TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

EEG Monitoring based on Automatic Detection of Seizures and Repetitive Discharges

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

Ao. Univ.Prof. Eugenijus Kaniusas

E354 Institute of Electrodynamics, Microwave and Circuit Engineering

eingereicht an der Technischen Universität Wien,
Fakultät für Elektrotechnik und Informationstechnik
von

Dipl. Ing. Franz Fürbass

Martrikelnummer: 9830935
Satzingerweg 49/5/11, A-1210 Wien

Contents

List of abbreviations

ACNS, American Clinical Neurophysiology Society

AEEG, amplitude integrated EEG

BSP, burst suppression pattern

ECG, Electrocardiography

EEG, Electroencephalography

EMG, Electromyography

EMU, epilepsy monitoring unit

EWS, Epileptiform wave sequence

FN, false negative

FP, false positive

GTCS, generalized tonic-clonic seizure

ICH, spontaneous intracerebral hemorrhage

ICU, intensive care unit

ITC, Ictal tachycardia

ILAE, International league against epilepsy

IRA, inter-rater agreement

PNES, Psychogenic nonepileptic seizures

PPG, Photoplethysmography

PWA, Periodic waveform analysis

SCORE, Standardized computer-based organized reporting of EEG

TN, true negative

TP, true positive

Abstract

This work deals with computer aided monitoring of electrical brain signals to detect disease-related patterns. Severe neurological disorders can trigger unusually strong firing of brain cells that distinguishes clearly from normal brain activity. A well-known example of such a disease is epilepsy. During an epileptic seizure, fast repeating electrical discharges on the head surface are often measurable. Epileptic seizures usually occur rarely or unnoticed by patients during night wherefore considerable effort is needed to properly evaluate and treat patients. Besides epilepsy, inflammatory brain diseases such as encephalopathies or traumatic brain injuries can trigger different types of patterns with repetitive discharges and seizures. The severity of these diseases and injuries often require intensive medical treatment and continuous monitoring of neurological activity. The automatic detection of epileptic seizures and repetitive patterns in measurements of electrical brain signals is a central part of this work.

Currently, diagnostic in neurological patients involves a wide spectrum of methods and tools. In addition to clinical observations the objective quantification of the brain status is the primary step in diagnosis. Imaging methods such as magnetic resonance imaging (MRI) are able to generate a snapshot of the brain morphology. To evaluate the brain activity over time the electroencephalography (EEG) is used. The EEG is able to continuously record the electrical activity of the cortex which makes this method to a central element in the diagnosis of neurological patients. Manual evaluation of EEG is done by visual inspection of a graphical representation of these signals. Specially trained medical staff interpret the EEG in 10 to 15 second sections throughout the whole recording period. With a usual recording time of 4 days more than 34,000 EEG pages need to be evaluated, which requires a considerably amount of time for EEG interpretation. The correct evaluation of the EEG curves requires a high degree of experience in order to avoid misinterpretations. Despite intensive training, different interpretations of an EEG by different reviewers and differences in the evaluation of an EEG by the same reviewer at different time points are fundamental problems. Further, distortions from electrical signals of muscles (electromyography, EMG) can overlay with the EEG, which may lead to misinterpretations. In order to make medical findings comparable in the context of international studies, standardized descriptions of EEG patterns have been developed which are, however, require experienced staff to be assigned correctly. If the EEG is used for real-time monitoring of patients with serious illnesses, the evaluation of the EEGs must be carried out promptly in order to achieve an improved treatment. However, this places extremely high demands on the staff. Many of the mentioned problems in manual EEG analysis can be addressed by using computer-aided evaluation methods. A computer algorithm is able to reduce the cost and time of analysis and provides perfect repeatability of the result. Despite these obvious advantages, correct automatic analysis of the highly complex EEG signal is an unsolved problem that prevents the widespread use of such computer algorithms.

In this work novel computer algorithms for the automatic interpretation and monitoring of EEG signals are presented that were published in eight papers of highly-ranked peer reviewed journals. The overall aim is to make EEG evaluation considerably easier by automatically marking important time points in real-time. The focus is on the detection of epileptic seizures and patterns with repetitive discharges. Although the EEG is the primary data source for the algorithms, EMG interferences have to be treated adequately in order to achieve the highest possible precision. To raise sensitivity of the automatic seizure detection algorithm even further the electrocardiography (ECG) signal was additionally evaluated to find seizure related activity. The use of computer algorithms for real-time monitoring of EEG activity is intended to improve treatment of patients and to increase patient safety.

A fundamentally new approach for the detection of EEG discharges was developed in this work that can be applied to a wide range of pathological patterns. By combining individual discharges into groups that are extended spatially and over time, different types of patterns are modelled. Important measures such as frequency and amplitude can then be found by simply averaging the group elements. Furthermore, the temporal progression of patterns is used to quantify changes. The time-domain algorithm therefore creates the basis for analysis of seizures and other EEG patterns. The

classification algorithms that utilize this information then allow the detection of seizures as well as the quantification of EEG patterns in intensive care patients. The results of the computer algorithms can be read and interpreted efficiently by means of a newly developed graphical visualization.

The clinical validation of computer algorithms is an essential part of this work. The quality of the algorithms can only be determined with statistical significance by diagnostic studies including a high number of patients. Results of the algorithms were compared to manual annotations from experts to measure sensitivity and specificity. In this work, four multi-center studies and some smaller preliminary investigations were carried out for different medical questions and algorithms. In total, EEGs of 621 patients from 6 centers in Europe and the USA were used for validation. The results show that seizure alarming is possible with a sensitivity of 81% and a false alarm rate of 7 false alarms per day. A time delay of only 3 seconds was measured from the seizure pattern to the alarm. In the detection of seizures based on existing EEG files, the algorithms achieved a high sensitivity of 86% which is required for efficient evaluation. Special epilepsy types such as temporal lobe epilepsy showed a sensitivity of 94%. The detection of different patterns in the EEG of intensive care patients yielded in sensitivities between 85% and 93% and specificities in the range of 90% and 96%. Improved treatment of patients as well as a reduction in workload for medical staff are thus possible.

In the future, mobile EEG systems in the outpatient setting can represent a further significant improvement in diagnostic. Patients with rarely occurring seizures can save themselves from protracted hospital stays. Moreover, the use of mobile sleep diagnostic including EEG can increase the quality of life and save costs. At present, mobile EEG systems are still suffering from problems such as high time expenditure for the attachment of the electrodes and complex wiring. This results in low patient acceptance and prevents their widespread use. As soon as these difficulties are solved computer algorithms can be utilized to evaluate such mobile EEG systems. A large number of medical applications are then conceivable in which the quality of the algorithms will play a central role.

Kurzfassung

Diese Arbeit beschäftigt sich mit der computerunterstützen Überwachung von krankheitsbedingten Mustern in elektrischen Hirnsignalen. Schwerwiegende neurologische Krankheiten können ein ungewöhnlich starkes Feuern von Hirnzellen auslösen, welches eindeutig von normaler Aktivität unterscheidbar ist. Das bekannteste Beispiel einer solchen Krankheit ist die Epilepsie. Während eines epileptischen Anfalls sind oft schnell wiederholende elektrische Entladungen auf der Kopfoberfläche messbar. Gewöhnlich treten epileptische Anfälle sehr selten, oder auch vom Patienten unbemerkt in der Nacht auf, weshalb der Aufwand bei der Diagnose und Behandlung erheblich ist. Neben der Epilepsie, können auch entzündliche Hirnerkrankungen wie Enzephalopathien oder Schädel-Hirn-Traumata unterschiedliche repetitive Entladungen und Anfälle auslösen. Die Schwere dieser Krankheiten und Verletzungen erfordert oft eine intensivmedizinische Behandlung und eine laufende Überwachung der neurologischen Aktivität. Das automatische Auffinden von epileptischen Anfällen und repetitiven Entladungen in Messungen von elektrischen Hirnsignalen bildet den Schwerpunkt dieser Arbeit.

Zurzeit wird eine Vielzahl von Methoden und Werkzeugen in der Diagnostik von neurologischen Patienten eingesetzt. Neben bildgebenden Verfahren wie der Magnetresonanztomographie (MRT), ist die Elektroenzephalografie (EEG) ein zentraler Bestandteil in der Diagnostik von neurologischen Patienten. Im Gegensatz zur MRT, das eine Momentaufnahme der Morphologie des Gehirns darstellt, kann das EEG die elektrische Aktivität im Kortex kontinuierlich aufzeichnen. Die Auswertung der oft über mehrere Tage dauernden Aufzeichnungen erfolgt durch visuelle Begutachtung einer graphischen Darstellung dieser Signale. Speziell ausgebildetes medizinisches Personal interpretiert das EEG in 10 bis 15 Sekunden langen Abschnitten. Bei einer üblichen Aufnahmedauer von 4 Tagen fallen so über 34000 EEG Seiten für die Auswertung an, wodurch der zeitliche Aufwand der EEG Interpretation erheblich wird. Die korrekte Beurteilung der EEG Kurven erfordert ein hohes Maß an Erfahrung um Fehlinterpretationen zu vermeiden. Trotz intensiver Ausbildung stellen unterschiedliche Auslegungen eines EEGs von verschiedenen Begutachtern, sowie Unterschiede in der Bewertung eines EEGs vom selben Begutachter zu unterschiedlichen Zeitpunkten grundsätzliche Probleme dar. Auch können elektrische Signale von Muskeln (Elektromyographie, EMG) das EEG überlagern, wodurch es zu Fehlinterpretationen kommen kann. Um medizinische Erkenntnisse im Kontext von internationalen Studien vergleichbar zu machen, wurden standardisierte Beschreibungen von EEG Mustern entwickelt die sich allerdings wieder auf eine manuelle Auslegung stützen. Wird das EEG zur Echtzeit-Überwachung von Patienten mit schwerwiegenden Erkrankungen verwendet, muss die Interpretation des EEGs zeitnah erfolgen um eine verbesserte Behandlung zu erzielen. Allerdings werden dadurch extrem hohe Anforderungen an das Personal gestellt. Viele der genannten Probleme in der manuellen EEG Analyse können durch den Einsatz von computerunterstützten Auswerteverfahren adressiert werden. Ein Computeralgorithmus kann Kosten und Zeit der Analyse reduzieren und schafft eine perfekte Wiederholbarkeit des Ergebnisses. Trotz dieser offensichtlichen Vorteile stellt die korrekte Analyse des hochkomplexen EEG Signals ein ungelöstes Problem dar, das den weit verbreiteten Einsatz von Computeralgorithmen verhindert.

In dieser Arbeit werden neuartige Computeralgorithmen zur automatischen Interpretation und Überwachung von EEG-Ableitungen vorgestellt, die zuvor in acht Beiträgen von angesehenen wissenschaftlichen Zeitschriften veröffentlicht wurden. Ziel ist eine wesentliche Arbeitserleichterung in der EEG Auswertung durch automatische Markierung wichtiger Stellen in Echtzeit. Das Hauptaugenmerk liegt dabei auf der Detektion von epileptischen Anfällen und Mustern mit repetitiven Entladungen. Obwohl das EEG die primäre Datenquelle für die Algorithmen darstellt, müssen auch Überlagerungen von EMG adäquat behandelt werden um eine möglichst hohe Präzision zu erreichen. Um die Sensitivität der automatischen Anfallsdetektion noch weiter zu steigern wurde zusätzlich das elektrische Signal vom Herzen (Elektrokardiographie, EKG) ausgewertet, dass sich im Anfall ebenfalls stark verändert. Der Einsatz von Algorithmen für die Echtzeit-Überwachung von EEG-Aktivität soll die Behandlung von Patienten verbessern und die Patientensicherheit erhöhen.

Der in dieser Arbeit entwickelte und grundlegend neue Ansatz zur Auffindung von EEG Entladungen kann ganz allgemein auf eine Vielzahl von pathologischen Mustern angewandt werden. Durch Zusammenfassung von Einzelentladungen zu zeitlich und räumlich ausgedehnten Gruppen werden unterschiedliche Mustertypen modelliert. Wichtige Messwerte wie Frequenz und Amplitude können dann durch einfache Mittelwertbildung der Gruppenelemente gefunden werden. Weiteres wird der zeitliche Verlauf von Mustern verwendet um Änderungen zu quantifizieren. Dadurch wird die Basis für die Analyse von Anfällen und anderer Muster im EEG geschaffen. Die auf all diesen Informationen basierenden Klassifikations-Algorithmen erlauben dann neben der Detektion von Anfällen auch die Quantifizierung von EEG Mustern bei Intensivpatienten. Durch eine neu entwickelte graphische Visualisierung können die Ergebnisse der Computeralgorithmen effizient abgelesen und interpretiert werden.

Die klinische Validierung der Computeralgorithmen ist ein wesentlicher Teil dieser Arbeit. Erst durch Diagnosestudien die eine hohe Anzahl von Patienten inkludieren, kann die Güte der Algorithmen mit statistischer Signifikanz festgestellt werden. Dazu wurden die Ergebnisse der Algorithmen mit manuellen Bewertungen von Experten verglichen und Sensitivität sowie Spezifität festgestellt. Im Rahmen der Arbeit wurden für die unterschiedlichen medizinischen Fragestellungen und Algorithmen vier multizentrische Studien sowie weitere kleinere Voruntersuchungen durchgeführt. Insgesamt wurden für die Validierung EEGs von 621 Patienten aus 6 Zentren in Europa und den USA verwendet. Die Ergebnisse zeigen, dass die automatische Alarmierung bei epileptischen Anfällen mit nur 3 Sekunden Verzögerung zum Anfallsmuster mit einer Sensitivität von 81% und einer Fehlalarmrate von 7 Fehlalarmen pro Tag möglich ist. Auch bei der Detektion von Anfällen in bereits vorliegenden EEGs wurde eine hohe Sensitivität von 86% erreicht, die für eine effiziente Auswertung benötigt wird. Spezielle Epilepsie-Arten wie Temporallappenepilepsie erreichten hier eine Sensitivität von 94%. Die Detektion von unterschiedlichen Mustern im EEG von Intensivpatienten zeigte eine Sensitivität zwischen 85% und 93% und eine Spezifität im Bereich von 90% und 96%. Durch die Ergebnisse dieser Arbeit wurde gezeigt, dass der Einsatz von automatischen Systemen zur Überwachung von EEG Aktivität mit hoher Genauigkeit möglich ist. Eine verbesserte Behandlung von Patienten sowie eine Arbeitserleichterung für medizinisches Personal werden somit möglich.

Zukünftig kann der mobile Einsatz von Langzeit-EEGs außerhalb des Krankenhauses eine weitere wesentliche Verbesserung in der Diagnostik darstellen. Viele Patienten mit selten auftretenden Epilepsien können sich dadurch langwierige Krankenausaufenthalte ersparen. Auch der Einsatz von mobiler Schlafdiagnostik mittels EEG kann die Lebensqualität erhöhen und Kosten sparen. Derzeit leiden mobile EEG Systeme noch an Problemen wie hohem Zeitaufwand für die Anbringung der Elektroden und aufwendiger Verkabelung. Die dadurch geringe Akzeptanz bei Patienten verhindert deren weit verbreiteten Einsatz. Werden diese Schwierigkeiten behoben, können Computeralgorithmen für die Auswertung dieser Systeme eingesetzt werden. Eine Vielzahl von medizinischen Anwendungen sind dann denkbar bei denen die Qualität der Algorithmen eine zentrale Rolle spielen wird.

Chapter 1
Introduction

This work will present computer algorithms for automatic assessment of electroencephalographic signals (EEG). Patients with epilepsy and patients of the neurological intensive care unit (ICU) are commonly diagnosed with EEG and are in the focus of this work. Some general concepts and definitions will be described in this chapter that are essential to understand the problems and requirements of diagnosis with EEG.

1.1 Diagnosis of epilepsy patients

An epileptic seizure is a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain [1]. Recurrent and unprovoked seizures define epilepsy, a severe disease which about one percent of the world's population suffers from. The exact definition is given by Fisher et al, 2005: "Epilepsy is a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures and by the neurobiologic, cognitive, psychological, and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure" [1]. While 63 to 70% of epilepsy patients can become seizure free using antiepileptic drugs, the remaining are most difficult to treat and therefore suffer from medically refractory epilepsy [2], [3]. In 2017 the international league against epilepsy (ILAE) defined three fundamental seizure types based on the origin of their onset [4]. Seizures either begin focally in one hemisphere of the brain, have a generalized onset appearing from both hemispheres or are of unknown onset [4].

Assessment of clinical symptoms and anamnesis of patients is the primary source to diagnose epilepsy. Because the rate of misdiagnosis in epilepsy is as high as 23% evaluation of cerebral imaging and electrophysiological measurements is highly advisable [5]. Magnetic resonance imaging (MRI) shows a snapshot of the brain morphology at a certain time point and is used to find structural deviations. Besides MRI the most important measurement for diagnosis of epilepsy is the EEG. The EEG picks up electrical activity of postsynaptic potentials from large groups of pyramidal cells in the cortex. In normal brain activity cells are activated asynchronous which generate only small electrical potentials on the scalp (Figure 1, A). To measure increased amplitudes by electrodes on the scalp a few square centimetres of cortex have to be activated synchronously. Especially in pathologic conditions such large groups of cells are activated simultaneously and therefore show high EEG amplitudes. The time course of pathologic EEG activity is often abruptly and discontinuous but can also show sinusoidal activation patterns (Figure 1, B).

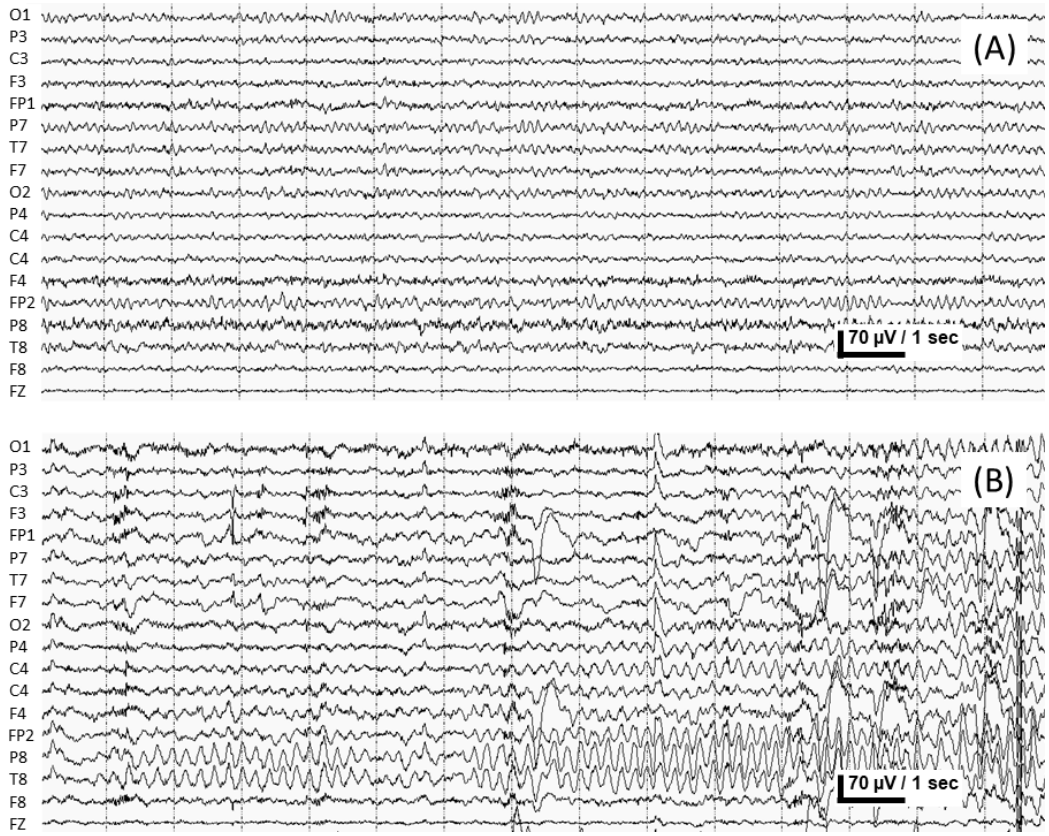

Figure 1: Normal EEG compared to EEG including a seizure: (A) Normal EEG of 15 seconds; (B) EEG of the same patient having an epileptic seizure visible as rhythmic activity starting on electrodes P8 and T8.

