Seizure detection with automated EEG analysis: A validation study focusing on periodic patterns

Alba Sierra-Marcos [a], Mark L. Scheuer [b], Andrea O. Rossetti [a,*]

[a] Department of Clinical Neurosciences, Centre Hospitalier Universitaire Vaudois (CHUV), and University of Lausanne, Lausanne, Switzerland
[b] Persyst Development Corporation, San Diego, CA, USA

HIGHLIGHTS

- Automated EEG interpretation is increasingly used in ICU recordings, where periodic discharges are relatively common.
- This study shows that the overall performance of automated EEG interpretation is good.
- However, the presence of periodic discharges seems to reduce its reliability.

ABSTRACT

Objective: To evaluate an automated seizure detection (ASD) algorithm in EEGs with periodic and other challenging patterns.
Methods: Selected EEGs recorded in patients over 1 year old were classified into four groups: A. Periodic lateralized epileptiform discharges (PLEDs) with intermixed electrical seizures. B. PLEDs without seizures. C. Electrical seizures and no PLEDs. D. No PLEDs or seizures. Recordings were analyzed by the Persyst P12 software, and compared to the raw EEG, interpreted by two experienced neurophysiologists; Positive percent agreement (PPA) and false-positive rates/hour (FPR) were calculated.
Results: We assessed 98 recordings (Group A = 21 patients; B = 29, C = 17, D = 31). Total duration was 82.7 h (median: 1 h); containing 268 seizures. The software detected 204 (=76.1%) seizures; all ictal events were captured in 29/38 (76.3%) patients; in only in 3 (7.7%) no seizures were detected. Median PPA was 100% (range 0–100; interquartile range 50–100), and the median FPR 0/h (range 0–75.8; interquartile range 0–4.5); however, lower performances were seen in the groups containing periodic discharges.
Conclusion: This analysis provides data regarding the yield of the ASD in a particularly difficult subset of EEG recordings, showing that periodic discharges may bias the results.
Significance: Ongoing refinements in this technique might enhance its utility and lead to a more extensive application.

© 2014 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Electroencephalography (EEG) represents a standard examination of brain function in patients living with epilepsy, and it is also increasingly used in subjects with acute consciousness impairment in an intensive care setting, where seizures may occur in 8–34% of patients, mostly with subclinical presentations (Privitera et al.,

1994; DeLorenzo et al., 1998; Towne et al., 2000; Scheuer, 2002; Claassen et al., 2004). In both clinical situations, long-term EEG monitoring allows continuous surveillance of the cerebral activity, but it implies a thorough and time-consuming interpretation by trained neurophysiologists. Over the last few decades, automatic methods have been developed to highlight significant electrographic events, providing insight into the EEG trends, reducing evaluation time and potentially increasing patient security by alerting medical staff sooner concerning the presence of seizures. Seizure detection algorithms are available in commercial software packages, based on the analysis of rhythmic patterns with a certain

waveform morphology, distribution, and evolution over time, until a threshold is achieved. Nevertheless, a major problem is the inter-patient (and, at times, intra-patient) variability of ictal patterns, ranging from quasi rhythmic or periodic discharges, over high frequency activities variation or abrupt phase changes, to more irregular groups of epileptiform transients (Gotman, 1990; Furbass et al., 2012).

Despite the increasing implementation of these automated systems and the previous literature regarding the sensitivity or specificity rates to detect seizures (Gotman, 1982, 1990; Pauri et al., 1992; Gabor, 1998; Wilson et al., 2004; Saab and Gotman, 2005; Wilson, 2005; Meier et al., 2008; Kelly et al., 2010; Hartmann et al., 2011), the application of these algorithms to patients with particular challenging EEG records has received little attention to date. Our study was conducted to address this aspect.

2. Methods

2.1. Study population

Routine and long-term EEG recordings, acquired using 23 scalp electrodes placed according to the international 10–20 system in adults and children over 1 year of age, were used for this analysis. All studies were recorded on a digital system (Viasys Neurocare, Madison, WI, USA), for standard evaluations (with intermittent photic stimulation and hyperventilation in certain cases), or in the in-patient clinics including the Intensive Care Unit (ICU), between January 2012 and March 2013.

