (Feinstein and Cicchetti, 1990; Gwet, 2008, 2010). We categorized j values following the convention: slight agreement 10–20%; fair agreement 20–40%; moderate agreement 40–60%; substantial agreement 60–80%; near perfect agreement 80–100% (Landis and Koch, 1977). 95% confidence intervals were calculated for the estimated κ values.

We plotted agreement matrices to identify sources of disagreement. The "gold standard" or "best answer" for both spectrogram and raw EEG images was defined to be the majority response.

Finally, we tabulated the frequency with which each raw EEG pattern corresponded with each spectrogram category as an estimate of the probability. These frequencies serve as an estimate of the probability that, when one sees a given spectrogram pattern, that review of the corresponding raw EEG will reveal seizures, IIC or other patterns.

3. Results

Rater characteristics for the 10 electroencephalographers who participated in the study are shown in Table 1. All had at least two years of experience reading ICU EEG, had passed the ACNS ICU EEG certification test, and routinely used quantitative EEG.

Fig. 3 shows the percentages of observed agreement and estimated chance-corrected level of IRA (κ value) with 95% confidence intervals for spectrograms and raw EEG. For spectrogram patterns, there was "near perfect" agreement for solid flames, low power, and artifact categories. There was "substantial agreement" for all the other spectrogram categories. As two of the raters (MBW and ESR) had been involved in the design and execution of prior work on the nomenclature (Amorim et al., 2017), we repeated analysis

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

after excluding these raters and found similar inter-rater agreement: Solid flames: kappa 94.56 [91.36–97.76]; Irregular flames: kappa 77.97 [71.10–84.38]; broadband-monotonous: kappa 78.40 [71.26–85.54]; narrowband-monotonous: kappa 74.77 [67.28–82.27]; stripes: kappa 79.49 [73.08–85.90]; low power: kappa 93.16 [89.52–96.80]; artifact: kappa 94.10 [91.26–96.94].

For raw EEG patterns, there was “near perfect” agreement for GPDs, LRDA, GRDA, seizures, burst-suppression and artifact. The remaining patterns all showed “substantial agreement”.

Agreement matrices showing patterns of disagreement are shown in Fig. 4. For both spectrogram and raw EEG the incorrect responses were generally close or similar to the correct or gold standard. For example, when the majority choice was solid flames, a small minority chose irregular flames. Similarly, when the majority answer was broadband-monotonous, most deviant responses were either irregular flames or narrowband-monotonous. A similar pattern was seen for raw EEG ratings (Fig. 4b). For example, when the majority response was seizures, the majority of the incorrect responses were for periodic and rhythmic patterns. When the majority response was generalized slowing, most incorrect responses were GRDA or suppression. This is likely a result of the challenge in choosing between categories for borderline or overlapping cases.

Fig. 4c shows the frequency with which raw EEG patterns correspond to each spectrogram category. Solid flames represented seizures or IIC EEG patterns 69.4% of the time. Irregular flames represented seizures or IIC EEG patterns 38.7% of the time, and focal or generalized slowing 32.3% of the time. Broadband-monotonous spectrograms primarily represented seizures or IIC patterns (54.3%) and focal/generalized slowing (32.9%). Narrowband-monotonous spectrograms primarily represented focal or generalized slowing (43.8%), followed by IIC EEG patterns (24.2%). Stripes primarily represented burst suppression (37.2%), and generalized suppression (34.4%). Low power spectrograms almost entirely represented generalized suppression (94%). The atlas provided in Supplement 1 shows examples of spectrogram patterns and the frequency of corresponding raw EEG findings.

Although the study was not designed to determine sensitivity of spectrogram patterns, we evaluated the diagnostic accuracy of solid flames for detection of seizures. The “gold standard” for seizures on raw EEG clips was determined by majority consensus (i.e. when seizures was the single most common response). The “gold standard” for solid flames as determined by majority consensus (i.e. solid flame was the single most common response). The Solid flames had a sensitivity of 87.5% (7/8) and specificity of 92.5% (99/107) for the detection of seizures. The positive likelihood ratio (LR+) was 12 and the overall accuracy of solid flames for detection of seizures was 92.2%. The pre-test probability of seizures was 7% and post-test probability was 47% (Fagan’s nomogram shown in Supplement 2).