EEG is a non-invasive diagnostic technique that allows continuous recording of several days. The signal provides high time resolutions in milliseconds but only average spatial resolution. Recording durations of several days are required to catch clinically important seizure activity for assessment of spatial origin and occurrence frequency. In principle two types of events are of interest: events that are seizures, called 'ictal' and events in between seizures called 'interictal'. Typical interictal events are epileptic spikes (short term sharp transients lasting 40-200ms) that occur with high prevalence during night especially in slow wave sleep [6]. Ictal events like the seizure shown in Figure 1B may occur very seldom. The average seizure rate varies greatly between different epilepsy syndromes. Patients with an epileptic focus in the frontal lobe may have several seizures in a single day but other epilepsies may show only a few seizures per year. Epileptic seizures typically show quasi periodic activity in the EEG with frequencies of 4 to 12 Hz. The medical term for such irregular but repetitive EEG activity is 'rhythmic pattern'. Commonly evolution in frequency and amplitude of rhythmic patterns can be observed during seizures [7]. Signal distortions from muscle activity, body movements, or from electrodes with high impedance (> 10 kΩ) can overlay with the EEG. These signal artefacts are a common source of misinterpretation and overtreatment of patients [8].

Seizure activity with rhythmic waveforms is typically observed in patients having mesial temporal lobe epilepsy (Figure 1B). But such 'rhythmic patterns' include a variety of repeating transients that are only loosely related to a signal that is described in technical terms as periodic signal. Figure 2 compares two examples of rhythmic seizure activity by showing a single channel signal and the corresponding power spectrum. In EEG terminology, both have the same level of 'rhythmicity' but the technical evaluation using a power spectrum estimation shows major differences between EEG signals (A) and (C). Although signal (A) seems much more distorted by high frequency noise the pattern exhibits a concentrated power spectrum at about 4 Hz (B). Signal (C) includes EEG activity with irregular gaps between discharges that shows equally power levels in a broad frequency band. This makes such waveform especially hard to analyse with standard signal analysis methods.

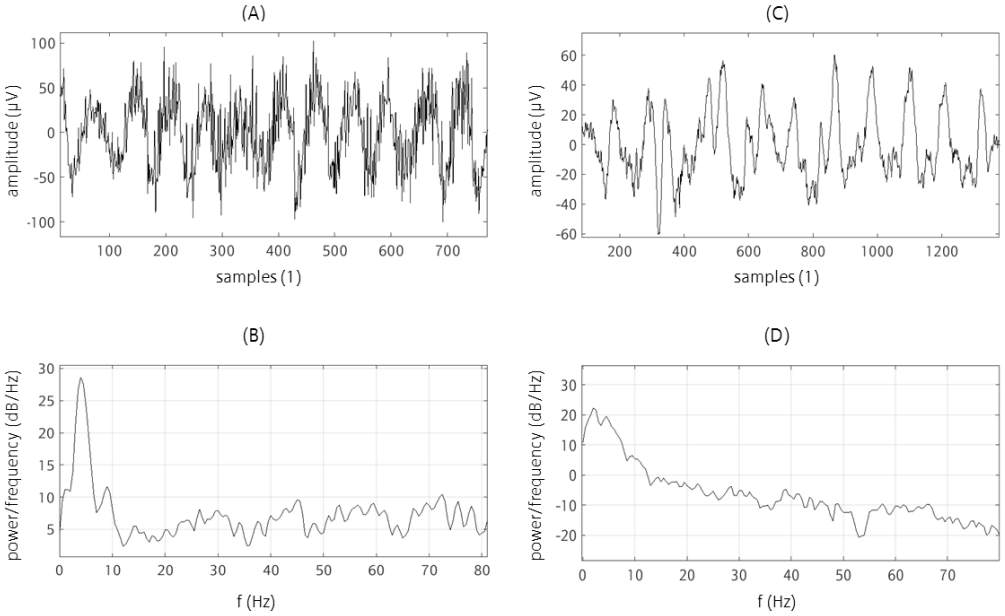

Figure 2: Two examples of rhythmic EEG: Ictal EEG with rhythmic morphology and with muscle artifacts (A) compared to the power spectrum of the same signal (B). In contrast, an ictal EEG with irregular morphology (C) leads to a fuzzy spectrum after transformation in the frequency domain (D). The smeared power spectrum has equally high components from 2 to 5 Hz.

Different reasons for EEG assessment exist. These clinical indications differ in their medical question to answer and require different equipment and recording durations. Common clinical indications for long-term EEG are differential diagnosis of epilepsy, psychogenic non-epileptic seizures (PNES), or other neurological diseases. Routine EEG of up to 30 minutes allows primary evaluation of brain activity in the ambulatory setting. Recording of such EEGs is often done using EEG caps that can be applied rapidly. Recording of seizures require much longer observation times and are therefore mostly done in stationary settings. Epilepsy monitoring units (EMUs) are hospital departments with specialized staff to diagnose or rule out epilepsy using long-term EEG and video monitoring. Evaluation of the recorded EEGs is mostly done on regularly basis with a time delay of several hours. But time delay of evaluation can be critical. Low latency of evaluation allows early treatment of seizures and neurological evaluation in the postictal phase. This increases the yield of the whole diagnostic by providing valuable clinical information of the seizure. Furthermore, safety of patients need to be addressed carefully. Immediate staff response is required in case of a seizure to avoid life threating conditions like post ictal apnoea. The percentage of time in which patients are observed continuously vary greatly between institutions. E.g. Atkinson et al. showed in a study including 20 patients that only 40% of seizures showed staff response [24].

In clinical praxis EEG signals are recorded by at least 21 electrodes on the scalp to ensure proper spatial resolution of cortical activity [7]. Commonly silver–silver chloride (Ag/AgCl) or gold disk electrodes are applied to standardized positions on the head with electrode impedances below 5 kΩ. The international 10-10 electrode placement system defines these standard positions for each sensor [16]. This allows interpretation of the signal location by channel names only. Figure 3 shows electrode positions of the international 10-10 electrode placement system. E.g. channel label C3 represents the EEG signal of the left motor area just above the central sulcus. The reduced 10-20 electrode placement system includes 21 electrodes in relative distances (10% or 20%) between the cranial landmarks over the head [16].

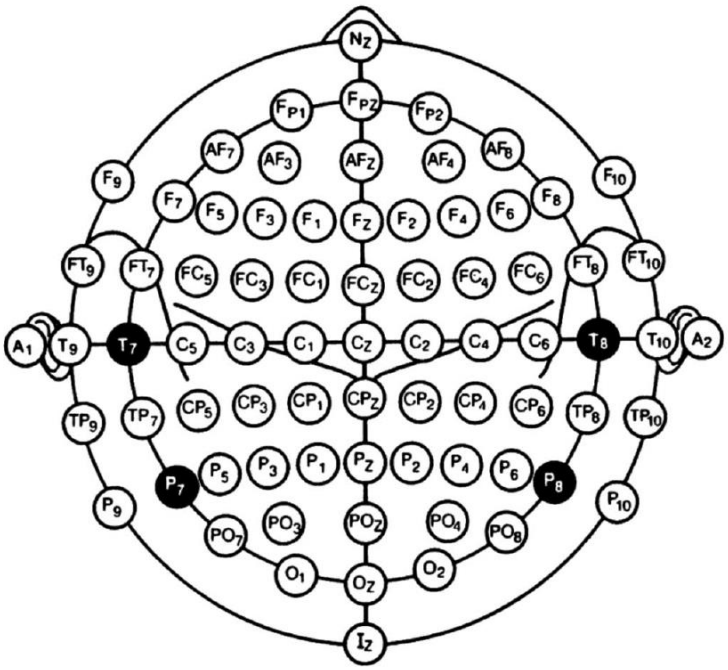

Figure 3: EEG electrode standard positions defined by the American Clinical Neurophysiology Society [16, p. 6].

Analysis of EEGs need to be repeated from medical staff with persistent high quality using standardized wording to describe findings. The SCORE initiative (Standardized Computer-based Organized Reporting of EEG) addresses the issue of standardized reporting on EEG by providing a predefined dictionary of terms to describe EEG findings [17]. By using SCORE unified EEG reporting is possible that will greatly facilitate clinical research. Despite all these efforts the issue of finding and interpreting EEG patterns of long term recordings remain one of the most time consuming and critical tasks.

1.2 Monitoring ICU patients with EEG

Recently, increasing attention has been paid to the treatment of seizures in critical care patients, which also triggered an increased use of continuous EEG. In patients of the ICU, epileptic seizures are often caused by traumatic brain injury, post-anoxic coma, or intracerebral haemorrhage. Because most patients are intubated and received muscle relaxants epileptic seizures are only visible in EEG without showing any or only subtle clinical signs (e.g. body movements). Seizures therefore often remain undetected and can further damage grey matter tissue by an excessive increase in metabolic demand of pyramidal cells. In a retrospective study including 102 patients with spontaneous intracerebral hemorrhage seizures occurred in 31% but over half were purely electrographic and only one patient had clinical seizures [18]. Seizures were associated with expanding haemorrhages, periodic discharges, and poor outcome [18]. These recent findings support the thesis that the number of seizures in critical care patients is underestimated.

Detection of seizures and EEG patterns require continuous EEG monitoring to reach high sensitivities. Claassen et al showed that 88% of the non-convulsive seizures of ICU patients require at least 24 hours of continuous EEG recording and that 12% of the patients needed more than 48 hours of recording to detect their first seizure [21]. To establish EEG for ICU monitoring in the first place, routine EEGs with short durations are used (e.g. < 30 minutes). Such diagnostic is available temporarily in most neurological ICUs. An analysis of 30 minute EEGs of 32 ICU patients showed that epileptiform discharges within these recording are predictive for the occurrence of seizures and rhythmic or periodic patterns of ictal–interictal uncertainty [22].

While patients with epilepsy mostly have clear-cut electrographic seizure patterns the situation is quite different for critical care patients. There, the epileptic seizure pattern itself is altered by medication and generally shows slower rhythmic EEG activity as well as discontinuous EEG patterns [19]. The Ictal – interictal continuum describes EEG patterns ranging from unequivocal seizure activity in the 4-8 Hz range to seizure related patterns with unclear significance [20]. The American Clinical Neurophysiology Society (ACNS) developed a research nomenclature for such EEG patterns of ICU patients with unclear significance to unify interpretation of these patterns and to establish the basis for research [9].

Automatic analysis of EEG from critical care patients is especially challenging. Commonly discontinuous EEG patterns can be observed which include periods of distinct low amplitude intervals. An example of such a pattern is the periodic pattern that consists of repetitive transients followed by low amplitude inter-discharge intervals. Another example of a discontinuous pattern is the burst suppression pattern (BSP). Burst suppressions are EEG patterns consisting of intermittent periods of very low voltage brain electrical activity (''suppression"), alternating in a quasi-periodic fashion with periods of higher amplitude activity (''bursts") [7], [9]. The major difference to periodic patterns is that discharges have more than three signal deflections and durations longer than 0.5 seconds. The close relationship between periodic and burst suppression patterns can be observed during rewarming of patients in hypothermia were burst suppression evolve into periodic patterns [7]. Figure 4 shows examples of periodic and burst suppression patterns.

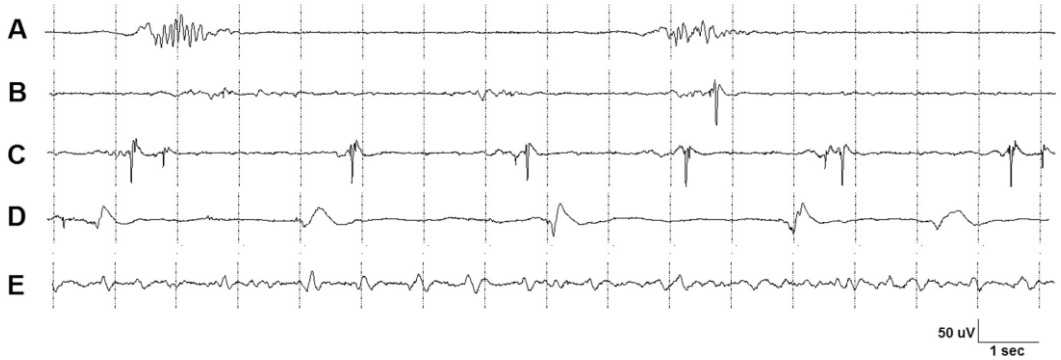

Figure 4: EEG examples of discontinuous patterns: EEG with discontinuous patterns is commonly observed in EEG of critical care patients. (A) Burst suppression pattern; (B) Burst suppression pattern with low amplitude bursts followed by a final burst including a high amplitude discharge; (C) Periodic pattern with surrounding low amplitude burst suppression pattern. (D) Periodic pattern with less than 3 phases in all discharges; (E) Periodic pattern with low amplitude discharges repeat with an inter-discharge interval of less than 1 second [A4].

BSPs are found in a wide range of pathological and clinically-induced conditions, including anesthetic-induced coma, hypothermia [10], [11] deep hypothermia [12], [13], or arising spontaneously as a result of anoxic brain injury [14], [15]. In addition, BSP can be found during myoclonic jerks in patients having hypoxic encephalopathies which are not directly epileptic seizures [7] but seizure are often develop in parallel to this severe functional brain disorder. In automatic computer analysis, such signal waveforms are problematic as no technical periodic signal components exist because of the random inter-discharge interval. Therefore, periodic and burst suppression patterns in general cannot be analysed with frequency domain methods like the Fourier transform.

1.3 State of the art in EEG processing

Starting in 1970 the length of EEG recordings continuously increases from 30 minutes up to one week for close inspection of patients with rare seizures [23]. But evaluation procedure continued to be tedious and time consuming for such lengthy recordings as pure visual inspection of raw signals on consecutive 10 second pages was state of the art at that time. Computer supported evaluation of EEG potentially reduces time spend to scan EEG segments without relevant information. To support this assumption, we conducted a small retrospective study on 14 randomly selected patients from an EMU that underwent video-EEG monitoring. Results showed that by using automatic EEG software it required only 3 minutes of time to find 83% of all seizures in 24 hours of recording on average [B3][1]. This shows that most of review time is spent to find time positions of epileptiform patterns without deducing valuable information. Driven by these existing issues of EEG diagnostic and with the ongoing wide spread use of continuous EEG automatic evaluation methods were developed.

1.3.1 EEG based seizure detection

One of the first algorithms for automatic seizure detection was published in 1982 by Jean Gotman. In that work, already 24 recordings were used for evaluation of detection performance [25]. From there on numerous approaches to automatic EEG evaluation and seizure detection were reported in the literature.

In principle two main use cases should be distinguished. Seizure alarming or online seizure detection algorithms evaluate the EEG data stream during recording. If a seizure was detected a notification is triggered to allow immediate staff response and neurological testing. Such alarming algorithms are trimmed to have short time delay and a low number of false alarms. As only past information of the patient can be used the task is considered as more challenging as to use all recorded EEG at once. In addition, short time delay does not allow extensive use of pipeline based signal processing and will therefore be suboptimal in calculation time. One of the first dedicated online detection algorithms already showed 100% sensitivity and a false alarm rate of 4 in 24 hours [26] on a number of selected recordings. By trimming online seizure detection algorithms to specific patients, the high inter-patient variability of different seizure types can be avoided. E.g. Zheng et al. showed that their machine learning based approach can reach a sensitivity of 92% while having an average false alarm rate of 4.08 in 24 hours [27]. From engineering perspective, such patient specific seizure detection algorithms and devices are highly problematic as each of these devices will need a separate clinical trial to be cleared as medical device. A common problem in neurological diagnostic of critical care patients is the unavailability of EEG equipment and of trained medical staff for EEG interpretation. The time delay of evaluation is especially critical in this setting as prolonged non-convulsive seizures need to be treated within short time delay to improve patient outcome [28], [29]. Therefore, evaluation has to be done ad-hoc or in regular short intervals.

Another use case of automatic EEG evaluation is supporting review of data after recording, called offline seizure detection or just seizure detection. There, automatic detection algorithms in general show higher sensitivity as online algorithms but also higher false detection rates. In clinical workflow such algorithms do not trigger alarms but are used to rapidly review data. The higher false detection rate is usually accepted as detections can be evaluated quickly. In this research area a high number of publications exists. In 2014 Jin et al. published an editorial that summarizes the most interesting approaches until that time [30] including [31]–[37]. Most recently Hopfengärter et all showed promising results mostly based on temporal lobe epilepsy patients in a large retrospective study [38].

[1] References [B1]...[B3] are scientific abstracts including major contributions of the author.

1.3.2 Non-EEG based seizure detection

In addition to electrical brain activity measured by EEG the electrocardiography (**ECG**) is another physiological signal expressing seizure activity. Seizures causing activation of the central autonomic network followed by changes in heart rhythm at seizure onset. Ictal tachycardia (ITC) represents the most frequent change in heart rhythm and can be observed in 65-86% of seizures [39], [40]. The advantage of seizure detection based on heart rate is the ease of measurement compared to EEG. A single ECG lead is sufficient to retrieve time information of QRS complexes. Further, the photoplethysmography (PPG) is a viable alternative to ECG based measurements resulting in similar timing information of heart beats by optically measuring changes in volume of blood vessels. Compared to evaluation of multi-channel EEG computational effort of ECG or PPG based algorithms is lower as computational complexity approximately rises linearly with the number of signal channels. Several heart beat based seizure detection algorithms were published in the past showing good results on patients having complex partial seizures [41], tonic and myoclonic seizures [42]. Osorio et al presented results of an ECG based seizure detection on 81 patients and showed that the delay of ictal changes in the ECG is short compared to intracranial EEG. The time delay between ictal EEG and ECG seizure onset ranged from 0.8 to 13.8 seconds [43]. Recently Jeppesen et al. showed that all seizures of 13 out of 17 patient could be detected using a ECG based algorithm [44] [45]. The disadvantage of heart beat based detection is the limited applicability as ITC is not present in all seizure types of epilepsy patients and is usually not observed in seizures of ICU patients. Another common problem of heart beat based detections are false detections triggered by normal physical activity.

Another electrical signal modality for detection of epileptic seizures is the electromyography (**EMG**) representing the electrical activity of muscle fibres. During convulsive seizures, pathologic muscle contraction can be observed that distinguishes from physiological contraction by its excessive strength which can even lead to torn muscle fibres. The commonly known generalized tonic-clonic seizure (GTCS) where patients show symptoms like stiffened muscles and jerking movements is a major risk factor for sudden unexpected death in epilepsy [46]. Mostly a single lead surface electrode is used as electrophysiological setup and detection performance reported in literature is generally high for GTCS (95% in [47], 100% in [48]). Disadvantage of this approach is the limited applicability as GTCS represent the minority of all seizure types and focal motor seizure or focal seizures with impairment of consciousness can only be detected with EEG or ECG.

Chapter 2
Methods

In this chapter computer algorithms for detection of epileptic seizures and patterns with repetitive discharges in EEG, ECG, and EMG signals will be presented. This work aims to improve on the current state of the start in automatic biosignal evaluation by using algorithms that imitate clinical validated methodologies. The presented algorithms are based on features that capture physiological knowledge of human experts and therefore facilitate comprehensibility. Further, rule-based classification schemes that imitate decision trees of EEG experts are implemented. To allow real-time monitoring of patients, algorithms must evaluate the EEG data stream continuously. All methods in this work avoid patient specific parameter settings to allow universal application to new patients and therefore maximize efficacy in clinical practice. This will greatly increase applicability in clinical workflow.