Recordings were retrospectively selected for this study according to their features, and classified into the following groups:

A. Periodic lateralized epileptiform discharges (PLEDs, synonymous with lateralized periodic discharges (Hirsch et al., 2013)) or periodic epileptiform discharges (PEDs, synonymous with generalized periodic discharges (Hirsch et al., 2013)) with intermixed electrical seizures, defined as acceleration of their frequency over at least 2 Hz within a few seconds and/or progressive change in field and morphology, of more than 1 s of duration (Fig. 1).
B. Monotonous PLEDs/PEDs with no seizures (first control group).
C. Electrical seizures without PLEDs/PEDs, defined as any variation of the background amplitude, together with an acceleration of frequency over 4 Hz (Hirsch et al., 2013) or a spike-and wave pattern below 4 Hz consistent with typical seizures associated with idiopathic (genetic) generalized epilepsy, and/or progressive change in field and morphology, regardless of duration.
D. No PLEDs/PEDs nor seizures (second control group), in recordings suggestive of changes in vigilance states, containing considerable artifacts, or epileptiform-looking variants of normality suggesting seizures to an untrained reader, such as prolonged runs of rhythmic mid-temporal discharges, or wicket spikes.

2.2. Persyst 12 seizure detection description

This algorithm (version 2013.06.25), is designed to evaluate the scalp EEG signal for changes in background activity exhibiting rhythmicity, evolution in amplitude and/or frequency, and asymmetry, and produce outputs concerning electrographic seizure activity. The algorithm is built by combining the output of many small artificial neural networks (few input and hidden nodes), each of which were trained to recognize a particular feature, e.g. frequency evolution. Uncertainty is propagated through each level

of the processing, resulting in a single "is-seizure" output (0–1) per each one-second epoch. Two outputs of the algorithm are available: the identification of discrete electrographic seizure events with a minimum duration of 11 seconds, which we used for this analysis, and a seizure probability curve that displays the probability that any one-second epoch would be marked as "seizure". The algorithm was trained on a set of varied EEG recordings containing seizures identified by human experts, drawn from Epilepsy Monitoring Unit, Intensive Care Unit, and ambulatory settings. Non-seizure records were also utilized in training the detection algorithm; these contained a broad sampling of EEG patterns and states, including records from the ICU. There was no algorithm training to specifically or systematically attempt to identify and differentiate certain types of patterns, like lateralized or generalized periodic discharges, or gray area ICU patterns that fall into the ictal-interictal continuum (where expert electroencephalographers might have legitimate disagreements in interpretation).

2.3. Algorithm interpretation

Each recording was analyzed using the automated seizure detection (ASD) function, which presents on the screen as a red mark above a given threshold, being considered as a binary variable for practical purposes. We also inspected the color-coded rhythmicity and fast Fourier transformed (FFT) power spectrograms (both averaged for each hemisphere), in order to determine the presence of patterns suggestive of electrographic seizures. These analyses were always assessed using the artifact reduction device, on a time scale between 30 min and 1 h. The visual interpretation of the corresponding whole raw recording represented the gold standard comparator; all analyses were the result of agreement between two experienced neurophysiologists (ASM, AOR).

2.4. Data analysis

On the whole dataset, we calculated the percentage of seizures detected (either with the ASD only, or using also the spectrograms), as well as the rate of patients with seizures in whom at least one event was detected by the software. False negatives (FN) were defined as expert-marked seizures on raw recordings, not detected by any automated method. Expert-marked seizures that were also detected by the automated software represented true positives (TP). Positive percentage agreement (PPA), a term preferred rather than "sensitivity" to describe the comparison of a new test to a non-reference standard, was calculated for each record individually, regarding both automated approaches: TP was divided by the sum of TP and FN for each patient, and described with a median (given the assumed non-normal distributions) for each group, and in order to reduce any biasing by individual patients having many seizures.

False positive (FP) detections were defined as any event identified by the ASD as a seizure, but not corresponding to any expert-marked seizure. The false positive rate (FPR), corresponding to FP divided by recording time, was described using the median of each individual FPRs across a given group; to complement this approach we also calculated FPR using the sum of individual FP divided by the whole recording time by group.

3. Results

We analyzed 98 recordings: 21 patients in Group A (with 170 seizures), 29 in group B, 17 in group C (with 98 seizures), and 31 in group D; globally, these recordings contained 268 seizures (groups A and C). In group A, the median number of seizures per

(a)