4. Discussion

This study indicates that a spectrogram nomenclature for EEG patterns has a high inter-rater agreement and can be reliably used by expert electroencephalographers. In evaluating the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

performance of the nomenclature we found that seizures were nearly always categorized as solid flames, supporting our preliminary study that this is a very sensitive spectrogram signature (Amorim et al., 2017). Not surprisingly, there was more variability in the spectrogram categories attributed to IIC EEG patterns; however those were nearly always categorized under the flames or broadband categories, supporting the potential use of this nomenclature not only to screen for seizures, but also for IIC EEG patterns. Though not the primary focus of our study, we also found a high inter-rater agreement for raw EEG patterns; this high agreement may be because participating raters were all familiar with the standardized ACNS nomenclature which has now been available for several years and is being increasingly utilized for critical care EEG reporting.

The atlas we developed (Supplement 1) can serve as a tool to support more nuanced screening, seizure surveillance and “alarm activation” by bedside providers. Once an EEG spectrogram category is identified, the atlas can be used as a benchmark to gauge the likelihood of seizure or IIC risk for that pattern, allowing providers to go beyond an “all or none” approach to seizure screening and to calibrate the level of concern for events that might be actionable or require further investigation. For example, a solid flame alarm carries a very different seizure risk compared to a stripes or low power pattern. A prospective study of comparing use of the proposed nomenclature and atlas with current standard EEG screening practices both among expert neurophysiologists and non neurophysiologist bed side providers is needed to better define the potential clinical impact of the nomenclature.

Identifying the specific EEG spectrogram pattern that corresponds with a seizure already identified on raw EEG, and vice-versa, has the potential to reduce the frequency of false alarms and increase specificity and reduce timing of seizure recurrence diagnosis. This study showed high inter-rater agreement for both our EEG spectrogram nomenclature system and for the corresponding raw EEG classification using the ACNS ICU EEG terminology, supporting previous studies showing that the spectrogram and raw EEG signatures rating is consistent despite the heterogeneity of the raw EEG examples (Gerber et al., 2008; Mani et al., 2012; Gaspard et al., 2014; Amorim et al., 2017). That was true for EEG spectrograms associated with seizure and IIC EEG patterns (flames and broadband), but also for those infrequently associated with epileptiform abnormalities (stripes or low power). This result further motivates the implementation of this nomenclature in future studies of seizure screening by expert and non-expert bedside providers in the critical care setting.

Raw EEG review is a time-intensive task for expert electroencephalographers. Review of 24-hours of raw EEG data by an experienced neurophysiologist requires 38 minutes on average (Moura et al., 2014). The reviewing time can be further reduced to an average of 8 minutes when spectrograms or quantitative EEG displays are used (Moura et al., 2014). Moreover, several studies have shown that quantitative EEG displays can aid rapid and early detection of seizures both by neurophysiologists and non-neurophysiologists (Stewart et al., 2010; Williamson et al., 2014; Swisher et al., 2015; Haider et al., 2016). Sensitivity of quantitative EEG guided seizure screening in these studies varies widely between clinicians (66–92%) and nurses (64–95%), highlighting the heterogeneity in study methodology, training of non-experts, and examples of seizure and IIC EEG patterns included. False alarms are an important limitation of seizure screening systems using quantitative EEG alone, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

previous studies identified up to one false alarm every 6-h (Amorim et al., 2017). Our spectrogram atlas may simplify training for and implementation of seizure screening with quantitative EEG, and may assist bedside providers - both expert electroencephalographers and non-experts - in shortening EEG screening time and streamline earlier identification, confirmation on raw EEG, and treatment of seizures or IIC EEG patterns.

There are limitations to this study. First, the 115 examples selected aimed to cover a wide range of variations of seizures and IIC EEG patterns. The prevalence of these EEG findings in our study may not reflect routine clinical practice, and the increased difficulty of adding more challenging-to-categorize patterns might mean that the estimates provided in our atlas underestimate the true level of inter rater agreement and seizure/IIC risk that can be achieved in practice. Second, the inter-rater agreement evaluation was restricted to expert electroencephalographers, and did not include trainees, nurses or other bedside providers. However, the high interrater agreement of our EEG spectrogram nomenclature and atlas indicate that this tool might be useful in training non-expert providers in workshops as done in previous studies. Third, experts involved in the evaluation of the study were blinded to patient history and were only provided with 1-h of spectrogram data, therefore it is possible that real-world interrater agreement performance could have been further improved by collateral information and additional EEG data. Finally, raters in our study had to categorize spectrogram and EEG patterns by selecting from a limited number of options. It is possible that inter-rater agreement would be different if the options had been different, or if raters had been permitted to answer using free form responses.