2.1 Time domain analysis of EEG

A major problem in automatic EEG analysis is the irregularity of patterns as outlined in section 1.1 and Figure 2. Discontinuous patterns like periodic and burst suppression patterns (Figure 4) basically consist of transient discharges occurring randomly in time. To create the basis for analysis of all these patterns a novel approach for signal analysis is sought. The idea to solve this problem was triggered by the observation of irregular seizure patterns that showed wide spread energy distribution of frequency bins although visual interpretation of a rhythmic pattern was obvious (Figure 2). In this section time-domain methods for analysis of such patterns are investigated.

2.1.1 Epileptiform wave sequence algorithm

In a first approach to automatically analyse EEG in the time-domain the epileptiform wave sequence (EWS) algorithm was developed **[A1]**[2]. The algorithm especially addresses a group of ictal patterns that show moderate irregular rhythmic signal patterns, abrupt phase changes and distortions. Figure 2 shows an example of a rhythmic EEG pattern with a dominating frequency (A, B) compared to (C, D) where the energy is spread over a wide frequency range. Both signals were annotated by board certified EEG experts as rhythmic ictal activity and need a highly flexible approach for automatic detection. The EWS algorithm directly scans time series data of EEG for consecutive maxima of ictal waves. A wave must fulfil certain criteria on instantaneous frequency, dynamic, and the relative amount of high frequency noise. All three measurements are defined in the time domain (Equation 1, 2, and 3) of **[A1]**. High frequency noise is estimated by the squared sum of all amplitude signal points in the wave. Then waves are clustered using the instantaneous frequency and sequences are defined of non-overlapping waves of the same cluster. This algorithm copes with the irregularity of distinct waves in the rhythmic pattern and creates a compact information entity including the number of waves, average frequency, and amplitude. Figure 5 gives an example of the time position of maximum points that were included in a rhythmic pattern group **[A1]**. Such definition of rhythmic groups would lead to a very high false detection rate if compared to seizures markers of EEG experts. The reason is that the regularity of waves need to be incorporated in classification. The next step in the algorithm therefore checks if waves show a certain amount of average regularity in morphology. The similarity value γ defines such measure as the quotient of the variance of the average wave to the average variance of all waves w_n in a group of N waves.

[2] References [A1]...[A8] are peer reviewed papers including major contributions of the author.

$$\gamma = \left| \frac{\mathbf{var}_{\tau}\left(\mathbf{E}_{\mathbf{n}}\left[\mathbf{w}_{\mathbf{n}}(\tau)\right]\right)}{\mathbf{E}_{\mathbf{n}}\left[\mathbf{var}_{\tau}\left(\mathbf{w}_{\mathbf{n}}(\tau)\right)\right]} \right| \quad \mathbf{n} \in \mathbf{1}..\mathbf{N}, \tau \in \tau_{\mathbf{n}}^{\mathbf{max}}..\tau_{\mathbf{n+1}}^{\mathbf{max}} \tag{1}$$

The value of γ reaches 1 if all waves are exact replicates of each other. In this case the averaged wave will be equal to each wave w_n. Wave groups with very different waves will result in a low value of γ with a lower bound of zero.

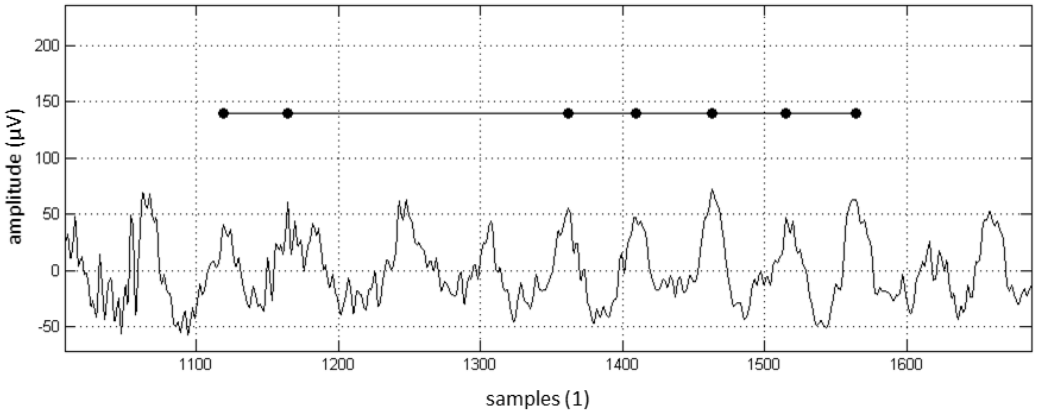

Figure 5: Wave sequence marked by dots on the maximum position of waves found by the EWS algorithm.

A limitation of the EWS algorithm is that waves describe the transition between two consecutive discharges and not a single discharge. Noisy signal components in between two discharges may avoid detection of the sequence. Therefore, wave sequences without major interruption are required in general. This implies that analysis of EEG patterns with large inter-discharge intervals like periodic pattern or burst suppression pattern is hardly feasible. Analysis of such patterns in addition to unequivocal rhythmic seizure patterns will be possible with an advancement of the idea of the EWS algorithm called direct discharge segmentation.

2.1.2 Direct discharge segmentation

Based on the idea of the epileptiform wave sequence (EWS) algorithm **[A1]** a universal segmentation algorithm was developed. The algorithm works directly on the time series data without using templates for correlation analysis. The working principle is to capture transient discharges of signal sources that show large amplitudes compared to other signal components. For example, an epileptic seizure activity can be modelled as sequence of discharges without inter-discharge intervals. Periodic patterns are based on the same model but show distinct inter-discharge intervals with lower amplitude levels. Surprisingly, this approach can be extended to non-EEG signals. In the ECG three defections with the largest amplitudes correspond to the electrical signal of the ventricular contraction called QRS complexes. This typical waveform repeats for every heartbeat and dominates all other electrical activity of the heart like the electrical signal of the atrial contraction and the ventricular repolarisation. QRS complexes can be therefore detected with the same computer algorithm than EEG discharges. A limitation of the direct discharge segmentation is that superposition of two signal sources with the same amplitude levels cannot be differentiated. E.g. two sine waves with different frequencies but same amplitudes are not separable by analysing the dominating summation peaks in the time domain. Another counter example are burst suppression patterns that include an overlay of several small discharges in each burst and are therefore indistinguishable in the time domain (e.g. Figure 4, signal A and B).

Direct discharge segmentation defines a list of wave segments that represent discharges. Each discharge is defined by start, end, and maximum time point. Constraints are defined that control how fine grained the signal is segmented. These constraints include the minimum discharge amplitude and the minimum discharge duration. The algorithm starts by defining local minimum and maximum time points of the time series. By using each maximum X as start point the time span is extended along the signal samples in the surrounding of X to reach the minimum amplitude and duration constraints (Figure 6, A). After that, a secondary search is conducted to extend the discharge. Two lines g_1 and g_2 starting at the borders of the primary discharge (M_s, M_e) are defined. Signal samples with smaller amplitudes than the values of g_1 and g_2 at the same time position are included in addition and will extend the initial discharge (Figure 6, B).

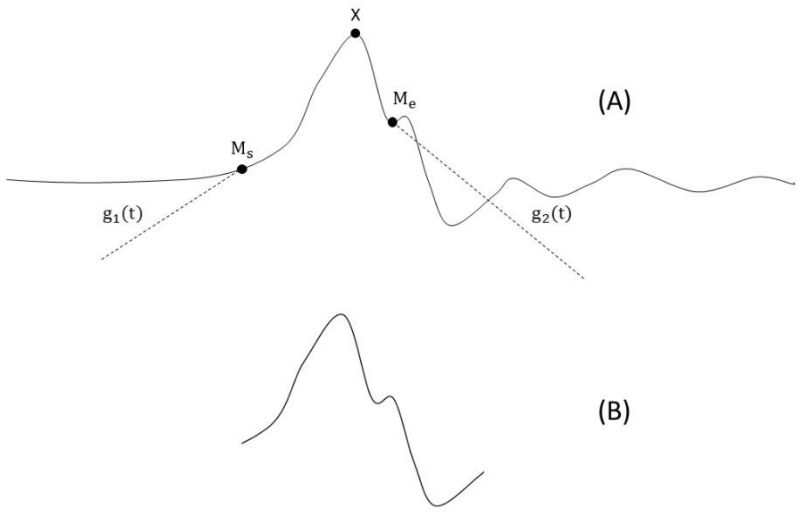

Figure 6: Direct discharge segmentation exemplified on a single EEG discharge. First maximum point X, start and end time points $\mathbf{M_s, M_e}$ are detected in the primary search (A). Two extensions lines $\mathbf{g_1, g_2}$ define the final extent of the discharge segment (B). Discharge segments are the basis for automatic detection of seizure patterns in EEG and ECG.

All extracted discharges are then classified as slow wave, sharp wave, spikes, theta, or alpha waves depending on the morphology, duration, and amplitude of the wave. The resulting list of discharges will then be used as input for pattern detection algorithms. Comparing this method to the Fast Fourier Transform (FFT) or the continuous wavelet transform (CWT) algorithm reveals some major differences. First, no signal templates are needed that act as eigenvectors for signal analysis. In direct discharge segmentation, the morphology of discharges is found directly without correlation of eigenvectors with the signal. This avoids definition of a previously unknown signal morphology with the advantage that series of discharges with random morphology are segmented with equal quality. Analysis quality in continuous wavelet transform relates to the number of wavelet coefficients that include most of the signal energy. This approach avoids such templates and will therefore show high signal analysis quality at low computational cost.

2.1.3 Quantification of discontinuous EEG

The level of EEG discontinuity is an important clinical marker in the EEG of critical care patients. To lay the basis for quantification of discontinuous EEG like burst suppressions patterns a simple time-domain method was developed. The primary goal is to measure any drop in signal amplitude that exceeds a factor of two over a certain time period, as defined by the ACNS terminology for critical care EEG patterns [9]. The ACNS terminology requires that bursts have a duration of at least 0.5 seconds and that more than fifty percent of the recording consists of suppressions. No restriction to the signal waveform in the non-suppressed segments are made to allow maximal flexibility. The

algorithm works on fixed non-overlapping analysis windows of 15 seconds. First, for each EEG channel the peak-to-peak amplitude of non-overlapping chunks of 0.4 length is measured in each segment **[A4]**. Amplitude values are smoothed using moving average filters with 0.5 and 1.5 second length and are then used for detection of discontinuous segments. If the average amplitude of chunks in a time period of at least 1.5 seconds is less than half of the amplitude of a foregoing or following interval of at least 0.5 seconds, then a discontinuous EEG segment was detected. The average amplitude value of the whole burst and suppression period is used as separation value to assign chunks in the segment to either the burst or the suppression group. If no suppression was found all chunks get assigned to the burst group. The average amplitude values of the burst and suppression group is evaluated. The result of this quantification will then be used for detection of burst suppression patterns.

2.2 Detection of repetitive patterns

In this work, computer algorithms for detection of distinct types of EEG patterns with repetitive elements were developed. For simplicity, such patters are referred as 'repetitive patterns' in this text. Rhythmic and periodic patterns are two examples that are closely related to each other when automatically analyzed in the time-domain. Therefore, a single algorithm is used detect both. Burst suppression patterns are detected based on a different approach that will be described in section 2.2.2.

2.2.1 Detection of rhythmic and periodic patterns

Rhythmic and periodic patters are signal waveforms consisting of a repetitive sequence of discharges. According to the standardized critical care EEG terminology [14] a rhythmic activity is a "repetition of a waveform with relatively uniform morphology and duration, and without an interval between consecutive waveforms" whereas a periodic discharge is a "repetition of a waveform with relatively uniform morphology and duration with a quantifiable inter-discharge interval between consecutive waveforms and recurrence of the waveform at nearly regular intervals". These signal waveforms are the basis for automatic EEG monitoring and detection of epileptic seizures which will be the major goals of this work. The probability to observe these patterns in the EEG is not equally likely. In general, ICU patients with rhythmic or periodic patterns below 4 Hz are three times more frequent than ICU patients with electrographic seizures [49]. Further, rhythmic and periodic patterns may not be clearly ictal or inter-ictal depending on the patient under investigation [20] [50] and are subject of current clinical research [51]. This shows that an automatic detection algorithm needs to objectively measure patterns in a large frequency range without further classification as ictal or non-ictal.

In this work, an algorithm for automatic detection of rhythmic and periodic patterns was developed. The basis of the algorithm are segments that were found by the direct discharge segmentation algorithm. The algorithm starts by combination of similar discharges that occur at the same time in different signal channels **[A3]**. This will be the case in EEG when several channels may pick up signal from the same group of neurons that fire synchronously. This spatial combination utilizes a hierarchical cluster algorithm where the maximum point of each segment is used as distance metric. This early integration of spatial information allows exclusion of signalling distortions by checking the amplitude and phase and by exploiting information of absolute sensor positions on the head. Then consecutive multi-channel discharges are used to build groups with similar properties in spatial extent, duration, and amplitude. In each group, the algorithm scans for sequences of discharges that fulfil criteria related to the ACNS definition **[A3]**. Properties of the group are calculated by averaging values of all included multi-channel discharges. The algorithm therefore results in information elements describing rhythmic or periodic patterns that span over several seconds and channels.

2.2.2 Detection of burst suppression patterns

Burst suppressions are patterns with repetitive bursts in the EEG of critical care patients. Like periodic patterns, burst suppression patterns show distinct low amplitude intervals that are clinically described as discontinuous EEG. Despite these similarities, the non-suppressed period includes a superposition of random discharges that are not separable. For automatic detection and monitoring, a different approach than the grouping of single discharges is needed. In general, not only burst suppression (suppressed below 10 µV) but also burst attenuation patterns (suppressed but higher than 10 µV) are interesting clinical markers and should be considered during EEG evaluation. In this work, it was decided to not only detect burst suppression but also to model burst attenuation patterns and other variants with a single detection algorithm. This approach will result in a much richer information for clinicians as also transitional states can be observed over time. Detections of burst suppression patterns will be continuously visualized in real-time on a graphical user interface to allow bedside monitoring (see section 2.3).

The algorithm is based on the quantification for discontinuous EEG presented in section 2.1.3. First, consecutive burst and suppression chunks are connected over time to form larger segments. A hierarchical cluster algorithm spatially combines this channel-wise information. The relative time duration between detected burst and suppression segments, the amplitude relation between suppression/attenuation and burst, as well as spatial distribution is used to quantify detected patterns. A complete description of the algorithm can be found in section 2 of **[A4]**. Figure 7 shows an EEG example and detected burst as well as suppression areas. The major algorithmic steps are outlined in addition.

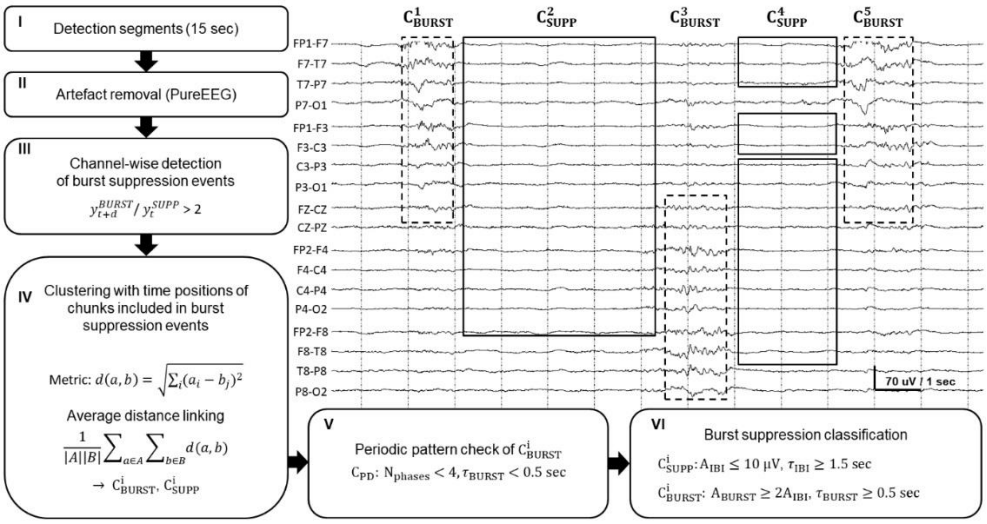

Figure 7: Block diagram of the burst suppression detection algorithm including EEG example showing burst and suppression clusters **[A4]**.

2.3 EEG monitoring with NeuroTrend

The graphical visualisation of detected patterns and other simple EEG measures over a time axis is called quantitative EEG (qEEG). The amplitude integrated EEG (aEEG) is a simple but effective implementation of qEEG that is widely used to monitor neurological state of neonates [53]. The goal is to visualise the time course of EEG properties on a compressed time axis to allow visual assessment of trends without the need to continuously check and interpret the raw EEG curves. In this work, a graphical user interface called NeuroTrend **[B2]** was designed to visualise a more sophisticated quantification of EEG: detections of patterns with repetitive discharges and seizures. Figure 8 gives an example of NeuroTrend for an EEG of a critical care patient. The idea was to use major properties of patterns instead of a directly visualizing the amplitude of the whole EEG. NeuroTrend shows the spatial localisation of rhythmic and periodic patterns, their frequency and amplitude. The major difference and advantage to aEEG is the visual amplification of focal rhythmic activity with small amplitudes by information reduction. AEEG that will show one amplitude value of one EEG channel or of one hemisphere for each time point. Pattern detections will show only values for EEG segments including rhythmic activity. This pre-filtering of information reduces the number of visual elements and will allow the reviewer to focus on most important time segments. In addition, focal rhythmic activity on a single channel is not averaged with non-rhythmic activity before the amplitude value is calculated. This way a single channel with rhythmic or periodic activity will contribute in the same way to a visualisation element on the quantitative EEG screen than a rhythmic pattern distributed over hemispheres. Burst suppression patterns are visualized by showing detections in red bars. In addition, a quantification of amplitude and time relations between burst and suppression segments are shown to allow recognition of pattern transitions between continuous and discontinuous EEG. The relative amplitude drop V_{drop} of a burst suppression is calculated by the average burst amplitude minus the average inter-burst interval (IBI) amplitude divided by the average suppression amplitude (Equation 2). The relative suppression time T_{supp} is defined as the average IBI time duration divided by the sum of burst and IBI time duration (Equation 3).

$$\boldsymbol{V_{\mathrm{drop}}}(\%) = \frac{\boldsymbol{V_{\mathrm{burst}} - V_{\mathrm{ibi}}}}{\boldsymbol{V_{\mathrm{burst}}}} * \mathbf{100} \tag{2}$$

$$\boldsymbol{T_{\mathrm{supp}}}(\%) = \frac{\boldsymbol{T_{\mathrm{ibi}}}}{\boldsymbol{T_{\mathrm{burst}} + T_{\mathrm{ibi}}}} * \mathbf{100} \tag{3}$$

Both measures are given in percent, whereas the relative suppression time is visualized with an inverted y axis. This axis inversion will reduce relative distance on the y-axis for burst suppression patterns with either a long suppression times or very reduced suppression amplitudes. In this work NeuroTrend is used to monitor detections in real-time and as visual inspection tool for EEG review **[A3, A4, A6]**.