(b)
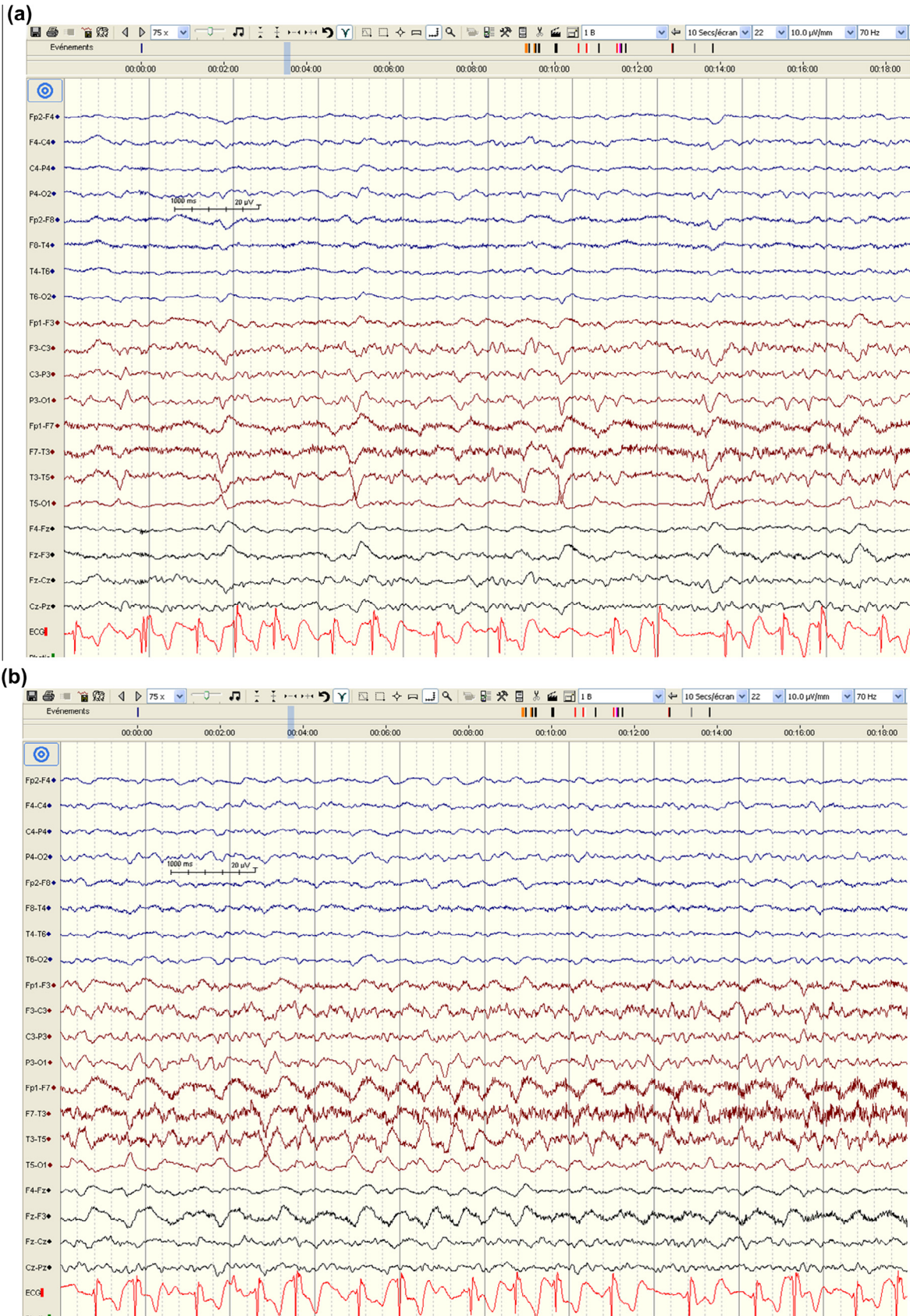
Fig. 1. (A, B): two consecutive screenshots of a patient with pseudo-periodic discharges on the left posterior region, evolving to a semi-rhythmic activity over the left temporal region. Longitudinal bipolar montage (right over left), 30 mm/sec, 10 µV/mm.

Table 1

Results regarding the positive percent agreement (PPA) and false positive ratio (FPR) per hour in each group and in the whole studied cohort. PPA 1: Median PPA using only the automated seizure detection (ASD) function. PPA 2: Median PPA including ASD function and spectrograms. FPR 1 (per hour): Median false positive rate using the ASD function only. FPR 2 (per hour): Median false positive rate through both methods. FPR* relates to the sum of false positive across each groups, divided by the total recording time in the group. Results are formulated in terms of median (range), interquartile range (IQR), except in the last column.