5. Conclusion

A simple nomenclature for screening EEG for IIC vs. other patterns using EEG spectrograms has high inter-rater agreement among expert electroencephalographers. The seizure and IIC risk score/probabilities provided in the atlas is expected to be able to help experts and non-expert bedside providers who have quantitative EEG tools available to screen the EEG more rapidly, and to better recognize which types of changes in EEG spectrogram patterns are more likely to be seizures or IIC patterns, and thus warrant more urgent direct review of the raw EEG. Future prospective studies are needed to test if this decision support tool may improve early and accurate identification of seizures, expedite treatment, and reduce false-alarm rates.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank members of the Critical Care EEG Monitoring Research Consortium for valuable feedback on the study.

Funding/Disclosures

This research was supported by funding from NIH-NINDS K23NS114201 (PI- SFZ). MBW received funding from NIH-NINDS (NIH-NINDS 1R01NS102190, 1R01NS102574, 1R01NS107291) not related to this work. SFZ, ESR, JJ, and MBW received funding from SAGE Therapeutics not related to this work. JWL does contract work for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SleepMed/DigiTrace, Advance Medical; Site PI for Engage Therapeutics; Research funding from NINDS, not related to this work. HAH received consultant support from Ceribell, Inc.; author royalties from UpToDate Inc. and Springer Publishing not related to this work. EA received funding from the Neurocritical Care Society, American Heart Association, and Massachusetts Institute of Technology-Philips Alliance not related to this work Amorim et al., 2017; Bokil et al., 2010; Claassen et al., 2007; De Marchis et al., 2016; Feinstein and Cicchetti, 1990; Gaspard et al., 2014; Gavvala et al., 2014; Gerber et al., 2008; Gwet, 2010, 2008; Haider et al., 2016; Hirsch et al., 2013; Kang et al., 2019; Kurtz et al., 2014; Landis and Koch, 1977; Mani et al., 2012; Moura et al., 2014; Oddo et al., 2009; Sivaraju and Gilmore, 2016; Stewart et al., 2010; Swisher et al., 2015; Thomson, 1982; Williamson et al., 2014; Young et al., 1996; Zafar et al., 2018.

Abbreviations:

ACNS	American Clinical Neurophysiology Society
cEEG	continuous EEG
GRDA	generalized rhythmic delta activity
GPDs	generalized periodic discharges
ICU	intensive care unit
IIC	ictal-interictal continuum
IRA	Inter rater agreement
LPDs	lateralized periodic discharges
LRDA	lateralized rhythmic delta activity
MGH	Massachusetts General Hospital

References

Amorim E, Williamson CA, Moura LM, Shafi MM, Gaspard N, Rosenthal ES, et al. Performance of spectrogram-based seizure identification of adult EEGs by critical care nurses and neurophysiologists. J Clin Neurophysiol. 2017;34:359–64. [PubMed: 27930420]

Bokil H, Andrews P, Kulkarni JE, Mehta S, Mitra PP. Chronux: a platform for analyzing neural signals. J Neurosci Methods 2010;192(1):146–51. 10.1016/j.jneumeth.2010.06.020. [PubMed: 20637804]

Claassen J, Jette N, Chum F, Green R, Schmidt M, Choi H, et al. Electrographic seizures and periodic discharges after intracerebral hemorrhage. Neurology 2007;69:1356–65. [PubMed: 17893296]

De Marchis G, Pugin D, Meyers E, Velasquez A, Suwatcharangkoon, Park S, et al. Seizure Burden in Subarachnoid Hemorrhage Associated With Functional and Cognitive Outcome. Neurology 2016;86:253–60. [PubMed: 26701381]

Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43:543–9. [PubMed: 2348207]

Gaspard N, Hirsch LJ, LaRoche SM, Hahn CD, Westover MB. Critical Care EEG Monitoring Research Consortium. Interrater agreement for critical care EEG terminology. Epilepsia. 2014;55:1366–73. [PubMed: 24888711]

Gavvala J, Abend N, LaRoche S, Hahn C, Herman ST, Claassen J, et al. Continuous EEG monitoring: a survey of neurophysiologists and neurointensivists. Epilepsia 2014;55:1864–71. [PubMed: 25266728]

Gerber PA, Chapman KE, Chung SS, Drees C, Maganti RK, Ng Y-T, et al. Interobserver agreement in the interpretation of EEG patterns in critically ill adults. J Clin Neurophysiol 2008;25:241–9. [PubMed: 18791475]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

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]

Gwet KL. Handbook of inter-rater reliability: the definitive guide to measuring the extent of agreement among raters. Gaithersburg, MD: Advanced Analytics, LLC; 2010.