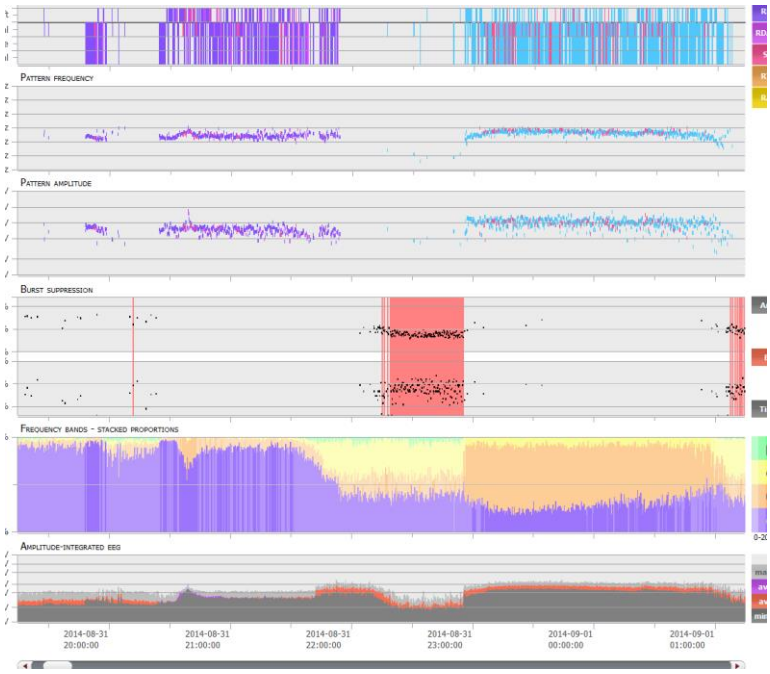

Figure 8: NeuroTrend as EEG Monitor: The graphical visualisation tool NeuroTrend showing results of a continuous EEG recording of a critical care patient in real-time.

2.4 Automatic seizure detection

To find epileptic seizures in long-term EEG recordings is the major goal in epilepsy monitoring units. Although some seizures are clinically observed through medical staff or relatives, many seizures are missed during sleep. In this work, algorithms for automatic detection of seizures were developed to alarm hospital staff in case of a seizure and to raise efficiency during EEG review.

2.4.1 Seizure alarming

In epilepsy monitoring units (EMU) seizure alarming devices are used to ensure patient safety and to initiate neurological tests after the seizure. Seizure alarming devices evaluate the signal ad-hoc or "online" while recording of the signal. The algorithm first has to detect the seizure in the data stream and then to trigger an alarm as fast as possible.

EpiScan is an seizure detection and alarming algorithm built for the use in epilepsy monitoring units [33]. The algorithm is trimmed to detect seizure activity with short time delay in EEG signals. In this work the EWS algorithm was introduced into the initial version of the EpiScan algorithm [33] to further increase detection sensitivity by automatic detection of irregular ictal EEG patterns **[A1]**. EpiScan finally consisted of a continuous wavelet transformation algorithm called periodic waveform analysis (PWA) for detection of small amplitude regular rhythmic activity of deep cortical sources, the EWS algorithm for detection of irregular rhythmic activity, and a simple muscle activity detector **[A2]**. The false alarm rate of an alarming device should be as low as possible to avoid ignorance of relevant alarms by staff members. To use the EWS algorithm in EpiScan a patient adaption mechanism had to be developed. Patient adaption uses recorded feature vectors of the past to evaluate the normal EEG activity of the patient over time and for each channel. Feature vectors include measures like frequency and amplitude of the rhythmic group. Goal is to build a spatiotemporal model depending on frequency that is used to avoid repeating alarms. False alarms can be caused by physiological non-ictal patterns like alpha background activity or other benign theta variants **[A2]**. After removing statistical outliers of collected measures the remaining maximum amplitude value of each channel and frequency band is used to normalize currently detected feature values. Experiment series showed best adaption results for a time window of 4 hours. Applied to the EWS adaption model this means that physiological events of a certain pattern frequency and location will reduce the alarm probability of patterns with similar frequency and location of up to 4 hours in the future. The resulting combined algorithms therefore build a novelty detection system [52]. This approach imitates the behaviour of EEG technicians that first learn commonly present EEG patterns of a patient and later decides if a pattern is pathologic or physiologic based on that initial evaluation.

An issue with the seizure alarming approach of EpiScan is that a very high computational effort is required to reduce the detection delay. Pipeline processing to calculate values for later use cannot be utilized to reduce alarming delay as much as possible. The reason is that delays of filters, detection algorithms of rhythmic groups, and classification algorithms would exceed the delay requirement if calculated sequentially. In addition, computation of spatiotemporal models including information of several hours of EEG activity are very processing intensive. These two issues illustrate that post-hoc detection of seizures in long-term EEG requires different algorithmic structures to reduce calculation complexity.

2.4.2 Seizure detection

Compared to seizure alarming devices that are designed to detect epileptic seizure as early as possible at low false alarm rates, the use case for offline seizure detectors are different. Here, the primary goal is to find most of the seizures to reduce work load in EEG evaluation by avoiding scanning through EEG segments without valuable information. Seizure detectors are therefore trimmed to have a very high sensitivity and fast computation times. False detections are accepted as long as most of uninteresting EEG segments can be skipped. Of highest importance is the overall gain in work load reduction by using automatic detection systems. Based on the algorithms for automatic detection of rhythmic and periodic patterns and the knowledge gained by the EpiScan algorithm an automatic seizure detection algorithm for post-hoc analysis was developed [A5]. The incorporation of all available signal modalities aims to raise detection sensitivity as high as possible.

The multimodal seizure detection algorithm exploits information from EEG and ECG data. The EEG signal is used to detect rhythmic EEG patterns with steadily increasing amplitudes in the delta, theta, and alpha Berger EEG bands (1-13 Hz). The detection algorithm for rhythmic pattern detection based on direct discharge segmentation was used for that purpose [A5]. Detected rhythmic patterns are used in a novelty detection approach to normalize current detections using past rhythmic activity as baseline. In addition, rhythmic pattern amplitudes must exceed a frequency dependent amplitude threshold (Figure 9, part B) that models the average physiological EEG activity. This frequency dependent EEG model will allow detection of rhythmic alpha activity (8-13Hz) with amplitudes below 20 µV, whereas delta activity needs a much higher rhythmic pattern amplitude to trigger detections.

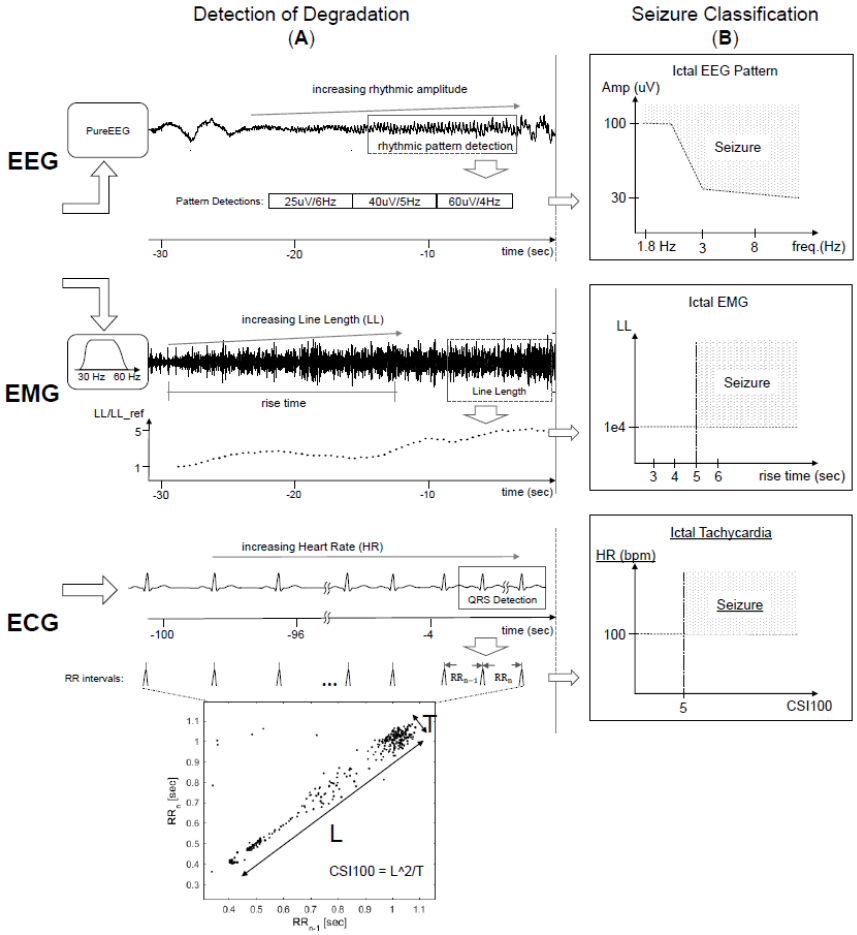

Figure 9: Multimodal seizure detection system with high sensitivity of [A5].

Generalized tonic-clonic seizures (GTCS) are commonly followed by temporary paralysis and post ictal apnea which is potentially life threatening for patients. Such seizures show extremely high muscle activity that differentiates from physical muscle movement. The surface EMG is able to capture muscle activity, but dedicated EMG electrodes are not part of the standard setup in the EMU. In this work electrical activity from chin, scalp, and face muscles captured by EEG electrodes are utilized for seizure detection. A bandpass filter between 30 and 60 Hz extracts EMG data from standard EEG signal. Signal strength was quantified using the line length method defined as the sum of distances between each consecutive data sample in non-overlapping windows [A5]. Similar to the detection scheme for rhythmic ictal activity, the line length measure is required to raise continuously over a time period of several seconds. Further, the absolute value must exceed the mean value of a baseline window by a certain factor.

Finally, ECG signals were used for detection of epileptic seizures. ECG signals from chest electrodes were used to measuring heart rate and for automatic detection of ictal tachycardia [A5]. In addition to a novelty detection scheme, the condition on increasing heart rate had to be defined more precisely. Heart rate increases are very common during normal physical activity and during arousals in sleep [7]. The cardiac sympathetic index (CSI) is a graphical method to define a steadily increasing signal with a certain slope [45]. In this work, 100 heart beats are used for CSI calculation and to avoid detections having low CSI values. Signal epochs that show high heart rates compared to baseline and having large CSI values trigger detections based on ECG.

2.5 Measuring algorithm performance

2.5.1 Quantifying detection performance

Detection performance of computer algorithms is quantified using sensitivity, specificity, and false detection rate. Seizure annotations from experienced epileptologists are used as reference. Time points of detected events are compared to these annotations to define four logical variables: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Seizure epochs were defined as true positive (TP) if at least one detection occurred within the epoch time range. Detections outside of seizure epochs were defined as false positives (FP). False positives occurring within a time span of less than 30 s were counted as a single false detection because validation of this event is done by assessment of the same screen of EEG. Seizure epochs without a matching detection were defined as false negative (FN). All other epochs are defined as TN. Figure 10 summaries these definitions. In this work a maximum seizure duration of 3 minutes is assumed which covers most seizures [54]. To handle also seizure annotations just before ictal patterns, a seizure epoch is defined ranging from 30 seconds before manual seizure annotation to 150 seconds after annotation.

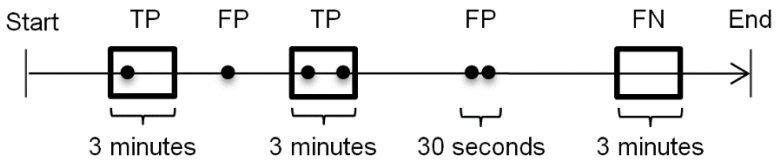

Figure 10: Definition of logical variables TP, FP, and FN. All remaining epochs are defined as TN.

Sensitivity (SE) is defined as the number of true positive events (n_{TP}) divided by the sum of true positives and false negatives (n_{FN}):

$$\mathrm{SE} = \frac{n_{\mathrm{TP}}}{n_{\mathrm{TP}} + n_{\mathrm{FN}}} \tag{4}$$

Specificity is defined as the number of true negative events (n_{TN}) divided by the sum of the number of true negatives and the number of false positives (n_{FP}):

$$\mathrm{SP} = \frac{n_{\mathrm{TN}}}{n_{\mathrm{TN}} + n_{\mathrm{FP}}} \tag{5}$$

Human interpretation of specificity is problematic for seizure detection algorithms as the number of TN exceeds the number of FP by an order of magnitude. E.g. an average patient in the epilepsy unit will has 3 seizures in 100 hours of observation time on average **[A2]**. Let us assume one-minute time intervals for all logical variables, 500 false detections ($n_{\mathrm{FP}} = 500$) as well as 3 seizures ($n_{\mathrm{TP}} = 3$) and 100 hours of recording ($n_{\mathrm{hour}} = 100$). In this example

$$n_{\mathrm{TN}} = 100 * 60 - 500 - 3 = 5497 \tag{6}$$

$$\mathrm{SP} = \frac{5497}{(5497 + 500)} \approx 0.916 = 91.6\% \tag{7}$$

The value of 91.6% misleadingly implies high specificity because of the unbalanced ratio between n_{TN} and n_{FP}. Much simpler interpretability can be reached by using false detections per time unit. An established unit in the community is number of false positives divided by the number of recording days ($n_{\mathrm{hour}}/24$) [55]:

$$\mathrm{FD/24h} = 24\frac{n_{\mathrm{FP}}}{n_{\mathrm{hour}}} \tag{8}$$

This value can be interpreted easily by physicians as the average number of false detections per day and linearly correlates to the amount of work needed for assessment of detections or alarms. For the given example

$$\mathrm{FD/24h} = 24\frac{500}{100} = 120 \tag{9}$$

which is considered as a too high false detection rate for most use cases of seizure detection systems.

2.5.2 Number of patients for validation

Sensitivity and false detection rate quantify detection performance of a single patient recording. To assess average sensitivity and false detection rate a high number of patients need to be involved to result in point estimates with small confidence intervals. In this work, an experiment was performed to determine the reliability of the mean and confidence interval of sensitivities in a seizure detection study with 30 and 94 patients [A2]. In this experiment, sensitivity results of 242 patients with seizures were used to simulate 1,000 virtual trials. For each virtual trial, 30 results out of 242 were drawn randomly to simulate a study with n=30. The average sensitivities and the 95% confidence intervals of all 1,000 virtual studies were calculated using a bootstrapping method [56]. The same protocol was applied to simulate a study with n=94 patients and results were compared to the n=30 studies. For trials with n=30 average sensitivities of single trials ranged from 43% to 92% with an average confidence interval of 25%. Open circles in Figure 11 shows results for all 1,000 trials. This high variance of means and confidence intervals demonstrates that a sample size of 30 is too small for reliably estimating the detection sensitivity. Most compelling is the fact that a single n=30 study achieved a mean sensitivity of 92% with a very small confidence interval of 14% simply by chance, although this result does not reflect the real detection performance at all. The same procedure was repeated for n=94 and plotted in comparison to the n=30 studies in Figure 11 (filled circles). Sensitivity of trials with n=94 are much more concentrated between 55% and 80%. The confidence intervals for detection sensitivity reduced to a range of 10% to 17% with an average confidence interval of 14%.

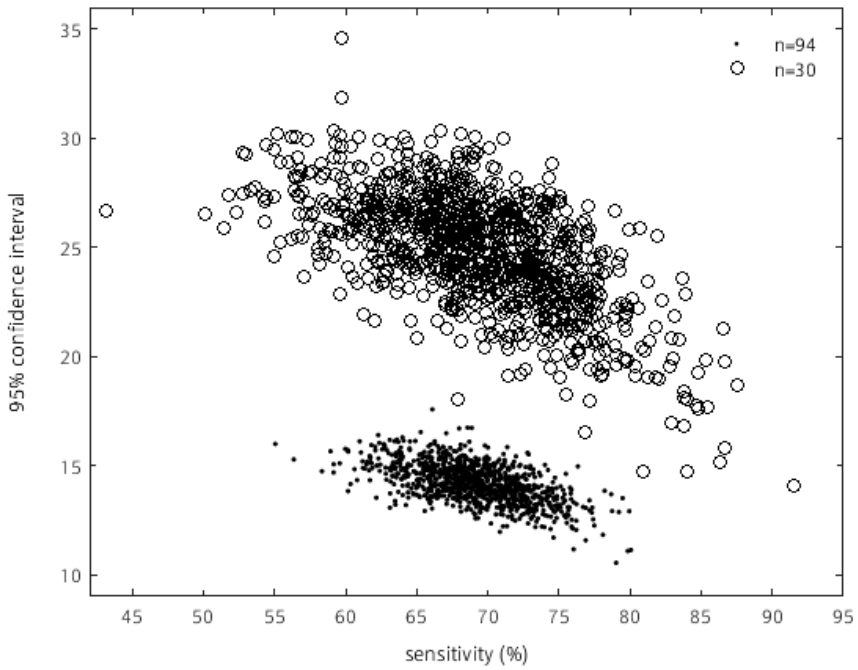

Figure 11: The 95% confidence interval (CI) of the mean sensitivity plotted over the mean sensitivity of 1,000 virtual studies with n=30 (open circles) and n=94 (filled circles) [A2].

Chapter 3
Results

Proper clinical validation of computer algorithms is of highest importance. The study design must be considered carefully to retrieve meaningful results. For validation, EEG data of a high number of patients has to be used for high statistical confidence. Several types of epilepsies should be included in the data so that results will correspond to real world situation. Further, whole recordings must be used for validation without modification or manual editing. Data of multiple recording sites will reduce the problem of hidden confounders. EEG data for development and data for testing should be separated to avoid overfitting of algorithms. In this work retrospective datasets were previously used for algorithm development; prospective datasets were not used for development of algorithms. In this chapter, results of clinical validation studies are presented that will show the performance of the developed computer algorithms.

3.1 Validation of repetitive pattern detections

3.1.1 Validation of rhythmic and periodic detections

To get a first impression on the clinical applicability of rhythmic and periodic pattern detections a small retrospective study was conducted with NeuroTrend. We randomly selected EEG recordings of 10 ICU patients including 187 hours of EEG with an associated clinical EEG report from the Comprehensive Epilepsy Center, Medical University of South Carolina (MUSC) [A3]. A clinical neurophysiologist from the Neurological Center Rosenhuegel (NCR) used NeuroTrend and the EEG to write a detection guided EEG report. Then, the detection guided report was compared to the content of the manual EEG report from neurophysiologists of MUSC to evaluate similarities and differences. The study was designed to use a restricted evaluation time of 10 minutes only, to show if detected patterns lead to the most important EEG time points of the recording. The result showed that in three out of five patients with reported seizures all seizures from the original report were found [A3]. In two patients, additional features were found which shows the high sensitivity of the detection method. Not all seizures of another patient could be found because of minor EEG correlates. There, seizures were defined by clinical observation and video analysis. In one patient, the single seizure was not found because of no clear EEG activity. The result of this small retrospective study was promising but the sensitivity and specificity of pattern detections had to be shown with high statistical confidence.

A prospective multi-center study was designed and performed in two neurological ICUs including 68 patients to validate sensitivity and specificity of automatically detected rhythmic and periodic patterns [A7]. Annotations of two clinical neurophysiologist were used as gold standard in this study. Automatically detected patterns were compared to annotations to define sensitivity and specificity of the detection algorithm. The result showed an overall sensitivity of 94%. Periodic patterns detections showed a sensitivity of 80%, rhythmic patterns below 4 Hz a sensitivity of 82%. The study showed that an average specificity of 78% could be reached for 20 second annotation segments [A7].

In summary, high sensitivity but only average specificity could be measured for the detected patterns. Further, the possibility to find most important EEG segments using a graphical visualisation of detected patterns was shown but validation of the underlying EEG will be mandatory. Results also showed that further work on classification algorithms is needed to raise specificity.

Finally, two open questions for the application in the neurological ICU remained: 1) How does the quality of the detected patterns degrade when only few EEG electrodes are applied? 2) Can the graphical visualisation be used as bedside monitor for real-time monitoring of the neurological state? These two important questions were addressed in two separate follow-up studies (3.1.3 and 3.2).

3.1.2 Validation of burst suppression detections

Clinical validation of the burst suppression detection algorithm was conducted using prospectively collected data of multiple centers [A4]. Because of the lack of EEG annotations that describe the discontinuity in EEG signals directly, only the burst suppression detection of the algorithm was validated. Assessment included data of three recording centers including data from the Massachusetts General Hospital in Boston. In summary data of 88 patients with a total of 3982 hours of EEG was included. Sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) were measured by comparing detection results of the algorithm with EEG annotations of several reviewers. The inter-rater agreement (IRA) between human annotations of burst suppression and detection results was assessed to gain insight into the quality of human annotation on this pattern type.