	PPA 1 median (range)IQR	PPA 2 median (range)IQR	FPR 1 median (range)IQR	FPR 2 median (range)IQR	FPR*
Group A(periodic patterns and seizures)	75 (0–100)4.8–100	75 (0–100)50–100	0 (0–6.1)0–0	0 (0–17.1)0–0	0.76
Group B(periodic patterns only)			0 (0–27.0)0–0	6.06 (0–75.8)3–9.09	4.28
Group C(seizures only)	100 (0–100)13.33–100	100 (0–100)50–100	0 (0–6.1)0–0	0 (0–28.3)0–3	0.06
Group D(no seizures, no periodic patterns)			0 (0–0)0–0	0 (0–9.1)0–2.9	0.0
Total	100 (0–100)5.9–100	100 (0–100)50–100	0 (0–27.0)0–0	0 (0–75.8)0–4.5	0.97
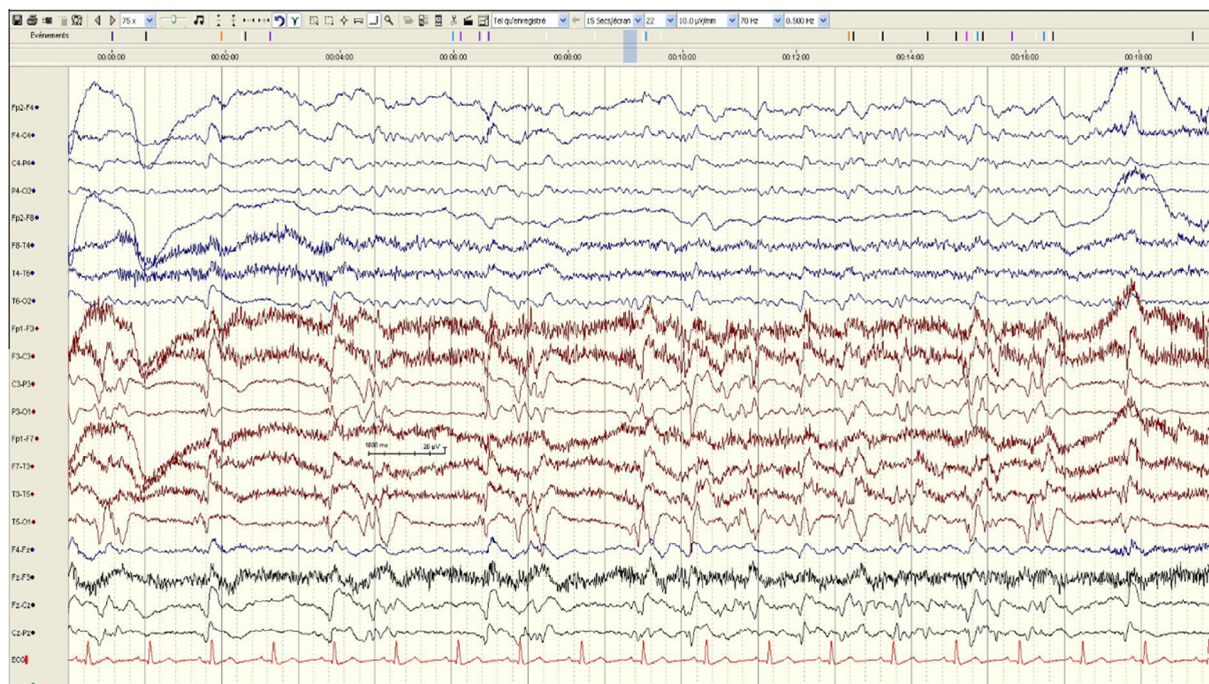
Fig. 2. A seizure from group A is detected by the visual analysis with the raw EEG (pseudo-periodic discharges on the left posterior region with intermixed faster rhythmic activity over less than 10 s), but is not marked on the automated detection. Longitudinal bipolar montage (right over left), 20 mm/s, 10 μV/mm.

recording was 4 (range 1–50); in group C the median was 3 (range 1–18). The total duration of all recordings was 82.7 hour, and the median recording time 1 hour (range: 20 minutes – 19 h).

Regarding the clinical characteristics of the patient population, median age was 57 [range: 3–88] years and 48 (49%) were women. The presumed etiology was structural-metabolic in 66 (67.3%), genetic in 7 (7.1%), and unknown in 25 patients (25.5%) (Berg et al., 2010).

Table 1 gives the overview of the results. Overall, the software detected 204/268 (=76.1%) of seizures. Mean detection rate per subject was 90.5% (median 100%, range 0–100; interquartile range 50–100). The 64 "undetected seizures" corresponded to subtle ictal patterns without clear evolution in frequency or amplitude in 30 (46.9%), short events (duration less than 10 s, Fig. 2) in 21 (32.8%), fast rhythms in 8 (12.5%) and to seizures masked by muscular artifacts in 5 (7.8%) cases. Considering the two groups presenting seizures (A and C), applying both detection methods (ASD function and spectrogram), all ictal events were captured in 29 out of 38 (76.3%) patients, whereas only in three subjects (7.7%) no seizures at all were detected. The median PPA in Group A (PLEDs plus intermixed seizures), was 75%; while in Group C (seizures only) was 100%, regardless of the methods.