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. [PubMed: 27466474]

Hirsch LJ, LaRoche SM, Gaspard N, Gerard E, Svoronos A, Herman ST, et al. American clinical neurophysiology Society's standardized critical care EEG terminology: 2012 version. J Clin Neurophysiol. 2013;30:1–27. [PubMed: 23377439]

Kang JH, Sherill GC, Sinha SR, Swisher CB. A trial of real-time electrographic seizure detection by neuro-ICU nurses using a panel of quantitative EEG trends. Neurocrit Care 2019;20:1–9.

Kurtz P, Gaspard N, Wahl AS, Bauer RM, Hirsch LJ, Wunsch H, et al. Continuous electroencephalography in a surgical intensive care unit. Intensive Care Med 2014;40:228–34. [PubMed: 24240843]

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74. [PubMed: 843571]

Mani R, Arif H, Hirsch LJ, Gerard EE, LaRoche SM. Interrater reliability of ICU EEG Research Terminology. J Clin Neurophysiol 2012;92:203–12.

Moura LM, Shafi MM, Ng M, Pati S, Cash SS, Cole AJ, et al. Spectrogram screening of adult EEGs is sensitive and efficient. Neurology 2014;83:56–64. [PubMed: 24857926]

Oddo M, Carrera E, Claassen J, Mayer SA, Hirsch LJ. Continuous electroencephalography in the medical intensive care unit. Critical Care Med 2009;37:2051–6. [PubMed: 19384197]

Sivaraju A, Gilmore EJ. Understanding and managing the ictal-interictal continuum in neurocritical care. Curr Treat Options Neurol 2016;18:8. [PubMed: 26874841]

Stewart CP, Otsubo H, Ochi A, Sharma R, Hutchison JS, Hahn CD. Seizure identification in the ICU using quantitative EEG displays. Neurology 2010;75:1501–8. [PubMed: 20861452]

Swisher CB, White CR, Mace BE, Dombrowski KE, Husain AM, Kolls BJ, et al. Diagnostic accuracy of electrographic seizure detection by neurophysiologists and non-neurophysiologists in the adult ICU using a panel of quantitative EEG trends. J Clin Neurophysiol 2015;32:324–30. [PubMed: 26241242]

Thomson DJ. Spectrum estimation and harmonic analysis. Proc IEEE 1982;70:1055–96.

Williamson CA, Wahlster S, Shafi MM, Westover MB. Sensitivity of compressed spectral arrays for detecting seizures in acutely ill adults. Neurocrit Care 2014;20:32–9. [PubMed: 24052456]

Young BG, 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. [PubMed: 8710130]

Zafar SF, Postma EN, Biswal S, Boyle EJ, Bechek S, O'Connor K, et al. Effect of epileptiform abnormality burden on neurologic outcome and antiepileptic drug management after subarachnoid hemorrhage. Clin Neurophysiol 2018;129:2219–27. [PubMed: 30212805]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

HIGHLIGHTS

- The proposed standardized spectrogram EEG nomenclature has high inter-rater agreement.
- The probability of ictal-interictal continuum (IIC) vs. other patterns with each nomenclature category can aid EEG screening.
- Prospective studies will determine if the nomenclature can expedite IIC detection and treatment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