Results showed substantial inter-rater agreement with a kappa value of 0.71 (0.68–0.74). Detailed analysis of the annotated data showed that EEGs from the Massachusetts General Hospital included significantly more burst suppression patterns than data of the two ICUs in Vienna. Based on the uneven distribution of review segments with burst suppression patterns it was decided to avoid patient-wise statistics. Instead, detection performance was analysed over all annotation segments of all patients.

Detection results based on consensus annotation of two reviewers yielded in an average sensitivity of 90% and an average specificity of 84%. A PPV of 64% (61–66) and a NPV of 96% (96–97) was measured. The high detection performance on prospectively collected data showed that automatic detection of burst suppression patterns is possible at high accuracy and that real-time monitoring of burst suppression patterns is feasible for ICU patients.

3.1.3 Effects of electrode reduction

The reduction of EEG electrodes plays an important role in the ICU and for ambulant EEG. Dedicated staff for EEG recording and evaluation is generally missing in ICUs. The standardized 10/20 electrode placement system including 19 electrodes is therefore a major obstacle to implement continuous EEG in the ICU. In this work, the first published systematic approach to evaluate detection sensitivity of rhythmic, periodic, and burst suppression patterns with reduced electrode sets was conducted [A8]. Based on the 10/20 electrode system including 19 sensors the number of electrode was reduced stepwise towards fours minimal electrode setup schemes know from literature. For each of the resulting 50 reduction schemes the sensitivity and specificity of the automatic pattern detections (section 2.2) was assessed.

Results showed that reduction of EEG electrodes rapidly reduces sensitivity for periodic patterns. Contrary, rhythmic and especially burst suppression patterns can be detected with much less electrodes at the same sensitivity level (Figure 12). Further, placement of the reduced number of sensors is important. Results showed that the forehead electrode placement system is most beneficial **[A8]**. The study has a major impact to the neurophysiological community as continuous EEG is the fastest growing market in the EEG segment and the costs for electrode application and maintenance are high [57].

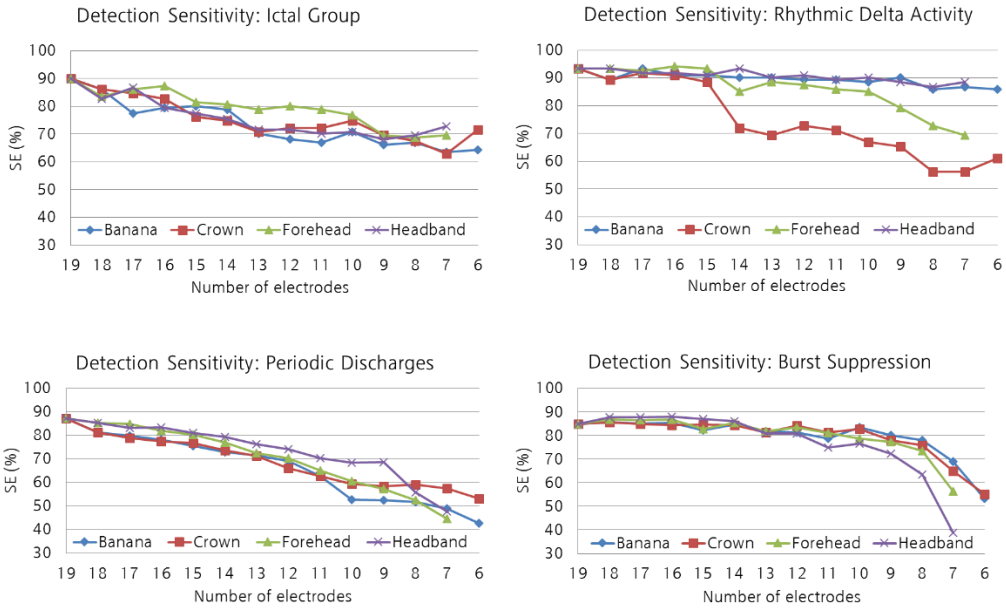

Figure 12: Results of systematic electrode reduction to pattern detection **[A8]**. Analysis was done on the group of ictal patterns, periodic patterns (PD), rhythmic delta activity (RDA), and burst suppression patterns (BS). Reduction of electrodes was done in four montage schemes called BAM, CRM, FOM, and HAM **[A8]**.

3.2 NeuroTrend as bedside monitor

The urgent need for automatic EEG evaluation in the neurological ICU triggered the design of a study involving the monitoring tool NeuroTrend. Automatic detections of seizures and patterns of ictal–interictal uncertainty can be used for post-hoc analysis and delayed patient treatment [49]. But of higher importance is continuous EEG evaluation of intensive care patients to allow early treatment of life threatening pathologies like cerebral ischemia or status epilepticus. Goal of the study "Applicability of NeuroTrend as a bedside monitor in the neuro ICU" **[A6]** was to assess if ICU caregivers will be able to correctly interpret NeuroTrend screens for treatment decisions. The study involved retrospective ratings of 120 NeuroTrend screenshots by 18 briefly trained nurses and biomedical analysts. Substantial agreement (60–80%) was found for electrographic seizure patterns, periodic discharges, and seizure suspicion. The study demonstrated that subclinical seizures can be detected by shortly trained caregivers using NeuroTrend.

3.3 Validation of seizure detections

3.3.1 Validation of seizure alarming

The seizure alarming system EpiScan includes the EWS algorithm for detection of rhythmic seizure activity (section 2.4.1) as well as the existing continuous wavelet based algorithm PWA [33]. First, the algorithm was validated in a retrospective study. In total EEGs of 275 patients including 22,463 hours of recording and 623 seizures were used to measure detection performance. By comparing performance of different parameter sets and combinations of detection features the most promising variant for a subsequent prospective multi-center trial should be evaluated. Highest detection performance was observed when combining the existing continuous wavelet transform algorithm for seizure detection [33] using high specific settings with the EWS algorithm **[A1]** (Figure 13, HYB_L1). The results proofs that time domain algorithms are capable of adding value to frequency domain based algorithms for detection of epileptic seizures in EEG.

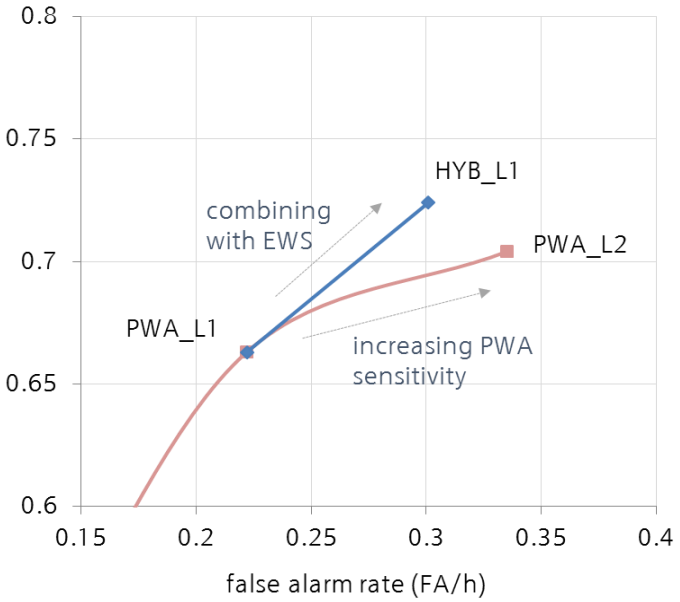

Figure 13: Detection performance of continuous wavelet transform based algorithm with parameter set 1 (PWA_L1), the parameter set L2 having higher sensitivity (PWA_L2), and combination of PWA_L1 and EWS detection (HYB_L1) **[A1]**.

In this work one of the largest published prospective multi-center studies for an automatic seizure alarming algorithm was conducted on EpiScan **[A2]**. Three epilepsy monitoring units in Austria and the Netherlands joint the research consortium for assessment of detection delay, sensitivity, and false alarm rate. Data of 539 long-term EEG recordings including over 42,000 hours of EEG was included consisting of 205 prospective and 334 retrospective EEGs. Overall 1836 seizures were recorded and annotated. Annotations of seizures in prospective data was done using a novel multi-level scheme that solved the problem of different levels of seizure perception. In addition, the one and only FDA and CE cleared seizure detection software at that time from Persyst (www.persyst.com) was evaluated using the same prospective dataset. Result showed that detection performance measured on ''probably yes'' seizures resulted in a sensitivity of 81% (74–86%) and a false alarm rate of 7.1 false alarms per day (FA/24h). Detection performance degraded to 78% sensitivity and 7.08 FA/24h when annotations with lower perception values were used for comparison. Finally, a sensitivity of 72% with similar false detection rate was achieved when all seizures from the standard

EMU review protocol were used for evaluation. Comparison to the Persyst seizure detection software showed that sensitivity was 81% for EpiScan whereas Persyst reached a value of 75%. A two-paired t-test showed that the difference of sensitivities was of no statistically significance. False alarm rate of EpiScan was lower compared to the result of Persyst (-27% or 2.68 FA/24 h). Although the overall detection sensitivity did not reach values over 85% similar results could be reached in all centers which allows the universal application of the method. Figure 14 depicts the main result of the EpiScan study.

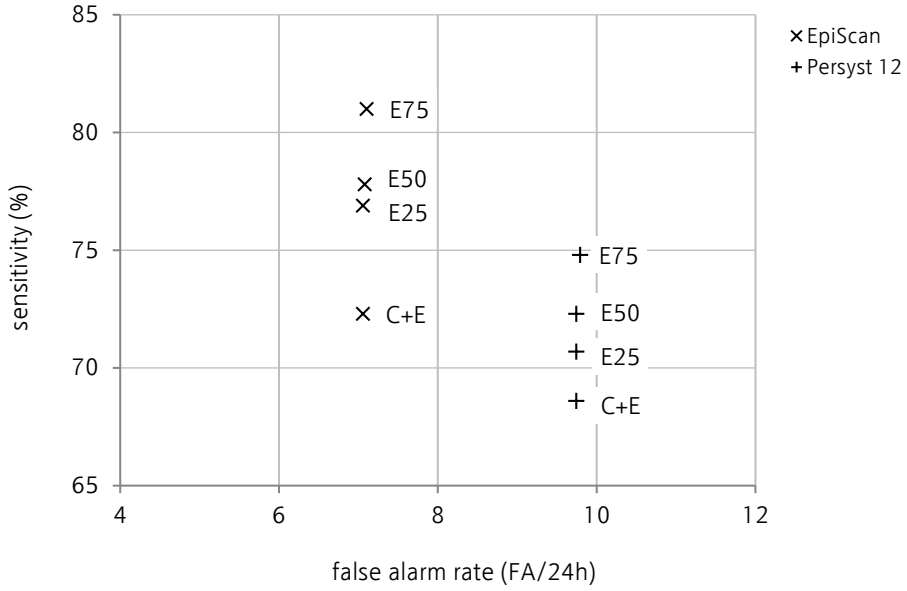

Figure 14: Result of the EpiScan study **[A2]**. Detection performance of EpiScan and Persyst 12 compared in four seizure categories using the same prospective data.

3.3.2 Validation of offline seizure detection

Clinical validation of the seizure detection algorithm described in section 2.4.2 involved retrospective recordings of two epilepsy monitoring units including EEG and ECG data of 94 patients with 494 epileptic seizures. Special attention was given to reduced electrode setups with only 8 electrodes. The use case of counting seizure using an automatic seizure detection algorithm for ambulant recordings of outpatients is of high interest to reduce costs and to increase quality of seizure therapy studies **[A5]**. Result showed that a high overall detection sensitivity of 86% could be reached with the multimodal seizure detection algorithm. An average false detection rate of 16.5 false detections in 24 hours was reached. The subgroup of temporal lobe epilepsy patients showed 94% detection sensitivity and generalized tonic-clonic seizures could be detected to 100%. The work also showed that by using 8 frontal and temporal electrodes the detection sensitivity reduces only by 5% compared to the full setup. This finding will allow reduction of EEG setup time by a factor of three and ensures patient compliance. Compared to manual evaluation of EEG the proposed method will increase overall detection sensitivity and efficiency of EEG review.

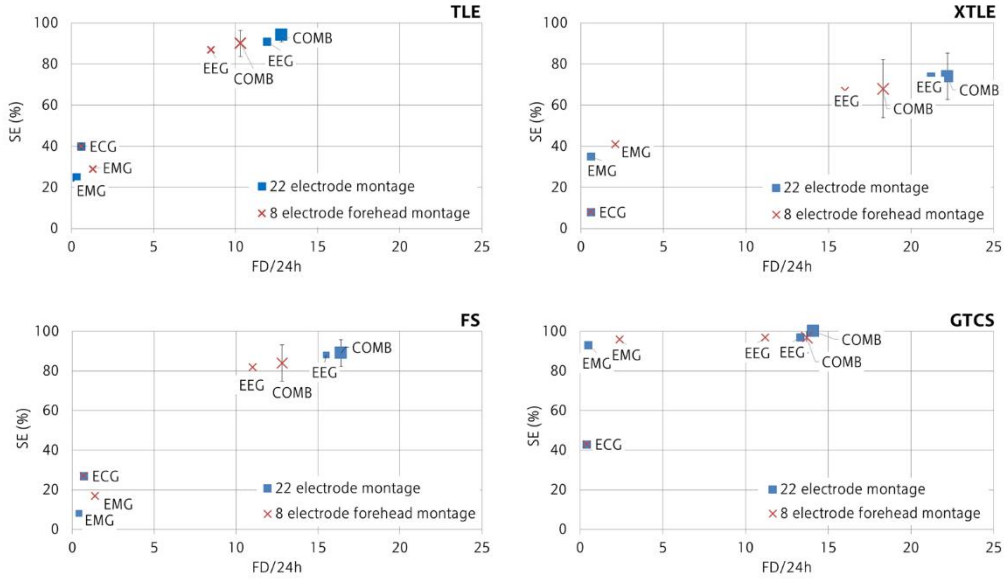

Figure 15: Results of the multimodal seizure detection algorithm. Reduced and full electrode sets were used to validate the algorithm on temporal lobe epilepsy patients (TLE), extra temporal lobe epilepsy patients (XTLE), on patients with focal seizures (FS), and on patients with generalized tonic-clonic seizures (GTCS).

Chapter 4
Conclusion

Aim of this work was to implement real-time monitoring of patients by using computer algorithms for automatic EEG evaluation. The detections of seizures and repetitive discharges capture neuronal activity of patients which are highly relevant for diagnostic. Through this approach several major improvements in the workflow of clinical neurophysiologists were achieved. First, real-time surveillance of seizure activity in epilepsy patients became possible. An alarming algorithm for epileptic seizures was developed that is able to notify medical staff of epilepsy monitoring units within a few seconds. The false alarm rate had to be as low as possible in order not to be ignored. Second, monitoring of critical care patients was implemented by extracting several kinds of patterns with repetitive discharges from the EEG. These detection algorithms are able to capture information of rhythmic, periodic, and burst suppression patterns. A graphical visualization tool called NeuroTrend was invented that allows monitoring of the neurological state through display of these patterns at bedside of critical care patients. Further, efficiency of post-hoc EEG review was raised by introducing a computer assisted EEG review workflow. A multimodal and high sensitive detection algorithm for epileptic seizures was developed for that purpose. By showing results of seizure and pattern detections in NeuroTrend the review time of EEG can be reduced while maintaining quality.

In this work, a novel and unique approach for signal analysis of biosignals was invented. The idea started with the observation that frequency transformation algorithms are not able to analyze seizure patterns at equal quality. The EWS algorithms was invented in this work to directly scan time-series EEG data for consecutive epileptiform discharges. By iteratively validating the waveform between two discharges an epileptiform wave sequence is build. The prove of concept that such feature is able to find at least some seizure activity with higher precision than a precursor wavelet based algorithm was shown in this work. This idea was taken further to build a more universal signal analysis framework. The direct discharge segmentation is able to separate pulsing discharges in time series data. Based on this segmentation the detection algorithms for seizures and other repetitive patterns were built. The algorithms mimic the visual approach of EEG reviewers by grouping discharges with similar amplitude, duration, and spatial location. Such pattern fields are then validated to fulfill minimal requirements on the number of discharges and finally represent EEG patterns of medical interest.

Real-time monitoring of critical care patients was implemented by developing algorithms that automatically detect standardized EEG patterns defined by the American Clinical Neurophysiology Society (ACNS). The algorithms scan the EEG for rhythmic, periodic, and burst suppression patterns and allow objective quantification of EEG content. By using the computer results a standardized and reproducible way of EEG evaluation regarding these patterns was created. Further, by showing these repetitive patterns on a graphical user interface a novel computer assisted workflow was introduced that is able to provide real-time assessment of the EEG. In this work, clinical validation of computer algorithms was conducted with high effort. Clinical diagnostic studies including data of multiple recording centers were used for each use case. For the validation of rhythmic and periodic patterns detections a clinical study including 68 patients of two neurological ICUs was completed. High detection sensitivity and specificity of the different pattern variants was proven. A major obstacle to implement continuous EEG in neurological ICUs is the high number of electrodes in the standard EEG setup. In this dissertation, the first published study on systematically reduced electrode sets was conducted. The effect on the algorithm performance and the reviewer impression was measured when reducing the number of EEG electrodes one by one. Results showed that some EEG patterns of critical care patients are more vulnerable to reduced electrodes sets than others. Further, clinical validation of the burst suppression pattern detection algorithm was conducted in cooperation with the Massachusetts General Hospital and two neurological intensive care units in Vienna. Results showed high performance levels above previously published data in this field. The algorithm is therefore able to monitor sedation levels as well as spontaneously arising discontinuities in the EEG of critical care patients in real-time.

In this dissertation, a novel seizure alarming algorithm was developed that will raise patient safety and efficiency of continuous EEG in the epilepsy monitoring unit. The problem of high inter-patient variability of ictal EEG had to be solved first. Small amplitude rhythmic patterns in a single channel may be the most pronounced seizure activity in one patient but may be present continuously as normal EEG activity in another patient. To solve this issue, a patient specific component had to be used that works without explicit training data. The spatiotemporal patient model collects past information of a fixed time window for baseline estimation and for feature normalization. This approach was then used to implement the seizure alarming system EpiScan that notifies medical staff in case of seizure. For clinical validation of EpiScan one of the largest published prospective multi-center studies was initiated and completed. The international consortium of three epilepsy centers gathered long-term EEG data of 205 patients including gold standard annotations of seizures. Results showed high sensitivity and acceptable false alarm rates in all centers. In addition, the prospectively collected EEG data was used to compare EpiScan to a commercially available seizure detection software. This unique approach is the first published comparison of two EEG based seizure detection algorithms on a large prospective dataset and showed the superior false alarm rate of EpiScan.

To raise effectiveness of EEG review was another goal of this work. The user interface NeuroTrend was designed to show detected patterns as simple colour coded elements and therefore implements a computer aided monitoring system for EEG. This work showed that review time can be reduced significantly when using NeuroTrend to review EEGs of epilepsy patients. To further raise quality of EEG diagnostic, a high sensitive seizure detection algorithm was developed that is based on detections of rhythmic EEG patterns. In addition, the algorithm incorporates information of ECG and EMG to raise sensitivity. A clinical study including recordings of 91 epilepsy patients of the Epilepsy Center Erlangen and the Neurological Center Rosenhügel in Vienna was used for validation. Results show that some seizure types can be detected with 100% sensitivity while having an acceptable false detection rate of 16.5 per day on average. Compared to manual EEG review the number of EEG screens for validation therefore reduces from over 34,000 to only 66 to find 86% of the seizures on average.