Eighty false positive events were identified by the detector, overall. Across the entire dataset (82.7 h), this gives an FPR of 0.97/h; if two extreme outliers in group B were removed (with a very high FPR, see Fig. 3; these events occurred in EEGs showing monotonous, prolonged GPEDs in both), the falsely detected events dropped to 20, with an FPR of 0.24/h. Of note, the median FPR across all recordings was 0 (range 0–75.8; interquartile range 0–4.5). False-positive detections were concentrated in groups A and B, both with periodic patterns, often due to artifacts (Fig. 4), or prolonged GPEDs. FPR calculated by adding all false positive events across each group, divided by the total recording time of the group, are somewhat higher.

Finally, the combination of the ASD function and the spectrogram analysis increased the FPR, particularly in group B. In the control group (D), no FPs were detected using only the ASD function.

4. Discussion

In this study, the Persyst P12 software was evaluated in terms of PPA and FPR in a particularly difficult group of EEG patterns, including patients with periodic patterns (Group B), and with epileptiform-looking variants of normality or artifacts suggestive of ictal events (Group D). This software detected 76.1% of all seizures. The medians show high PPA and low FPR overall, but with lower
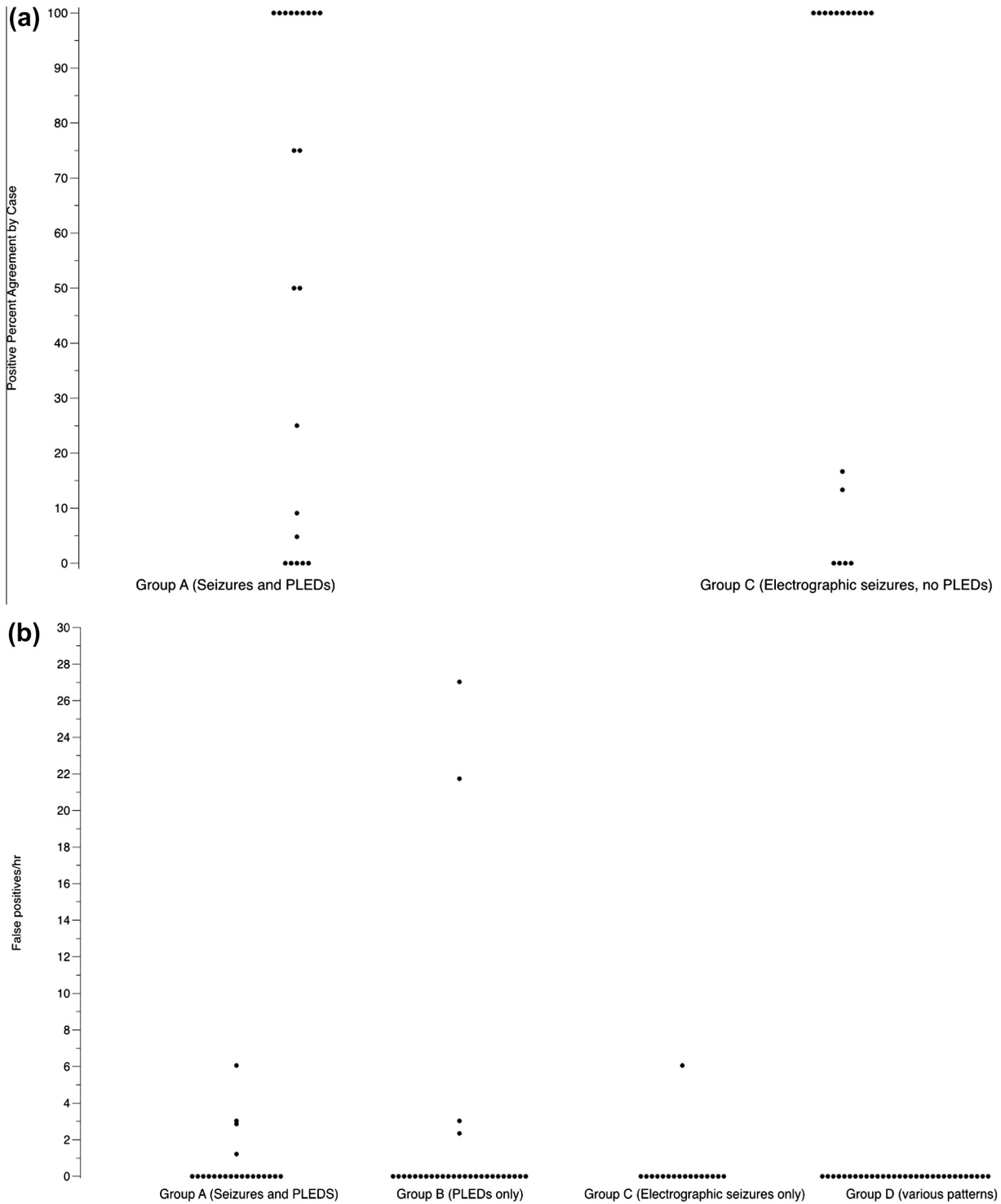
Fig. 3. (A, B): Graphs showing A (upper panel): positive percentage agreement (PPA) using only the automated seizure detection (ASD) function, and B (lower panel): false positive rates (FPR) by case for the ASD function.