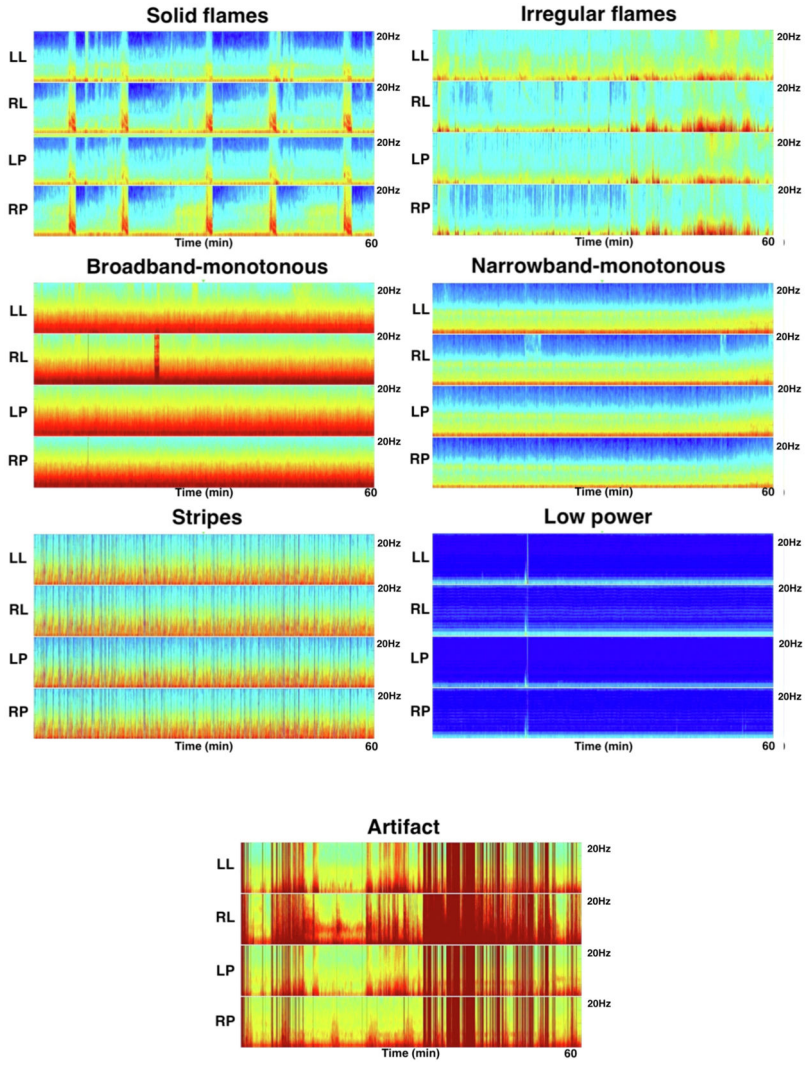

Fig. 1.

Quantitative EEG Spectrogram Nomenclature. Representative examples of the spectrogram nomenclature are shown. The scalp is divided into 4 different regions for spectrogram construction: Left lateral (Fp1, F7, T3, T5, O10; right lateral (Fp2, F8, T4, T6, O2); left parasagittal (Fp1, F3, C3, P3, O1); right parasagittal (Fp2, F4, C4, P4, O2). Each spectrogram image has four panels: left lateral (LL), right lateral (RL), left parasagittal (LP), right parasagittal (RP). Each spectrogram image shows one hour of recording. The vertical axis represents the spectrogram frequency from 0–20 Hz. “Solid flames” are characterized by an abrupt appearance of higher power and bandwidth, and have regular and have smooth edges. “Irregular flames” are characterized by choppiness and do not have smooth edges or a regular appearance. “Broadband-monotonous” is characterized by sustained higher power at low frequencies with minimal variation or very gradual waxing and waning of frequencies within the high-power band. “Narrowband-monotonous” is characterized by a sustained <5 Hz band of high power (yellow/red) with minimal variation within the high-power band. “Stripes” represent a burst suppressed background and characterized by rapid alternation between diffuse low power and high power and high frequencies. “Low power”

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

spectrograms are characterized by diffuse low power, and appear monotonous. “Artifact” is characterized by irregular high-power signal saturating all frequencies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