Clinical evaluation studies of this work already received attention in the scientific community. An review paper on EEG based seizure detection algorithms by Jean Gotman, who published the first computer algorithm of that kind in 1982, describes the EpiScan study **[A2]** as "probably a realistic assessment of the performance that is possible on current methods" [23]. Further, the work on electrode reduction received an editorial by Kenneth G. Jordan in the journal Clinical Neurophysiology [57].

This work showed that the highly complex problem of automatic EEG analysis can be solved to a certain extend and that new and innovative solutions are possible for diagnostic of EEG. Computer based monitoring of EEG can potentially increase quality of diagnostic by providing simple and easy to understand descriptions that can be interpreted not only by experts in the field. With the help of computer algorithms for EEG evaluation future applications like home monitoring and wearable seizure alarming devices are feasible. Such applications can further reduce the overall costs of the health care system and will increase quality of life for patients.

Chapter 5
Bibliography

5.1 Peer reviewed papers

A1 Fürbass 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: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2012. p. 1020–3.

A2 Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. Clin Neurophysiol. 2015 Jun;126(6):1124–31.

A3 Fürbass F, Hartmann MM, Halford JJ, Koren J, Herta J, Gruber A, et al. Automatic detection of rhythmic and periodic patterns in critical care EEG based on American Clinical Neurophysiology Society (ACNS) standardized terminology. Neurophysiol Clin Neurophysiol. 2015 Sep;45(3)

A4 Fürbass F, Herta J, Koren J, Westover MB, Hartmann MM, Gruber A, et al. Monitoring burst suppression in critically ill patients: Multi-centric evaluation of a novel method. Clin Neurophysiol. 2016 Apr;127(4):2038–46.

A5 Fürbass F, Kampusch S, Kaniusas E, Koren J, Pirker S, Hopfengärtner R, et al. Automatic multimodal detection for long-term seizure documentation in epilepsy. Clin Neurophysiol. 2017 Aug;128(8):1466–72

A6 Herta J, Koren J, Fürbass F, Zöchmeister A, Hartmann M, Hosmann A, et al. Applicability of NeuroTrend as a bedside monitor in the neuro ICU. Clin Neurophysiol. 2017 Jun;128(6):1000–7.

A7 Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: A new approach for screening continuous EEG in the intensive care unit. Epilepsy Behav. 2015 Aug;49:273–9.

A8 Herta J, Koren J, Fürbass F, Hartmann M, Gruber A, Baumgartner C. Reduced electrode arrays for the automated detection of rhythmic and periodic patterns in the intensive care unit: frequently tried, frequently failed? Clin Neurophysiol. 2017 Aug;128(8):1524–31.

5.2 Published abstracts

B1 F. Fürbass, J. Halford, J. Herta, J. Koren, H. Perko, K. Schnabel, M. Weinkopf, T. Kluge, A. Gruber, C. Baumgartner, M. Hartmann. NeuroTrend: Computational EEG analysis for critically ill patients based on the standardized terminology of the ACNS; World Congress of Neurology 2013, Wien; 21.09.2013 - 26.09.2013; in: "Journal of the Neurological Sciences 333 (2013)", Elsevier, 333 (2013), ISSN: 0022-510x; S. 285.

B2 Hartmann MM, Koren J, Fürbass F, Weinkopf M, Schnabel K, Halford JJ, et al. NeuroTrend: Rapid review of continuous EEGs from ICUs. ACNS 2014 Annual Meeting, Atlanta, GA, USA; 04.02.2014 - 09.02.2014; in: "Journal of Clinical Neurophysiology", Lippincott Williams & Wilkins, (2014), ISSN: 0736-0258; S. 1.

B3 M. Hartmann, F. Fürbass, G. Gritsch, A. Skupch, J. Koren, J. Herta, Ch. Baumgartner, T. Kluge. Rapid Identification of Ictal Events in Long-Term EEG Recordings from Epilepsy Monitoring; Epilepsia, The International League Against Epilepsy, (2015), ISSN: 1528-1167; S. 86 - 87.

5.3 References

[1] R. S. Fisher ***et al.***, "Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," ***Epilepsia***, vol. 46, no. 4, pp. 470–472, Apr. 2005.

[2] P. Kwan and M. J. Brodie, "Early identification of refractory epilepsy," ***N. Engl. J. Med.***, vol. 342, no. 5, pp. 314–319, Feb. 2000.

[3] J. W. Sander, "The use of antiepileptic drugs—principles and practice," ***Epilepsia***, vol. 45 Suppl 6, pp. 28–34, 2004.

[4] R. S. Fisher, "The New Classification of Seizures by the International League Against Epilepsy 2017," ***Curr. Neurol. Neurosci. Rep.***, vol. 17, no. 6, p. 48, Jun. 2017.

[5] B. Scheepers, P. Clough, and C. Pickles, "The misdiagnosis of epilepsy: findings of a population study," ***Seizure***, vol. 7, no. 5, pp. 403–406, Oct. 1998.

[6] B. Schmitt, "Sleep and epilepsy syndromes," ***Neuropediatrics***, vol. 46, no. 3, pp. 171–180, Jun. 2015.

[7] S. Zschocke and H.-C. Hansen, ***Klinische Elektroenzephalographie***, 3., Aktualisierte und erweiterte Auflage. Springer Berlin Heidelberg, 2011.

[8] W. O. Tatum, B. A. Dworetzky, and D. L. Schomer, "Artifact and Recording Concepts in EEG:," ***J. Clin. Neurophysiol.***, vol. 28, no. 3, pp. 252–263, Jun. 2011.

[9] L. J. Hirsch ***et al.***, "American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology: 2012 version," ***J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.***, vol. 30, no. 1, pp. 1–27, Feb. 2013.

[10] K. Nakashima, M. M. Todd, and D. S. Warner, "The relation between cerebral metabolic rate and ischemic depolarization. A comparison of the effects of hypothermia, pentobarbital, and isoflurane," ***Anesthesiology***, vol. 82, no. 5, pp. 1199–1208, May 1995.

[11] C. A. Pagni and J. Courjon, "ELECTROENCEPHALOGRAPHIC MODIFICATIONS INDUCED BY MODERATE AND DEEP HYPOTHERMIA IN MAN," ***Acta Neurochir. Suppl.***, vol. 14, p. SUPPL 13:35-49, 1964.

[12] S. Ching, P. L. Purdon, S. Vijayan, N. J. Kopell, and E. N. Brown, "A neurophysiological-metabolic model for burst suppression," ***Proc. Natl. Acad. Sci. U. S. A.***, vol. 109, no. 8, pp. 3095–3100, Feb. 2012.

[13] M. B. Westover ***et al.***, "The human burst suppression electroencephalogram of deep hypothermia," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 126, no. 10, pp. 1901–1914, Oct. 2015.

[14] E. Niedermeyer, D. L. Sherman, R. J. Geocadin, H. C. Hansen, and D. F. Hanley, "The burst-suppression electroencephalogram," ***Clin. EEG Electroencephalogr.***, vol. 30, no. 3, pp. 99–105, Jul. 1999.

[15] A. O. Rossetti, E. Carrera, and M. Oddo, "Early EEG correlates of neuronal injury after brain anoxia," ***Neurology***, vol. 78, no. 11, pp. 796–802, Mar. 2012.

[16] American Clinical Neurophysiology Society, "Guideline 6: a proposal for standard montages to be used in clinical EEG," ***Am. J. Electroneurodiagnostic Technol.***, vol. 46, no. 3, pp. 226–230, Sep. 2006.

[17] S. Beniczky ***et al.***, "Standardized computer-based organized reporting of EEG: SCORE," ***Epilepsia***, vol. 54, no. 6, pp. 1112–1124, Jun. 2013.

[18] J. Claassen ***et al.***, "Electrographic seizures and periodic discharges after intracerebral hemorrhage," ***Neurology***, vol. 69, no. 13, pp. 1356–1365, Sep. 2007.

[19] S. M. LaRoche, Ed., ***Handbook of ICU EEG Monitoring***, 1st ed. Demos Medical, 2012.

[20] V. Rodríguez, M. F. Rodden, and S. M. LaRoche, "Ictal-interictal continuum: A proposed treatment algorithm," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 127, no. 4, Apr. 2016.

[21] J. Claassen, S. A. Mayer, R. G. Kowalski, R. G. Emerson, and L. J. Hirsch, "Detection of electrographic seizures with continuous EEG monitoring in critically ill patients," ***Neurology***, vol. 62, no. 10, pp. 1743–1748, May 2004.

[22] J. Koren ***et al.***, "Prediction of rhythmic and periodic EEG patterns and seizures on continuous EEG with early epileptiform discharges," ***Epilepsy Behav. EB***, vol. 49, pp. 286–289, Aug. 2015.
[23] M. Bialer ***et al.***, "Seizure detection and neuromodulation: A summary of data presented at the XIII conference on new antiepileptic drug and devices (EILAT XIII)," ***Epilepsy Res.***, vol. 130, pp. 27–36, Feb. 2017.
[24] M. Atkinson, K. Hari, K. Schaefer, and A. Shah, "Improving safety outcomes in the epilepsy monitoring unit," ***Seizure J. Br. Epilepsy Assoc.***, vol. 21, no. 2, pp. 124–127, Mar. 2012.
[25] J. Gotman, "Automatic recognition of epileptic seizures in the EEG," ***Electroencephalogr. Clin. Neurophysiol.***, vol. 54, no. 5, pp. 530–540, Nov. 1982.
[26] H. Qu and J. Gotman, "A seizure warning system for long-term epilepsy monitoring," ***Neurology***, vol. 45, no. 12, pp. 2250–2254, Dec. 1995.
[27] Y. Zheng, J. Zhu, Y. Qi, X. Zheng, and J. Zhang, "An automatic patient-specific seizure onset detection method using intracranial electroencephalography," ***Neuromodulation J. Int. Neuromodulation Soc.***, vol. 18, no. 2, p. 79–84; discussion 84, Feb. 2015.
[28] J. Claassen ***et al.***, "Recommendations on the use of EEG monitoring in critically ill patients: consensus statement from the neurointensive care section of the ESICM," ***Intensive Care Med.***, vol. 39, no. 8, pp. 1337–1351, Aug. 2013.
[29] P. Kurtz ***et al.***, "Continuous electroencephalography in a surgical intensive care unit," ***Intensive Care Med.***, Nov. 2013.
[30] K. Jin and N. Nakasato, "Long-cherished dreams for epileptologists and clinical neurophysiologists: automatic seizure detection in long-term scalp EEG," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 125, no. 7, pp. 1289–1290, Jul. 2014.
[31] F. Pauri, F. Pierelli, G. E. Chatrian, and W. W. Erdly, "Long-term EEG-video-audio monitoring: computer detection of focal EEG seizure patterns," ***Electroencephalogr. Clin. Neurophysiol.***, vol. 82, no. 1, pp. 1–9, Jan. 1992.
[32] Scott B Wilson, M. L. Scheuer, R. G. Emerson, and A. J. Gabor, "Seizure detection: evaluation of the Reveal algorithm," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 115, no. 10, pp. 2280–2291, Oct. 2004.
[33] M. M. Hartmann ***et al.***, "EpiScan: Online seizure detection for epilepsy monitoring units," in ***Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE***, 2011, pp. 6096–6099.
[34] A. J. Gabor, "Seizure detection using a self-organizing neural network: validation and comparison with other detection strategies," ***Electroencephalogr. Clin. Neurophysiol.***, vol. 107, no. 1, pp. 27–32, Jul. 1998.
[35] R. Meier, H. Dittrich, A. Schulze-Bonhage, and A. Aertsen, "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. Off. Publ. Am. Electroencephalogr. Soc.***, vol. 25, no. 3, pp. 119–131, Jun. 2008.
[36] K. M. Kelly ***et al.***, "Assessment of a scalp EEG-based automated seizure detection system," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 121, no. 11, pp. 1832–1843, Nov. 2010.
[37] A. S. Zandi, G. A. Dumont, M. Javidan, and R. Tafreshi, "Detection of epileptic seizures in scalp electroencephalogram: an automated real-time wavelet-based approach," ***J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.***, vol. 29, no. 1, pp. 1–16, Feb. 2012.
[38] R. Hopfengärtner ***et al.***, "Automatic seizure detection in long-term scalp EEG using an adaptive thresholding technique: a validation study for clinical routine," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, vol. 125, no. 7, pp. 1346–1352, Jul. 2014.
[39] K. S. Eggleston, B. D. Olin, and R. S. Fisher, "Ictal tachycardia: the head-heart connection," ***Seizure***, vol. 23, no. 7, pp. 496–505, Aug. 2014.
[40] F. Leutmezer, C. Schernthaner, S. Lurger, K. Pötzelberger, and C. Baumgartner, "Electrocardiographic changes at the onset of epileptic seizures," ***Epilepsia***, vol. 44, no. 3, pp. 348–354, Mar. 2003.
[41] T. De Cooman, C. Varon, B. Hunyadi, W. Van Paesschen, L. Lagae, and S. Van Huffel, "Online Automated Seizure Detection in Temporal Lobe Epilepsy Patients Using Single-lead ECG," ***Int. J. Neural Syst.***, p. 1750022, Feb. 2017.
[42] W. J. C. van Elmpt, T. M. E. Nijsen, P. A. M. Griep, and J. B. A. M. Arends, "A model of heart rate changes to detect seizures in severe epilepsy," ***Seizure***, vol. 15, no. 6, pp. 366–375, Sep. 2006.
[43] I. Osorio, "Automated seizure detection using EKG," ***Int. J. Neural Syst.***, vol. 24, no. 2, p. 1450001, Mar. 2014.
[44] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen, "Detection of epileptic seizures with a modified heart rate variability algorithm based on Lorenz plot," ***Seizure***, vol. 24, pp. 1–7, Jan. 2015.

[45] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen, "Using Lorenz plot and Cardiac Sympathetic Index of heart rate variability for detecting seizures for patients with epilepsy," ***Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.***, vol. 2014, pp. 4563–4566, 2014.
[46] C. Harden ***et al.***, "Practice guideline summary: Sudden unexpected death in epilepsy incidence rates and risk factors: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the American Epilepsy Society," ***Neurology***, vol. 88, no. 17, pp. 1674–1680, Apr. 2017.
[47] C. Á. Szabó ***et al.***, "Electromyography-based seizure detector: Preliminary results comparing a generalized tonic-clonic seizure detection algorithm to video-EEG recordings," ***Epilepsia***, vol. 56, no. 9, pp. 1432–1437, Sep. 2015.
[48] S. N. Larsen, I. Conradsen, S. Beniczky, and H. B. D. Sorensen, "Detection of tonic epileptic seizures based on surface electromyography," ***Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.***, vol. 2014, pp. 942–945, 2014.
[49] J. P. Koren ***et al.***, "Rhythmic and periodic EEG patterns of 'ictal-interictal uncertainty' in critically ill neurological patients," ***Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol.***, Nov. 2015.
[50] A. Sivaraju and E. J. Gilmore, "Understanding and Managing the Ictal-Interictal Continuum in Neurocritical Care," ***Curr. Treat. Options Neurol.***, vol. 18, no. 2, p. 8, Feb. 2016.
[51] J. Claassen, "How I treat patients with EEG patterns on the ictal-interictal continuum in the neuro ICU," ***Neurocrit. Care***, vol. 11, no. 3, pp. 437–444, Dec. 2009.
[52] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," ***Signal Process.***, vol. 99, pp. 215–249, Jun. 2014.
[53] D. Zhang and H. Ding, "Calculation of compact amplitude-integrated EEG tracing and upper and lower margins using raw EEG data," ***Health (N. Y.)***, vol. 05, no. 05, pp. 885–891, 2013.
[54] E. Trinka, J. Höfler, and A. Zerbs, "Causes of status epilepticus," ***Epilepsia***, vol. 53 Suppl 4, pp. 127–138, Sep. 2012.
[55] P. M. Bossuyt ***et al.***, "STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies," ***BMJ***, vol. 351, p. h5527, Oct. 2015.
[56] T. J. DiCiccio and B. Efron, "Bootstrap confidence intervals," ***Stat. Sci.***, vol. 11, no. 3, pp. 189–228, Sep. 1996.
[57] K. G. Jordan, "Reduced Electrode Arrays for Acute Electroencephalography: Can Less be More?," ***Clin. Neurophysiol.***, vol. 128, no. 8, May 2017.

Chapter 6
Full text of publications

6.1 Paper A1: EWS in EpiScan

Title: Combining Time Series and Frequency Domain Analysis for a Automatic Seizure Detection

Authors: F. Fürbass, M. Hartmann, H. Perko, A. Skupch, P. Dollfuß, G. Gritsch, C. Baumgartner, T. Kluge

Published in: Conference Proceedings IEEE Engineering in Medicine and Biology Society

Year: 2012

Authors' contribution: Fürbass Franz developed the EWS algorithm and conducted the retrospective observational study including comparison to previous results of the PWA algorithm. Manuscript editing and corrections were done by Franz Fürbass. The PWA algorithm development and data annotations were done by coauthors.

34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Combining Time Series and Frequency Domain Analysis for a Automatic Seizure Detection

F. Fürbaß, M. Hartmann, H. Perko, A. Skupch, P. Dollfuß, G. Gritsch, C. Baumgartner, and T. Kluge

Abstract— **The detection of epileptic seizures in long-term electroencephalographic (EEG) recordings is a time-consuming and tedious task requiring specially trained medical experts. The EpiScan [1–4] seizure detection algorithm developed by the Austrian Institute of Technology (AIT) has proven to achieve high detection performance with a robust false alarm rate in the clinical setting. This paper introduces a novel time domain method for detection of epileptic seizure patterns with focus on irregular and distorted rhythmic activity. The method scans the EEG for sequences of similar epileptiform discharges and uses a combination of duration and similarity measure to decide for a seizure. The resulting method was tested on an EEG database with 275 patients including over 22000h of unselected and uncut EEG recording and 623 seizures. Used in combination with the EpiScan algorithm we increased the overall sensitivity from 70% to 73% while reducing the false alarm rate from 0.33 to 0.30 alarms per hour.**

I. Introduction

ELECTROENCEPHALOGRAPHY (EEG) is the medical standard for examination of patients suffering from epilepsy. Long-term EEG recordings lasting for several days are needed for pre-surgical evaluation of patients with refractory epilepsy types or patients having inacceptable medical side-effects. The unpleasant situation for patients monitored continuously with video and EEG is impaired with an increased risk of seizures as anti epileptic drugs are reduced. Not only is a thorough analysis of the long-term EEG involving medical experts required but also a 24 hour surveillance of the EEG in real-time. An automatic method that marks seizure events can reduce evaluation time drastically and increases patient security by alerting medical staff immediately.

A major problem in the automatic seizures detection is the inter-patient variability of ictal patterns ranging from quasi periodic patterns over patterns with high frequency variation or abrupt phase changes to completely irregular groups of ictal discharges. The existing EpiScan algorithm [1–5] identifies ictal activity with rhythmic or periodic morphology using a continuous wavelet transform approach. This method has reached a high overall sensitivity and a low false alarm rate in uncut, unselected clinical data [4]. Ictal patterns with high frequency variation and phase changes were partly recognized by the EpiScan algorithm, but lead to higher overall false alarm rates.

In this paper a time domain algorithm for detection of epileptic seizures called *Epileptiform Wave Sequence* (EWS) analysis will be presented that is designed to reliably detect epileptic seizures with rhythmic morphology and especially addresses the group of ictal patterns with a moderate irregular structure, abrupt phase changes or distortions. In this context an epileptiform wave is a pathologic discharge seen in the EEG and a sequence is an epoch dominated by waves with the same properties.