performances in patients having periodic patterns, intermixed or not with seizures.

A variety of quantitative EEG analyses and display techniques have been proposed to ease EEG interpretation. Review of quantitative displays usually quickly reveals an irregular structure, abrupt phase changes or distortions, appearance or increasing of focal slowing, generalized suppression, loss of faster frequencies, or increasing or decreasing EEG variability, suggesting an ictal event. However, the distinction between seizures from

artifact-related changes, variants of normality or repetitive epileptiform elements may remain very challenging.

Since the automatic EEG detection system described by Hjorth more than forty years ago (Hjorth, 1970), different algorithms have been tested regarding their yield in identifying epileptic seizures, and their results show a wide variability, which relates not only to the properties inherent to each software, but also to the types of EEG that were used for the validations. The "older" generation (Gotman, 1982, 1990; Pauri et al., 1992) showed results in terms
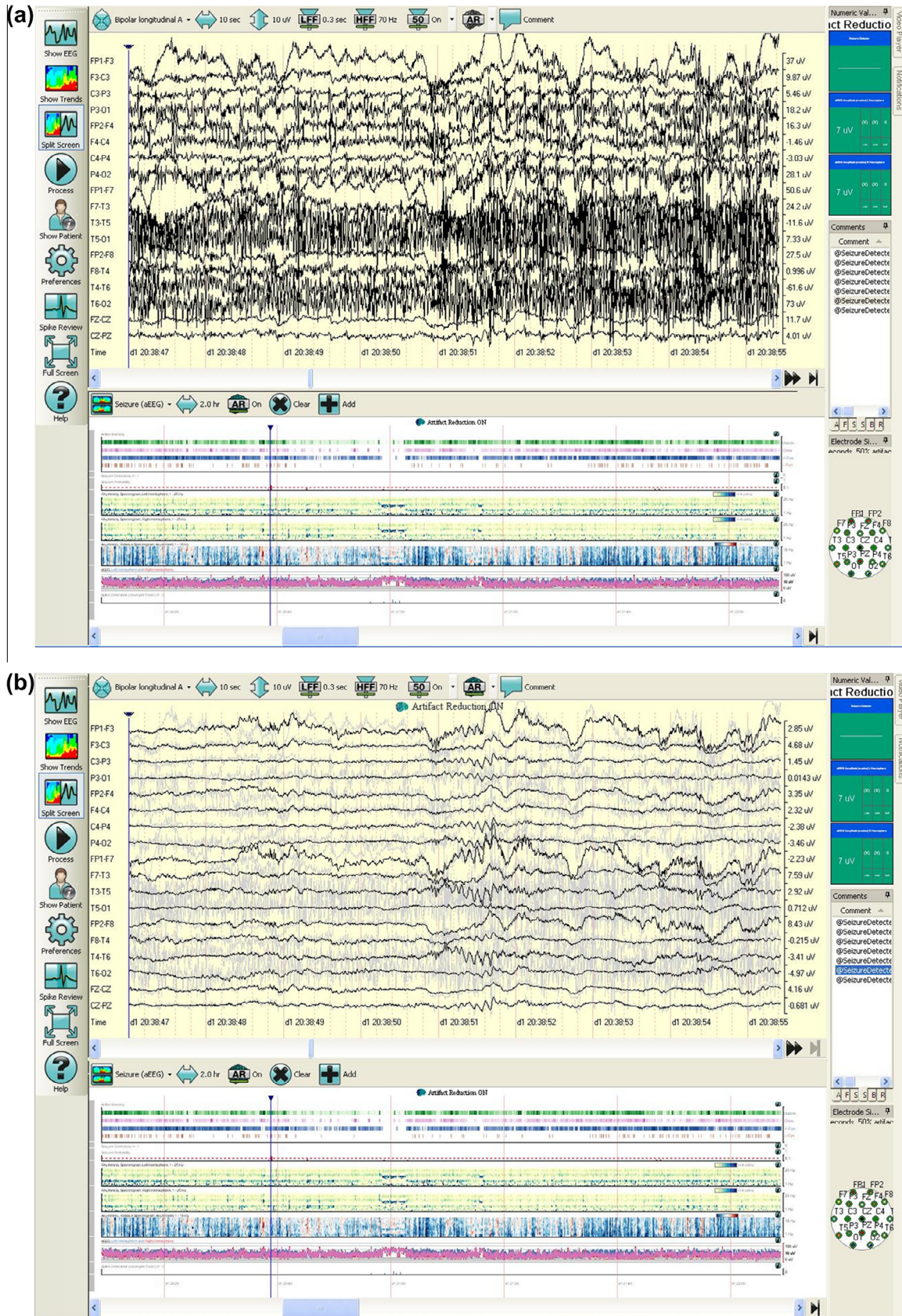
Fig. 4. (A, B): A seizure is "incorrectly" marked on the automated detection (vertical bar, bottom screen), while the raw EEG trace seems to show essentially artifacts. B: After activation of the artifact reduction function, no seizure is indeed observed (same position as Fig. 3A).