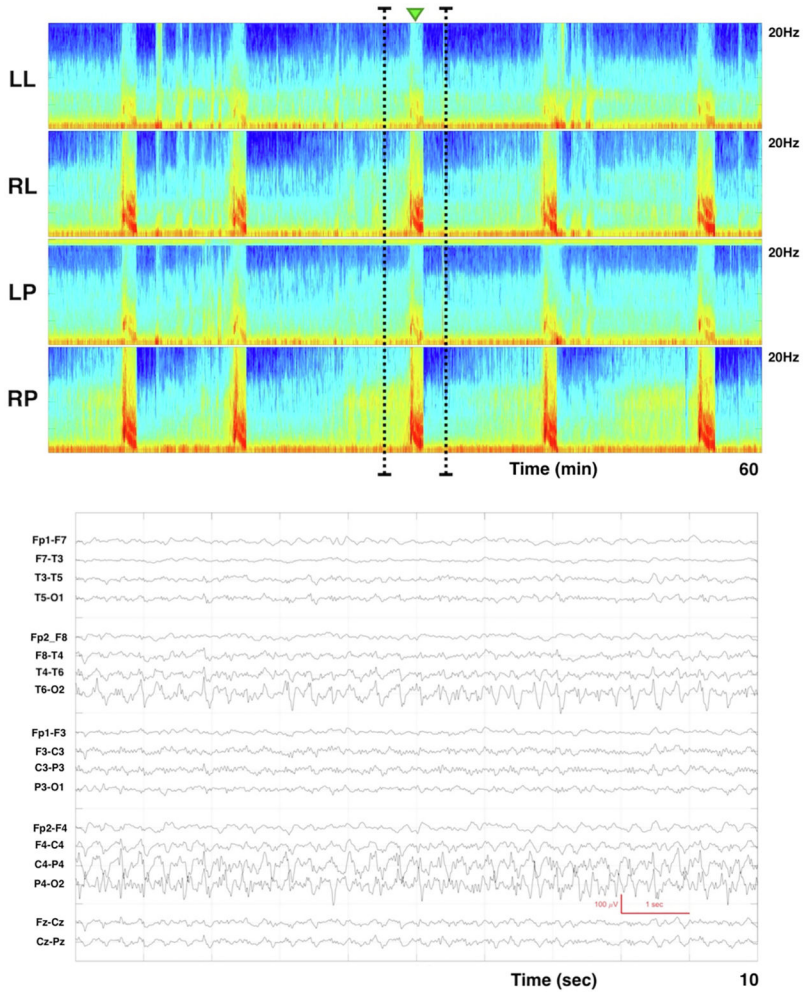

Fig. 2.

Example of spectrogram image and corresponding raw EEG clip. Each spectrogram image has four panels: Left lateral (LL), right lateral (RL), left parasagittal (LP), right parasagittal (RP). We divided the scalp into 4 different regions for spectrogram construction: Left lateral (Fp1, F7, T3, T5, O10; Right lateral (Fp2, F8, T4, T6, O2); Left Parasagittal (Fp1, F3, C3, P3, O1); Right Parasagittal (Fp2, F4, C4, P4, O2). Above each image is a small green arrow that represents the segment from which the corresponding raw EEG clip was obtained. Raters were asked to choose the spectrogram nomenclature that in their opinion was representative of the pattern immediately below the green arrow. The raw EEG image shows a 10 second clip of recording in a longitudinal bipolar montage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

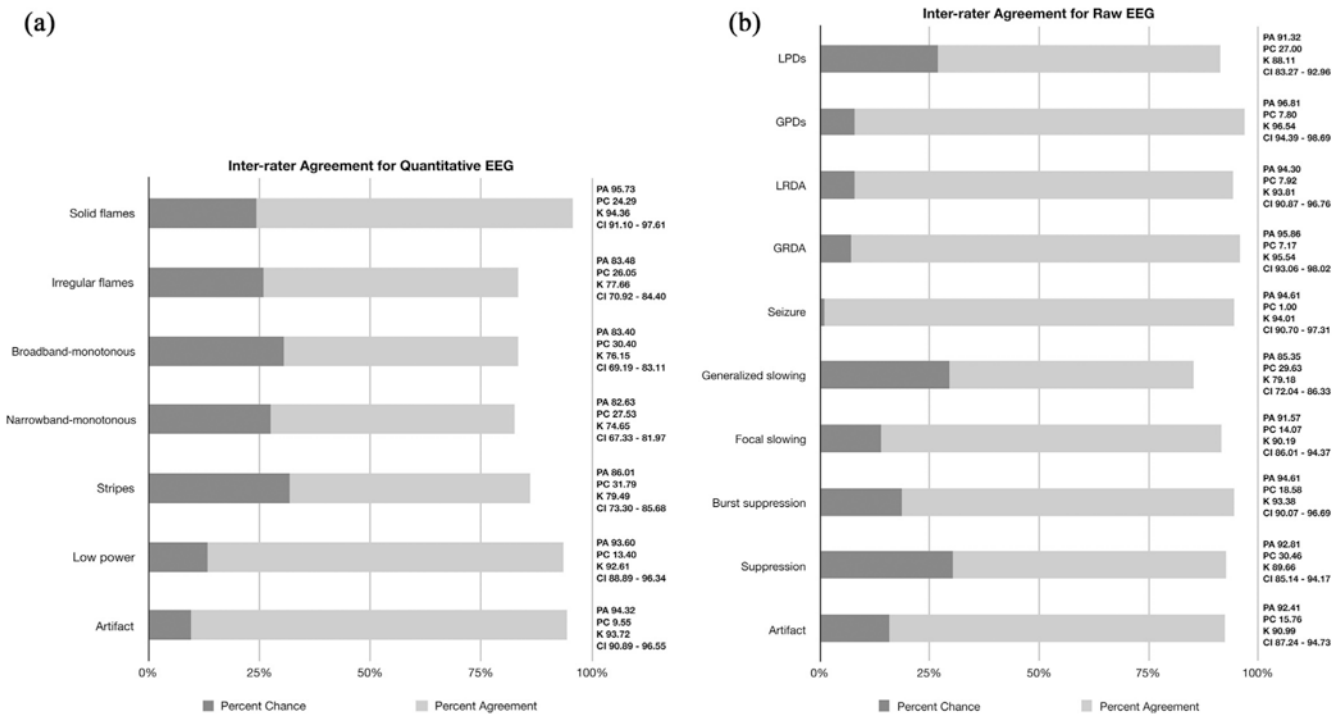

Fig. 3.