Such wave sequences result from repeating discharges of groups of cortical neurons with abnormal hypersychronous behavior [6]. The post-synaptic electrical potentials [6] coming from a synchronous firing neuronal group in the seizure onset zone mix non-linearly with other physiological signals, will be attenuated at the skull bone and finally sum up with artifacts from scalp muscles and technical electrode artifacts. Ictal patterns with moderate irregular morphology are often seen in patients with ictal slowing, rhythmic delta activity or in a secondary generalization phase of the seizure when the rhythmic pattern at onset (PAO) was obscured. Figure 1 shows an example of an ictal signal interfered with noise that can be easily analyzed by a method based on a

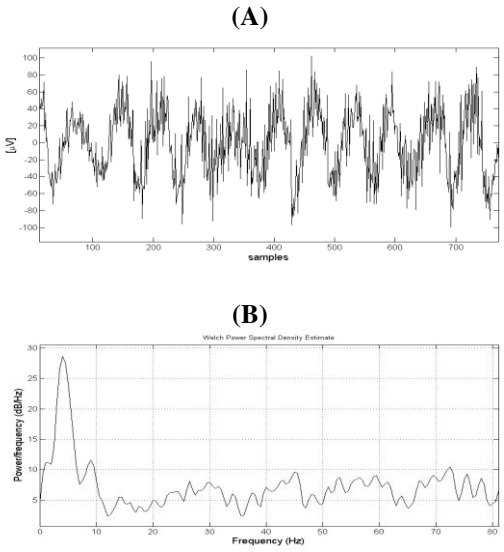

Figure 1 Ictal EEG with rhythmic morphology and with muscle artifacts **(A)** compared to the power spectrum of the same signal **(B)**. A method based on spectral analysis will be capable of detecting the underlying rhythmic pattern as the power spectrum reveals a strong peak at 4 Hz.

This work was supported in part by Grant WST3-T-81/014-2008 from the Provincial Government of Niederösterreich with the European Regional Development Fund, and by FWF Grant L585-B19.

Franz Fürbaß (corr. Author), Manfred M. Hartmann , Hannes Perko, Ana Skupch, Peter Dollfuß, Gerhard Gritsch, and Tilmann Kluge are with the Austrian Institute of Technology (AIT), Donau-City-Strasse 1, 1220 Vienna, Austria

Christoph Baumgartner is with the Neurological Department Rosenhügel at General Hospital Hietzing, Vienna, Austria.

spectral estimation. A time series analysis is preferable for the signal in Figure 2 because the more irregular distance between adjacent peaks do not result in a stable frequency for spectral estimation. Hence a combination of both approaches will be preferable.

Time-series algorithms that search for unique markers in the signal to segment and analyze the patterns are commonly used in the field of EEG analysis. Gotman e.g. [7–9] showed several applications with this approach. In this paper the method is evaluated using a comprehensive EEG database with statistical relevance.

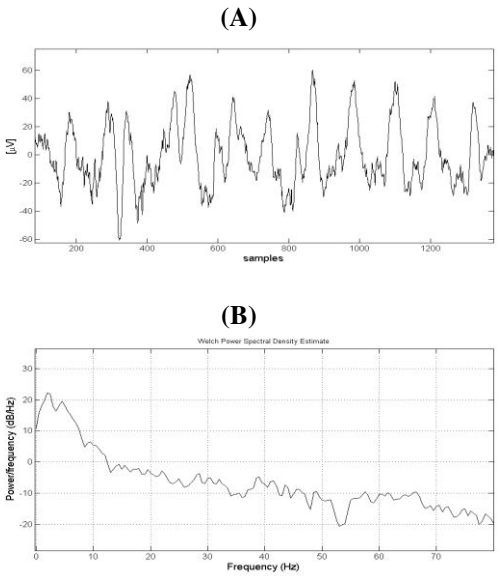

Figure 2 : Ictal EEG signal with an irregular morphology **(A)** that leads to a fuzzy spectrum after transformation in the frequency domain **(B)**. The power spectrum looks smeared and has equally high components from 2 to 5 Hz. The averaging of the spectral analysis obscures the simple structure of repeating discharges seen in the time series.

II. METHODS

A. Frequency domain analysis

EpiScan utilizes a continuous wavelet transformation algorithm called *Periodic Waveform Analysis* (PWA) to search for rhythmic patterns in the surface electrode EEG channels. More details can be found in [2], and a complete performance analysis using the AIT EEG database can be found in [3].

B. Time domain analysis

The *Epileptiform Wave Sequence* (EWS) analysis was designed to reliably detect epileptic seizure patterns with rhythmic morphology in the time domain and especially encounter the problems of high frequency variation, abrupt phase changes and signal distortions by muscle or electrode artifacts. The EWS analysis will proceed as follows:

1. wave classification
2. wave clustering
3. sequence creation
4. intra-sequence correlation

Step 1 will find interesting signals called waves, Step 2 group waves with the same properties using a clustering algorithm. Step 3 then creates a sequence from waves belonging to the same cluster. Step 4 calculates a correlation value acting as similarity measure for the sequence.

1) Wave classification

To find epileptogenic waves, the signal is scanned iteratively over time for maxima of ictal discharges. A wave is defined as the signal between two adjacent maxima that fulfills the following classification criteria:

- the instantaneous frequency f_n has to be in range
- the dynamic d_n of the wave has to be high enough
- the high frequency noise of a wave has to be small

The instantaneous frequency f_n of wave n is the ratio of the sampling frequency f_s to the duration of the wave measured between time points of two maximum peaks τ_k^{max}.

$$f_n = \frac{f_s}{\tau_{n+1}^{max} - \tau_n^{max}} \quad (1)$$

The dynamic of a wave d_n is measured using the two maximum values and the including minimum value t_{min_k}

$$d_n = \left|k(\tau_n^{min} - \tau_n^{max}) + x(\tau_n^{max}) - x\left(\tau_{n+1}^{min}\right)\right| \quad (2)$$

with the slope k defined as

$$k = \frac{x(\tau_n^{max}) - x(\tau_{n+1}^{max})}{\tau_{n+1}^{max} - \tau_n^{max}}. \quad (3)$$

The high frequency noise of a wave is measured using the sum of squares of all adjacent signal points $x(t)$ in the wave. This wave-extraction scheme will solve the problems of phase changes, signal distortion and high frequency variation as all waves are handled separately. Only a sequence of epileptogenic waves with similar morphology reliably specify a seizure pattern, so a sequence needs to be found.

2) Wave clustering

Waves will be clustered using the measures found in the wave classification step. The clustering algorithm uses a given variance to find a single subgroup that dominates the ictal EEG. The variances were found using a statistical analysis of the AIT EEG-database and cross-validation with knowledge from specially trained EEG experts. Clustering is done sequentially using the instantaneous frequency, wave amplitude and the noise measure.

3) Sequence creation

After wave clustering a sequence will be defined allowing gaps that correspond to signal distortions. This step will create the sequence using only clustered waves but leaving out signal epochs with artifacts. This will avoid mixing artifacts and interesting signals like in the spectral widening

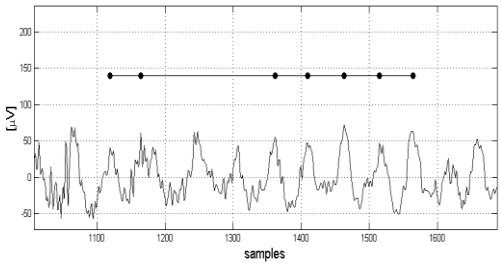

Figure 3 Ictal EEG (sampling frequency 256 Hz) with a repeating but irregular morphology and the detected sequence marked with filled circles at the end of the wave. Note that some waves are left out as they do not fit into the cluster or the sequence restrictions.

problem of the FFT [10]. An example of an ictal EEG in the delta band with marked waves as sequence is shown in Figure 3.

4) Intra-sequence correlation

The morphology of the waves in the sequence is used to separate artifacts from ictal patterns based on the observation that ictal sequences of epileptogenic discharges look similar to each other. A similarity value γ is calculated that models the similarity in a group of waves. Note that the morphology is not pre-defined but only must be similar in the sequence. This will avoid correlation artifacts as in cosine or wavelet transforms where a signal needs to be decomposable into the respective signal forms. The similarity value γ is calculated using the signal of N extracted waves $w_n(t)$.

$$\gamma = \frac{var_\tau\left(E_n\left[w_n(\tau)\right]\right)}{E_n\left[var_\tau\left(w_n(\tau)\right)\right]} \quad n \in 1..N, \tau \in \tau_n^{max}..\tau_{n+1}^{max} \tag{4}$$

The measure γ will then be used as replacement for the PWI feature in the EpiScan algorithm as described in [2].

C. Performance analysis

1) *Sensitivity:* The detection sensitivity was evaluated as follows: Each marker of electrographically visible seizures that intersects with a seizure alert from the algorithm is regarded as true positive event, whereas each seizure marker with no intersection is a false negative event. For each patient with recorded seizures the sensitivity is determined as the ratio of true positives and the total number of recorded seizures. We evaluate these sensitivities by calculating the mean over all patients with seizures. Averaging over patient-wise sensitivities is done since seizure counts of the patients are not equally distributed.

2) *False alarm rate:* The false alarm rate is also calculated patient-wise. Long contiguous markers from an automatic seizure detector create a higher review effort than short ones that can be inspected on a single EEG screen. In order to accommodate this fact, each seizure alert is divided into multiple sub-markers of maximally 30 seconds, meaning that each of these markers contributes to the false alarm rate. Each sub-alert that does not intersect with a true seizure marker (basic truth) is regarded as a false alarm. The number of false alarms for one patient divided by the total number of hours of EEG recordings for this patient gives the false alarm rate. False alarm rates are evaluated by calculating the mean over all patients.

D. Test set

The EEG database of the AIT [1] was used to validate the seizure detection performance. The database includes solely uncut and unselected EEG long-term recordings from several epilepsy monitoring units, mostly in 256 Hz sampling rate using the standard 10-10 or 10-20 international electrode system.

AIT EEG Database	Measure
# Patients	275
# Patients with epilepsy	159
# Patients with seizures	96
# Epileptic seizures	623
# Hours of EEG recordings	22463

Table 1 Overview of the AIT EEG database

III. Results

A. EWS Detection Performance

The EWS algorithm reached 100% detection sensitivity in a third of the patients (N=31). The mean of the detection sensitivity using all patients with all epilepsy types (N=96) is 53%. The overall false alarm rate of all patients (N=275) was 0.14 false alarms per hour (FA/h).

B. Combined EpiScan and EWS Detection Performance

To optimize the detection performance without further increasing the false alarm rate (compared to [4]) a PWA setting with reduced sensitivity was combined with the EWS algorithm. Figure 4 draws the detection sensitivity as function of the false alarm rate showing that the combination with the EWS algorithm is preferable to a further increase of the PWA sensitivity. The advantage of the combined version is an increase of 3% in sensitivity and a reduction of 0.034 false alarms per hour (FA/h) giving absolute values of 73% overall sensitivity (N=96) with an overall false alarms rate (N=275) of 0.3 FA/h. The combination leads to an improvement because irregular ictal patterns are detected more efficiently with the EWS algorithm.

A histogram showing the detection performance of the combined method is given in Figure 5. Note that the

majority of patients (N=52) had a detection sensitivity of 100%. The important subgroup of temporal lobe epilepsy (TLE) patients (N=61) is detected with high sensitivity of 83.6% and a false alarm rate of 0.29 FA/h.

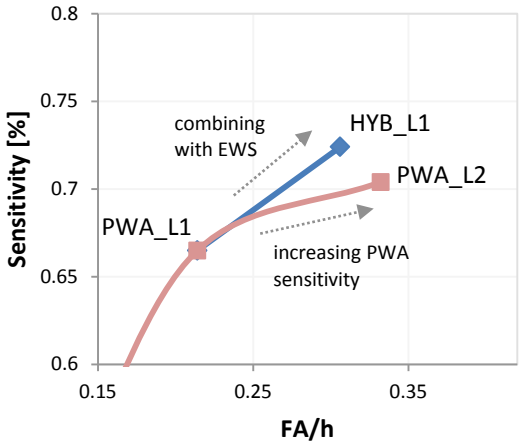

Figure 4 The plot of the algorithm operating characteristic (AOC) shows the overall detection sensitivity in percent against the false alarm rate per hour (FA/h)**.** The marker PWA_L2 correspond to EpiScan algorithm solely based on PWA published in [4]. The EWS algorithm is combined with the setting PWA_L1 to raise sensitivity while reducing FA/h compared to PWA_L2, giving the marker HYB_L1.

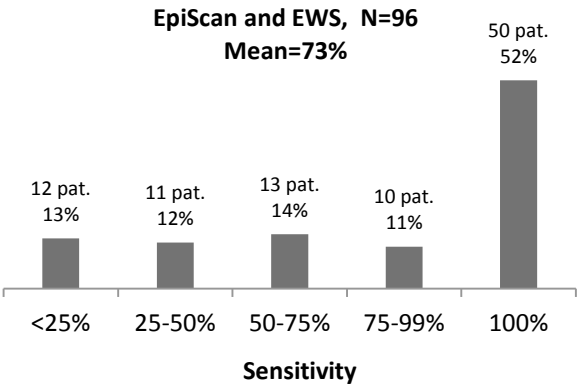

Figure 5 The histogram of the combined detection performance using the PWA_L1 and the EWS algorithm. Most of the 96 patients with seizures have a detection performance of 100%, patients having lower detection sensitivities are nearly equally distributed. The mean of the detection sensitivity is 73%.

IV. Discussion

A time domain approach of an epileptic seizure detector called EWS was presented that showed its effectiveness in detection of rhythmic seizure patterns with moderate irregular morphology or signal distortions. The combination with the EpiScan algorithm leads to a new hybrid detection scheme with a performance that could not be reached by one algorithm alone. An overall detection sensitivity of 73% while having a false alarm rate of 0.3 was reached that correspond to an increase of 3% in sensitivity and a reduction of 0.034 in false alarm rate compared to [4]. The performance of the important group of TLE patients reached a sensitivity of 83.6% with a false alarms rate of 0.29 alarms per hour.

The analysis of the problems and results imply that many feature extraction schemes working on biomedical signals face similar problems and that they will benefit from hybrid algorithm approaches as the strengths of both viewpoints are needed to bring performance to new levels.

V. References

[1] T. Kluge, „EpiScan“, *Advanced Algorithms for Brain Signal Analysis*. [Online]. Available: http://www.ait.ac.at/EpiScan.

[2] M. Hartmann, F. Fürbaß, H. Perko, A. Skupch, K. Lackmayer, C. Baumgartner, und T. Kluge, „EpiScan: Online seizure detection for epilepsy monitoring units“, in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.

[3] F. Fürbaß, H. Perko, M. Hartmann, C. Baumgartner, und T. Kluge, „Detection of Epileptic Seizures During Pre-Surgical Evaluation using Rhytmic EEG Patterns“, in *EFNS Budapest 2011*, Budapest, 2011.

[4] F. Fürbaß, H. Perko, M. Hartmann, G. Gritsch, C. Baumgartner, und T. Kluge, „Improved online seizure detection system for epilepsy monitoring units“, in *ENS 2012*, Prague, 2012.

[5] T. Kluge, M. Hartmann, C. Baumgartner, und H. Perko, „Automatic Detection of Epileptic Seizures in scalp EEG-Recordings Based on Subspace Projections“, *Epilepsia*, Bd. 50, Nr. 11, S. 26–27, Nov. 2009.

[6] J. G. Nicholls, A. R. Martin, B. G. Wallace, und P. A. Fuchs, *From Neuron to Brain: A Cellular and Molecular Approach to the Function of the Nervous System, Fourth Edition*, 4 Sub. Sinauer Associates, 2001.

[7] D. J. Koffler und J. Gotman, „Automatic detection of spike-and-wave bursts in ambulatory EEG recordings“, *Electroencephalogr Clin Neurophysiol*, Bd. 61, Nr. 2, S. 165–180, Aug. 1985.

[8] J. Gotman, „Automatic recognition of epileptic seizures in the EEG“, *Electroencephalogr Clin Neurophysiol*, Bd. 54, Nr. 5, S. 530–540, Nov. 1982.

[9] J. Gotman und P. Gloor, „Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG“, *Electroencephalography and Clinical Neurophysiology*, Bd. 41, Nr. 5, S. 513–529, Nov. 1976.

[10] A. Bruns, „Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches?“, *Journal of Neuroscience Methods*, Bd. 137, Nr. 2, S. 321–332, Aug. 2004.

6.2 Paper A2: EpiScan study

Title: Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring

Authors: Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, Elezi L, Pataraia E, Colon AJ, Baumgartner C, Kluge T.

Published in: Clinical neurophysiology

Year: 2015

Authors' contribution: Fürbass Franz developed the EWS algorithm and the patient adaptation which is part of the EpiScan seizure detector algorithm in this work. In addition data evaluation, statistical analysis, and comparison study to Persyst was done by Franz Fürbass. Initial writing of the manuscript, implementation of corrections from coauthors and reviewers was done by Franz Fürbass. Data annotations, data management, all clinical work and other algorithms included in EpiScan were implemented by coauthors.

Clinical Neurophysiology 126 (2015) 1124–1131

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units

F. Fürbass [a,*], P. Ossenblok [c], M. Hartmann [a], H. Perko [a], A.M. Skupch [a], G. Lindinger [d], L. Elezi [b], E. Pataraia [d], A.J. Colon [e], C. Baumgartner [b], T. Kluge [a]

[a] Department Safety & Security, AIT Austrian Institute of Technology GmbH, Vienna, Austria
[b] General Hospital Hietzing with Neurological Center Rosenhuegel, 2nd Neurological Department, Vienna, Austria
[c] Clinical Physics, Academic Center of Epileptology, Heeze & Maastricht UMC+, The Netherlands
[d] Department of Clinical Neurology, Medical University of Vienna, Vienna, Austria
[e] Department of Clinical Neurophysiology, Academic Center of Epileptology, Heeze & Maastricht UMC+, PO Box 61, 5590 AB Heeze, The Netherlands

ARTICLE INFO

Article history:
Accepted 24 September 2014
Available online 2 October 2014

Keywords:
Epileptic seizure detection
Automatic
Online
Parameter free
Alarm device
Prospective multi-center study

HIGHLIGHTS

- Large prospective multi-center study of an automatic seizure detection system including 205 patients.
- Comparison between two automatic seizure detection systems using the same prospectively recorded dataset.
- Performance numbers on the publicly available CHB–MIT dataset and on 310 retrospective patients datasets.

ABSTRACT

Objective: A method for automatic detection of epileptic seizures in long-term scalp-EEG recordings called EpiScan will be presented. EpiScan is used as alarm device to notify medical staff of epilepsy monitoring units (EMUs) in case of a seizure.
Methods: A prospective multi-center study was performed in three EMUs including 205 patients. A comparison between EpiScan and the Persyst seizure detector on the prospective data will be presented. In addition, the detection results of EpiScan on retrospective EEG data of 310 patients and the public available CHB–MIT dataset will be shown.
Results: A detection sensitivity of 81% was reached for unequivocal electrographic seizures with false alarm rate of only 7 per day. No statistical significant differences in the detection sensitivities could be found between the centers. The comparison to the Persyst seizure detector showed a lower false alarm rate of EpiScan but the difference was not of statistical significance.
Conclusions: The automatic seizure detection method EpiScan showed high sensitivity and low false alarm rate in a prospective multi-center study on a large number of patients.
Significance: The application as seizure alarm device in EMUs becomes feasible and will raise the efficiency of video-EEG monitoring and the safety levels of patients.

1. Introduction

Long-term video EEG-monitoring in epilepsy monitoring units (EMUs) plays a central role in pre-surgical evaluation of patients with epilepsy (Smith, 2005). This time-consuming procedure lasting for several days up to weeks requires high effort from staff to ensure patient safety and to evaluate the high amount of data. Safety in EMUs is an on-going discussion. It is generally accepted that precautions have to be in place to promptly detect seizures (Carlson, 2009) and to avoid additional harm to the patients. A study by Atkinson et al. (2012) with $N = 20$ patients showed that only 40% of seizures showed staff response. Changing the safety

* Corresponding author at: AIT Austrian Institute of Technology, Donau-City-Straße 1, 1220 Vienna, Austria. Tel.: +43 50550 4230; fax: +43 50550 4125.
E-mail address: franz.fuerbass@ait.ac.at (F. Fürbass).

http://dx.doi.org/10.1016/j.clinph.2014.09.023
1388-2457/

protocol for EMUs can thus lead to a decrease in patient accidents and an increase in detected seizures (Spanaki et al., 2012).