of sensitivity from 0.43 to 0.64, with FPR up to 3.3/h for Monitor 3.0 (Pauri et al., 1992). With software such as CNet and Monitor 8.0c, sensitivities of 0.93 and 0.74, and FPR of 1.4/ h and 3.0/h, respectively, were reported (Gabor, 1998). Otherwise, the Reveal system demonstrated a sensitivity between 0.74 and 0.8, with an FPR of 0.1–0.2/h only (Wilson et al., 2004). Using a neural network method for automatic and incremental learning applied to a patient-specific seizure detection (probabilistic neural network), a sensitivity of 0.89 and a FPR of 0.56/h were obtained in a small group of epilepsy monitoring and intensive care unit patients (Wilson, 2005). With the Stellate Harmonie system, the sensitivity was 0.76 and the FPR 0.34/h, with a median detection delay of 10 s, operating as an on-line seizure detection (Saab and Gotman, 2005). The overall detection sensitivity of IdentEvent in epilepsy monitoring unit patients was 0.80 with a very low FPR of 0.09/h (Kelly et al., 2010). Detection of seizures with the EpiScan has proved to achieve an overall sensitivity ranging from 0.73 to 0.83, with a FPR of 0.30/h (Hartmann et al., 2011; Furbass et al., 2012). Finally, with a multimorphologic ictal-pattern recognition system, the average correct detection rate was higher than 96%, with a mean false alarm rate of 0.25/h (Meier et al., 2008). It is of note that most of the studies of automated seizure detection algorithms were conducted primarily using recordings obtained on patients in the epilepsy monitoring units or EEG laboratories, and assessed very little data concerning seizures recorded in ICU settings. When the Ident-Event and Reveal algorithms were used to assess a small number ($N = 11$) of ICU recordings containing seizures, their detection sensitivities fell to 10.1% and 12.9%, respectively (Sackellares et al., 2011).

As compared to the aforementioned data, in the present study, including a large percentage of ICU patients with complex EEG patterns, we found a lower PPA and higher FPR in the subgroups with periodic patterns; this is probably related to the selection of recordings with particularly challenging EEGs. The Persyst 12 seizure detection algorithm functions as a general purpose seizure detector, and was not specifically designed to address ICU EEG patterns that can sometimes possess features bordering on electrographic ictal activity without clearly being seizures. Nevertheless, the characteristics of the detector showed good performances overall. The majority of missed seizures presented with a short duration (below 10 s), subtle diffuse electrodecrements, without any obvious evolution (Fig. 2). Conversely, false detections corresponded to muscle artifacts or electrode failures, or to prolonged GPEDs. It is important to underscore that overall only less than 8% of patients having at least one seizure did not get any detected.

This study has some limitations; first, the assessment between the two methods was not blinded. Second, the definition seizures, especially in patients having additional PLEDs (group A) did not rely on robust evidence-based criteria, but was somewhat subjective, integrating also the clinical situation, as it is common in clinical practice; for example, a patient with focal periodic discharges over the dominant temporal lobe and an intermixed acceleration-deceleration of the electrogenesis is generally regarded to be in status epilepticus and deserves a consequent treatment (Sutter and Kaplan, 2012). Third, the evaluation of spectrograms includes some subjectivity, as opposed to the ASD function. Finally, we analyzed a selected group of EEGs, therefore generalizability is not implicit.