Inter-rater agreement for quantitative and raw EEG. In inter-rater agreement calculation, a portion of the observed percent agreement (PA), is assumed to be chance attributed (PC). The inter-rater agreement statistical methods are used to determine the percent agreement beyond chance (kappa). Fig. 3a. PA, PC, κ and CI for quantitative EEG spectrogram nomenclature are shown. Fig. 3b. PA, PC, κ and CI for raw EEG are shown. PA: Percent agreement; PC: Percent chance; CI: confidence interval; κ: kappa; κ 10–20% = slight agreement; κ 20–40% =fair agreement; κ 40–60%=moderate agreement; κ 60–80%=substantial agreement; κ 80–100%= near perfect agreement (Landis and Koch, 1977). GPD: generalized periodic discharge; GRDA: generalized rhythmic delta activity; LPD: lateralized periodic discharges; LRDA: lateralized rhythmic delta activity.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

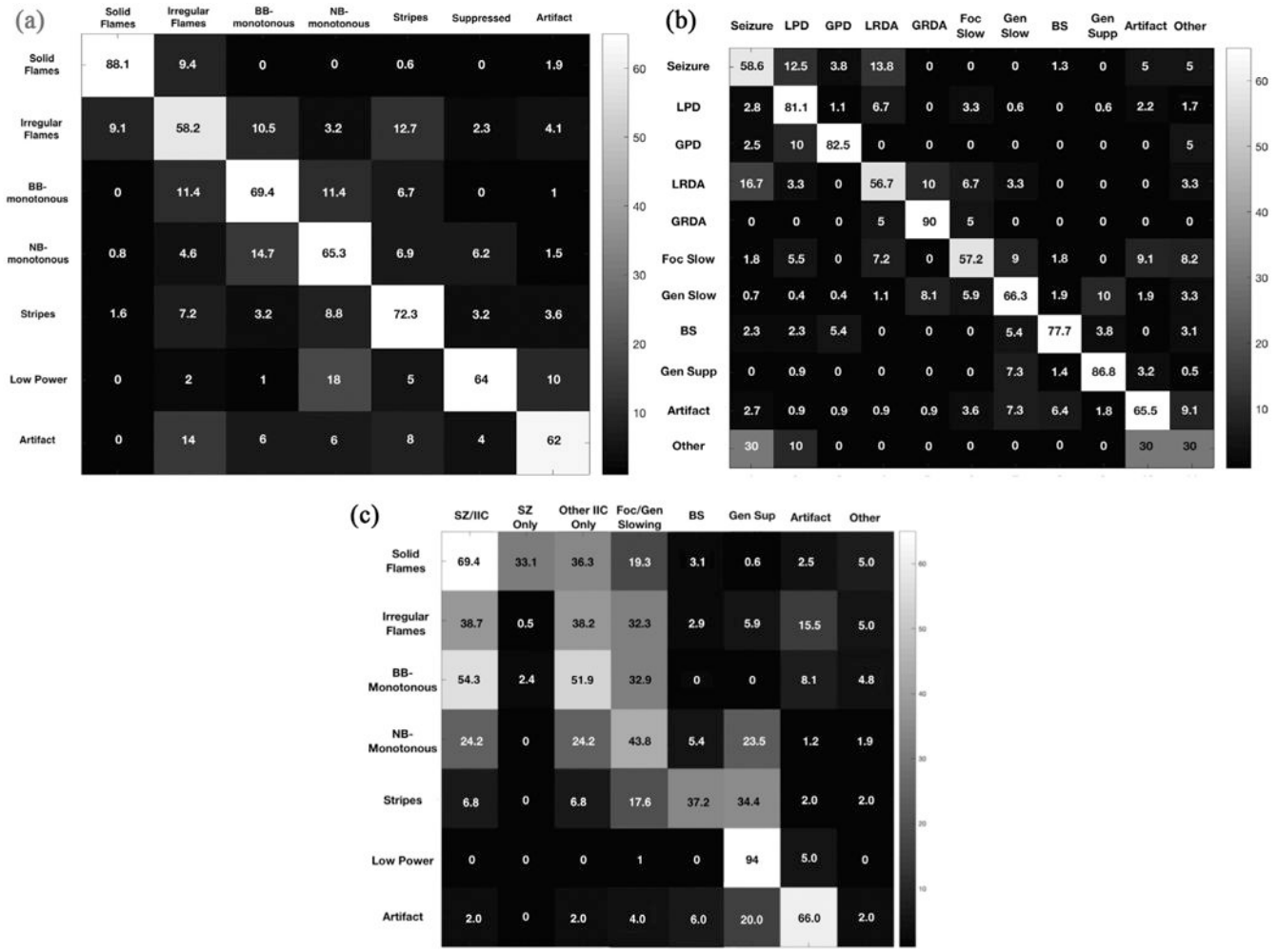

Fig. 4.

Agreement matrices. a. Agreement matrix showing sources of disagreement for spectrogram nomenclature is presented as a heat map. The spectrogram nomenclature categories are shown along the vertical and horizontal axes. Heat map intensities show the percentage of times each nomenclature option on the horizontal axis was chosen when the correct response (gold standard determined by majority consensus) was the one on the vertical axis. 100% agreement would produce a diagonal white line from upper left to lower right, with all the remaining squares not on that line being black. The figure demonstrates the majority had agreement in rating the spectrogram nomenclature. Any disagreement noted was primarily for patterns close or similar to each other. For example, when the gold standard response was solid flames, a small minority chose the incorrect response of irregular flames. Similarly when the correct answer was broadband-monotonous most incorrect responses were either irregular flames or narrowband-monotonous. BB-monotonous: Broadband-monotonous; NB-monotonous: Narrowband-monotonous. b. Agreement matrix showing sources of disagreement for raw EEG is presented as a heat map. The raw EEG categories are shown along the vertical and horizontal axes. Heat map intensities show the percentage of times each option on the horizontal axis was chosen when the correct response (gold standard determined by majority consensus) was the one on the vertical axis. 100% agreement would

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