In view of these results and the above mentioned literature, current computer-assisted EEG interpretation techniques may be considered an important help for the human observer (Scheuer, 2002). However, some limitations should be outlined: in the presence of periodic patterns, false negative and false positive results may become more frequent; conversely, the use of the rhythmicity and power spectrograms in addition to the ASD increases the

PPA, but also the FPR. It appears therefore important that EEG automated detections, especially in long-term monitoring, should be critically assessed for each specific patient, particularly at the beginning of the recording (Hartmann et al., 2011). Nevertheless, it is likely that ongoing refinements in seizure detection algorithms and EEG trending software will enhance their utility and lead to even more extensive applications.

This analysis provides data regarding the performance of an automated seizure detection algorithm in a difficult subset of EEG recordings, particularly including periodic patterns. This may help clinicians in the application of automated EEG algorithms, and provide developers with valuable data for possible improvements.

References

Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, van Emde Boas W, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009. Epilepsia 2010;51:676–685.

Claassen J, Mayer SA, Kowalski RG, Emerson RG, Hirsch LJ. Detection of electrographic seizures with continuous EEG monitoring in critically ill patients. Neurology 2004;62:1743–8.

DeLorenzo RJ, Waterhouse EJ, Towne AR, Boggs JG, Ko D, DeLorenzo GA, et al. Persistent nonconvulsive status epilepticus after the control of convulsive status epilepticus. Epilepsia 1998;39:833–40.

Furbass F, Hartmann M, Perko H, Skupch A, Dollfuss P, Gritsch G, et al. Combining time series and frequency domain analysis for a automatic seizure detection. In: Conference proceedings : annual international conference of the ieee engineering in medicine and biology society IEEE Engineering in Medicine and Biology Society Conference 2012;1020–1023.

Gabor AJ. Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies. Electroencephalogr Clin Neurophysiol 1998;107:27–32.

Gotman J. Automatic recognition of epileptic seizures in the EEG. Electroencephalogr Clin Neurophysiol 1982;54:530–40.

Gotman J. Automatic seizure detection: improvements and evaluation. Electroencephalogr Clin Neurophysiol 1990;76:317–24.

Hartmann MM, Furbass F, Perko H, Skupch A, Lackmayer K, Baumgartner C, et al. EpiScan: online seizure detection for epilepsy monitoring units. Conference proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference 2011:6096–6099.

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.

Hjorth B. EEG analysis based on time domain properties. Electroencephalogr Clin Neurophysiol 1970;29:306–10.

Kelly KM, Shiau DS, Kern RT, Chien JH, Yang MC, Yandora KA, et al. Assessment of a scalp EEG-based automated seizure detection system. Clin Neurophysiol 2010;121:1832–43.

Meier R, Dittrich H, Schulze-Bonhage A, Aertsen A. Detecting epileptic seizures in long-term human EEG: a new approach to automatic online and real-time detection and classification of polymorphic seizure patterns. J Clin Neurophysiol 2008;25:119–31.

Pauri F, Pierelli F, Chatrian GE, Erdly WW. Long-term EEG-video-audio monitoring: computer detection of focal EEG seizure patterns. Electroencephalogr Clin Neurophysiol 1992;82:1–9.

Privitera M, Hoffman M, Moore JL, Jester D. EEG detection of nontonic-clonic status epilepticus in patients with altered consciousness. Epilepsy Res 1994;18:155–66.

Saab ME, Gotman J. A system to detect the onset of epileptic seizures in scalp EEG. Clin Neurophysiol 2005;116:427–42.

Sackellares JC, Shiau DS, 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):69–73.

Scheuer ML. Continuous EEG monitoring in the intensive care unit. Epilepsia 2002;43(Suppl 3):114–27.

Sutter R, Kaplan PW. Electroencephalographic criteria for nonconvulsive status epilepticus: synopsis and comprehensive survey. Epilepsia 2012;53(Suppl 3):1–51.

Towne AR, Waterhouse EJ, Boggs JG, Garnett LK, Brown AJ, Smith JR, et al. Prevalence of nonconvulsive status epilepticus in comatose patients. Neurology 2000;54:340–5.

Wilson SB. A neural network method for automatic and incremental learning applied to patient-dependent seizure detection. Clin Neurophysiol 2005;116:1785–95.

Wilson SB, Scheuer ML, Emerson RG, Gabor AJ. Seizure detection: evaluation of the Reveal algorithm. Clin Neurophysiol 2004;115:2280–91.