produce a diagonal white line from upper left to lower right, with all the remaining squares not on that line being black. The figure demonstrates the majority had agreement in rating the raw EEG. Similar to spectrogram nomenclature, any disagreement was noted primarily for patterns close to each other. For example when the correct or gold standard response was seizures, the small number of incorrect responses were periodic and rhythmic patterns. When the correct response was generalized slowing, incorrect responses were GRDA or suppression. BS: burst suppression; Foc slow: focal slowing; Gen slow: generalized slowing; Gen supp: generalized suppression; GPD: generalized periodic discharge; GRDA: generalized rhythmic delta activity; LPD: lateralized periodic discharges; LRDA: lateralized rhythmic delta activity. c. Agreement matrix showing relation of spectrogram nomenclature with raw EEG patterns is presented as a heat map. Heat map intensities show the frequency (in percent) each option on the horizontal axis was chosen when the correct spectrogram response (gold standard determined by majority consensus) was the one on the vertical axis. SZ/IIC: The frequency with which seizure or any other IIC pattern (i.e. LPDs/GPDs/LRDA/GRDA) were selected. SZ only: The frequency with which seizure was selected; Other IIC only: The frequency with which IIC patterns other than seizure (i.e. LPD/GPDs/LRDA/GRDA) were selected; Foc/Gen Slow: The frequency with which focal or generalized slowing were selected; BS: The frequency with which burst suppression was selected. BB: broadband; BS: burst suppression; Foc slow: focal slowing; Gen slow: generalized slowing; Gen supp: generalized suppression; GPD: generalized periodic discharge; GRDA: generalized rhythmic delta activity; IIC: Ictal-interictal; LPD: lateralized periodic discharges; LRDA: lateralized rhythmic delta activity; NB: narrowband; SZ: seizures.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Characteristics of Electroenncephalographers.

	N (%)
Role group: Attending/faculty	10 (100%)
Years of experience reading EEGs:	
<2 years	0 (0%)
2–5 years	3 (30%)
5–10 years	5 (50%)
>10 years	2 (20%)
Years of experience reading quantitative EEGs:	
<2 years	0 (0%)
2–5 years	3 (30%)
5–10 years	5 (50%)
>10 years	2 (20%)
Quantitative EEG routinely incorporated in clinical practice	
Yes	10 (100%)
No	0 (0%)
Passed ACNS ICU EEG certification test	10 (100%)
Confidence in quantitative EEG interpretation	
Not confident	1 (10%)
Somewhat confident	1 (10%)
Very confident	5 (50%)
Extremely confident	3 (30%)

ACNS: American Clinical Neurophysiology Society; ICU: Intensive Care Unit.