Automatic epileptic seizure detection (ESD) is one method to improve patient safety and efficiency in the EMU. Although these systems have a long history of numerous methodical approaches that proved to be effective in some trials (Gotman and Gloor, 1976; Gotman, 1982, 1990) wide spread clinical application were not accomplished until now. Today, the small number of epilepsy monitoring units using such systems stays in contrast with the increasing awareness of patient security issues during long-term recording and the high costs of this examination method. A low false alarm rate is of major importance for alarm systems to avoid ignorance by staff as found by Lee and Shah (2013). Many epilepsy centers do not use automatic seizure detection systems because of a very high number of false detections.

Several publications proposed patient specific seizure detectors or detectors for certain seizure patterns (Beniczky et al., 2013). These approaches will be of limited value in clinical practice because details of the type of epilepsy or the localization of the seizure onset zone (SOZ) are mostly unknown. Attempts to use the first seizure of a patient for patient specific detectors are limited because of the long time delay to the first seizure. Several studies reported a delay between 2 and 3.7 days in EMUs for pre-surgical evaluation, depending on the type of epilepsy (Todorov et al., 1994). In addition, the average number of seizures that can be recorded in one week of video EEG is rather small (median of 3 in one week in our data). Furthermore, it is important to detect whether or not a patient has one or multiple types of seizures. This implies that detection systems cannot be efficiently trained or configured for patients in the EMU and that only parameter-free detection systems without restriction to seizure types are feasible.

Automatic analysis of the EEG can be done either ad-hoc during the recording of the patient or post hoc after the patient recoding has finished. These situations are also referred to as "online" or "offline" detection, respectively. This article will solely present results of the online seizure detector EpiScan but the major differences to offline detectors are depicted shortly to allow objective comparison to other publications. First of all, because online detection systems may be used as alarm devices whereas offline systems support the EEG evaluation after recording. Furthermore, online detection systems must have a very short time delay to trigger alarms. An artificially delayed alarm allows the collection of information about the trend of the supposed seizure and can avoid false alarms. A system reacting in the range of a few seconds is more close to an alarm device, whereas a system with a detection delay of several minutes or hours behaves like a typical post hoc system. When comparing the performance of ad-hoc to post hoc systems or ad-hoc systems with different delays care has to be taken.

The amount and kind of data to evaluate an automatic seizure detection system is an important and frequently discussed issue. A sufficient number of long-term patient recordings are needed in order to draw reliable conclusions about sensitivity, specificity or the differentiation between two competing systems or datasets. One critical point in assessment of seizure detectors is the estimation of the sensitivity. Seizures are rare events with high inter- and intra-patient variability. The detection sensitivity of an automatic system represents a random variable with high variance and unknown distribution. In statistics the central limit theorem states that a sampling distribution approaches the normal distribution if the sample size is sufficient, no matter how the population distribution was shaped. A sample size of $N = 30$ is considered as appropriate for moderately skewed population distributions and will give a rough estimate of the performance. Population distributions far from normal need a sample size of $N = 500$ or more. For the sensitivity and false alarm rate of a seizure detection system we cannot assume a distribution close to normal and thus have to carefully determine the amount of data necessary to get significant results.

However, sensitivity based on a high number of patients alone does not validate a clinical application if only parts of the recordings are used. Only complete and uncut datasets reflect the real clinical situation and can prove sensitivity and specificity at the same time. A detection system may easily be able to detect 100% of the seizures in a dataset when only ictal EEG fragments are used but will show an excessive false alarm rate when evaluated on full long-term recordings. In addition, changes of the EEG during the day/night cycle need to be included in the evaluation leading to a necessary continuous recording length of more than 24 h.

The Computational Encephalography research group (www.eeg-vienna.com) of the Austrian Institute of Technology (AIT) has developed an automatic seizure detection system for long-term scalp EEG recordings called EpiScan. The detection algorithm of EpiScan works as an alarm device which allows notification of medical staff in case of a seizure. The system does not require parameters or patients specific settings. In this article the results of a prospective multi-center study will be presented. The results of EpiScan will be compared to the results of the Persyst seizure detector using the same prospective dataset. A comparison to the EpiScan performance on the development dataset and the MIT–CHB dataset will be carried out.

2. Methods

2.1. Data analysis

EpiScan is based on a computational method, which automatically detects epileptic seizures in digitized EEG. This method was developed over several years by a team of physicians, mathematicians and medical experts (Schachinger et al., 2006; Perko et al., 2007; Kluge et al., 2009; Hartmann et al., 2011; Fürbass et al., 2012). It is intended to analyze the EEG ad-hoc and to act as an online detection system. The EpiScan method analyses the digital EEG during recording in intervals of a quarter-second. Frequencies below 0.7 Hz and above 99 Hz are removed by finite impulse response filters. Line noise is removed with notch filters at 50 and 60 Hz. EEG segments with artifacts like i.e. excessive amplitudes or artifacts from loose electrodes are removed automatically (Skupch et al., 2013) and are not used for detection. This will avoid false alarms based on measurement problems. The EEG is then scanned for rhythmic patterns in the time and frequency domain by algorithms called Epileptiform Wave Sequence Analysis (EWS) and Periodic Waveform Analysis (PWA), respectively (Hartmann et al., 2011; Fürbass et al., 2012). An energy detector scans for tonic or tonic–clonic seizures with strong muscle artifacts. All extracted features are normalized by a spatio-spectral model of the brain activity that is continuously updated by past information from the EEG. A set of classifiers is used to remove events with physiological origin. The use of these adaption and classification algorithms avoids repeated detections of physiological or pathological patterns that are no seizures and is therefore another important mechanism to avoid excessive false alarms. The parameters of the classifiers were optimized using an automatic parameter optimization method (Dollfuss et al., 2013).

2.2. Quantity and quality of data needed for evaluation

The amount of data in a study is a critical parameter for the reliability of the results. Standard measures in statistics like i.e. the mean or confidence intervals of a result assume a sufficient high number of replicates in order to be valid estimates. An objective estimate of the number of participants for a seizure detector study

is hardly feasible but it is easy to show that $N < 30$ is too low. Viewing the problem from the neurophysiological perspective it has to be considered that epilepsy is a symptomatic disease with numerous etiologies. Thus 30 patients will not adequately represent the full range of possible manifestations (see experiment using virtual trials in supplementary data).

2.3. Definition of seizures

An epileptic seizure is defined as a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain (Fisher et al., 2005). Although the EEG is an important tool for epilepsy diagnosis this definition does not state how well the seizure activity can be identified in the EEG. Some seizures are hardly recognized without using additional information from video or other clinical information because of artifact overlap or subtle EEG patterns. Furthermore, clinical practice often includes subtle seizure-like events in the list of seizures to support the neurologist.

In order to remove the bias of clinical procedures and EEG measurement issues, it is common to restrict the evaluation of EEG-based epileptic seizure detection systems on clearly visible electrographic seizures. In such an approach one or several experienced EEG reviewers select seizures according to a visual perception value and define seizure onset and duration. Such a pre-sorting of seizures is preferable during development of the detection system but not appropriate in the clinical practice of EMUs.

To take this into account we used a two-step procedure to evaluate our seizure detection system. In a first step the detection performance of EpiScan was assessed using seizures defined by clinical and electrographic observations without restriction to EEG correlates. This includes all seizures that were marked during recording of the EEG and seizures that were found retrospectively by the standard EMU review procedure using video, EEG, and observation reports from nursery staff. Results using seizures from this first step are referred to as $C + E$ evaluation group. In a second step the detection performance for seizures with different levels of EEG perception value (P) will be given. The perception values of seizures were assigned from experienced EEG technicians in several video-blinded reviewing sessions. The reviewers were asked to decide whether the EEG at a defined time point shows a seizure. They were allowed to switch montages and to review the EEG before and after the given time point. The possible answers included six levels of increasing perceptions values (see Table 1). Based on these six perception values data were divided to form four groups called $C + E$, E75, E50, and E25 for the evaluation of the seizure detection systems. The evaluation groups E75, E50 and E25 include all seizures that had at least at perception value of >75%, >50% or >25%, respectively (see supplementary data for examples). Reviewers will often use the middle of a scale or 50% if they are uncertain about the decision. This case was avoided by forcing a decision between "rather a seizure" (rather yes) or "rather not a seizure" (rather no).

2.4. Dataset

2.4.1. Prospective multi-center study

A prospective multi-center study was performed to evaluate the seizure alarm system EpiScan. During the study, long-term EEG recordings from 205 consecutive patients were evaluated. Data were recorded at three epilepsy-monitoring units, the 2nd Neurological Department of the General Hospital Hietzing with Neurological Center Rosenhuegel in Vienna (NCR), the Department of Clinical Neurology of the Medical University of Vienna (MUV) and the Epilepsy Center Kempenhaeghe in Heeze, the Netherlands (KEMP). Data were recorded between January 2012 and March 2013. All centers used the international 10–20 electrode placement system for data recording. The data was recorded using a sampling rate of 256 Hz in center MUV and NCR and a sampling rate of 200 Hz in center KEMP. The inclusion criteria were a signed patient agreement form and an age above 18.

An ITmed EEG recording machine was used in center MUV. Patients had to stay in bed to allow video-EEG in this facility. The center NCR uses a Micromed recording system including a headbox with internal memory that allowed unplugging of several minutes without loss of EEG. Due to technical reasons it was not possible to use the EEG in the unplugged time periods for the study (about 3% of the recorded data at NCR). The center KEMP uses a Stellate recording device with long patch cables. Here all patients stayed in a living-room like environment that allowed free movement. They were able to use fitness devices or the bathroom without disruption of the EEG recording. This environment induced lots of additional movement artifacts namely from cycling, chewing, and tooth brushing making this dataset especially challenging to analyze for an automatic detection system. In all three centers AEDs were withdrawn preceding or during the five day period of video-EEG monitoring depending on the patient. The amount of data that was collected at each center as well as number of patients is summarized in Table 2.

2.4.2. Retrospective data

EpiScan was developed using a dataset of 310 patients. This dataset will also be evaluated in this article to further increase

Table 2
Overview of the EEG data included in the prospective multicenter study. In total 205 patients participated, including 94 patient with seizures. The number of seizures and hours of recorded EEG is given.

Epilepsy center	N	N with sz.	Number of sz.	Hours of EEG
EEG data of the prospective study				
NCR	83	27	142	6513
KEMP	60	47	211	5127
MUV	62	20	173	4044

Table 1
The EEG perception value (P) of a seizure. Seizure markers received from EMUs have no perception value ($C + E$). A seizure perception value is assigned through EEG reviewing, the higher the more clearly a pattern was perceived as seizure in the EEG. An evaluation group includes all seizures that have at least a given minimum perception value.

Q: Is this EEG pattern a seizure?		Evaluation groups with included seizure			
Possible answer	P (%)	E75	E50	E25	$C + E$
Seizure perception value					
Surely yes	100–90	x	x	x	x
Probably yes	90–75	x	x	x	x
Rather yes	75–50		x	x	x
Rather no	50–25			x	x
Probably no	25–10				x
Surely no	10–0				x

the statistical relevance of the results and to show differences and similarities between a retrospective and prospective dataset. The development dataset was recorded at several different EMUs using the international 10–20 electrode placement system at a sampling rate of 256 Hz. The dataset included 693 markers that were a mixture of relevant information for the diagnosis and real seizures. The EMU review procedure on this development data had not been standardized for the use in a clinical study. Therefore a seizure evaluation group $C+E$ is undefined and thus no evaluation on $C+E$ seizures could be done. The dataset was evaluated retrospectively and the same protocol as for the prospective dataset to define seizures for the E25, E50, and E75 evaluation groups was applied.

Although EpiScan was developed and tested on data from patients with age above 18 an evaluation on a small pediatric dataset gives a first insight if a clinical application will be feasible. The CHB–MIT scalp EEG database was used. It was created by a team of investigators from Children's Hospital Boston (CHB) and the Massachusetts Institute of Technology (MIT) and is publicly available from the Physionet website (http://www.physionet.org/physiobank/database/chbmit/). The database includes data from 24 patients from 1.5 years to 22 years of age and a mean of 10 years (Goldberger et al., 2000; Shoeb, 2009; Hunyadi et al., 2012). Seizure markers were uses as given in the dataset without assigning perception values. Unfortunately, the CHB–MIT data is given in bipolar longitudinal montage only. A full montage set with other reference electrodes cannot be restored from this information. This will have a negative influence on the detection performance.

To mimic the behavior of prospective data, all retrospective patient recordings were used in their full length without restriction. No file selection and no segmentation of patient data were applied. The number of patients with seizures was defined based on the lowest available seizure perception value for each dataset. For the retrospective dataset E25 was used, for the prospective study data $C+E$ was used to define the number of patients with seizures. Table 3 gives an overview of the complete dataset used for this publication.

2.5. Definition of detection performance

EpiScan alarms were compared to manually defined seizure markers. Analysis was done separately for the four evaluation groups $C+E$, E25, E50, and E75. A seizure epoch was defined as a three minutes time range starting from the beginning of the seizure marker. An EpiScan alarm occurs on a specific time point without having time duration. An EpiScan alarm was defined as true positive (TP) when it appeared within a seizure epoch. Several EpiScan alarms in one seizure epoch were defined as one TP. Alarms outside of a seizure epoch were defined as false positives (FP). False positives occurring within a time span of less than 30 s were counted as a single false alarm. A seizure epoch without a matching EpiScan alarm was defined as false negative (FN). Fig. 1 summarizes these definitions.

The sensitivity of the automatic detection was calculated for each patient. It was defined as the ratio between the numbers of true positives (TP) to the number of all seizures (TP + FN). The false alarm rate of the automatic seizure detection was defined by the number of false alarms in 24 h (FA/24 h).

$$\text{Sensitivity} = \frac{\#\text{TP}}{\#\text{TP} + \#\text{FN}}$$

$$\text{FA/24h} = \frac{\#\text{FP}}{\text{duration of recording days}}$$

2.6. Comparison to Persyst seizure detection

Currently, Persyst is considered to be the most prevalent seizure detection system. We compared the results from the prospective study data with the results obtained from the Persyst seizure detection in Version 12 (Version 12, Rev. B, 2012.11.27, http://www.persyst.com/). All EEG datasets were converted to EDF format (http://www.edfplus.info/) with a maximum of 99,000 cycles per file. The correctness of the files was validated with the Polyman EDF checker (http://www.edfplus.info/downloads/). Each EDF file was processed separately with the Persyst seizure detection engine. The results of the "SzDetect" table were manually copied to Microsoft Excel tables. These tables were automatically analyzed by reading them into Matlab. The values for sensitivity and false alarm rate were calculated with the same procedures that were used for the EpiScan results.

2.7. Statistical methods

All confidence interval (CI) values are calculated using a parameter-free bootstrapping method for confidence intervals with 1000 bootstrap samples as described in DiCiccio and Efron (1996). The two-sample t-test was used to validate whether two samples came from a distribution with the same mean. The default alpha value was 0.05.

3. Results

3.1. Detection results of the prospective multi-center study

An overview of the dataset collected in the prospective multicenter study is given in Table 2. In total, 15,684 h of EEG including 205 patients with 526 seizures were evaluated. The results for the different perception values are shown in Fig. 2. For those seizures where all reviewers agreed on "probably yes" or higher (E75) EpiScan showed a mean sensitivity of 81% (95% confidence interval = 74–86%) combined with a false alarm rate of 7.1 false alarms per day. As expected including more ambiguous EEG pattern and thus lower perception value for the seizures led to a decrease in sensitivity. With a perception value of E50 ("rather yes" or higher) a mean sensitivity of 78% (95% confidence interval = 70–84%) with a false alarm rate of 7.08 per day was reached.

When calculating the average sensitivity for all 94 patients with seizures regardless of whether the seizure is visible in the EEG or not ($C+E$) we achieved a mean sensitivity of 72% (95% confidence

Table 3
Data used to evaluate EpiScan: Data of the prospective multicenter study (Study), the retrospective data of the development dataset (Devel), the public available pediatric dataset from MIT (CHB–MIT), and the cumulative dataset including all (ALL). The number of patients and the subgroup of the number of patients with seizures (*N* with sz.) are given. The total number of seizures and hours of recorded EEG are shown in the last two columns.

Dataset name	Prosp./Retrosp.	*N*	*N* with sz.	Number of sz.	Hours of EEG
EEG data for EpiScan evaluation					
Study	P	205	94	526	15,684
Devel	R	310	124	1113	25,567
CHB–MIT	R	24	24	197	1355
All	P/R	539	242	1836	42,594

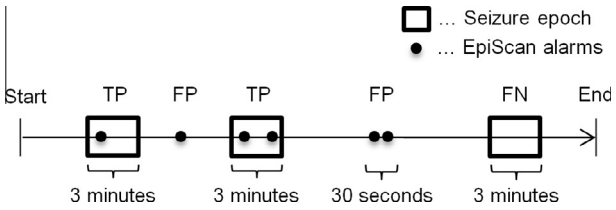

Fig. 1. Definition of true positives (TP), false positives (FP) and false negative (FN) detections by comparing seizure epochs defined by epileptologists to EpiScan alarms.

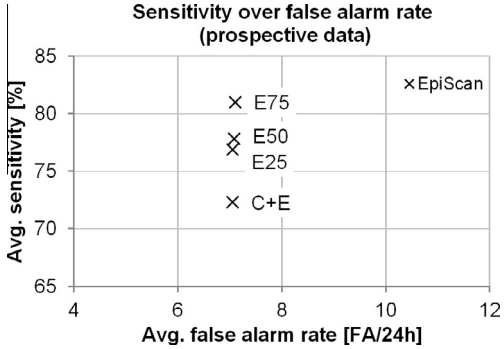

Fig. 2. Results of the prospective multi-center study of EpiScan for different levels of seizure perception values ($C+E$, E25, E50, E75). A steady increase in detection sensitivity can be observed for seizures groups with higher reviewer perception values.

interval = 65–79%). The false alarm rate reduced slightly to 7.05 false alarms per day.

Some seizures were not detected by the clinical protocol of the EMUs at the different centers. Reasons were subtle clinical signs, strong artifact superposition, unobtrusive visual EEG patterns or they were simply overlooked by the reader. During the prospective study, EpiScan detected 16 (3% of all seizures) previously undetected seizures.

We found no statistically significant difference in detection sensitivity between the three participating centers ($p > 0.06$) which shows the robustness of EpiScan against influences of different recording setups.

We further investigated how sensitivity and false alarm rate of individual patients are distributed in the dataset. This gives more insight into the performance of the detection system. We divided the patients in five groups: group one contained patients where EpiScan detected less than 25% of the seizures. Group two contained those patients where 25–50% of the seizures where detected. Group three, four and five contained patients where more than 50%, more than 75% and 100% of the seizures were detected, respectively. A histogram of the $C+E$ detection sensitivities is shown Fig. 3. The results were plotted separately for each center and for the complete dataset. The histograms were normalized to the number of patients of the given center to allow comparison of the results between the different recording sites. Fig. 3 reveals that in more than half of the patients 100% of the seizures were detected. The distribution of the detection sensitivity had a very similar pattern in all centers proving a very stable detection quality for different patient cohorts and recording conditions.

A similar analysis was performed for the false alarm rate. Patients were divided into groups with a false alarm rate per day of less than one, between one and five, five and ten, and between ten and 24 false alarms per day. Fig. 4 shows a histogram for all four centers as well as the combined data. The difference of the false alarm rate from center MUV was statistically significant ($p < 0.05$) compared to the other two centers.

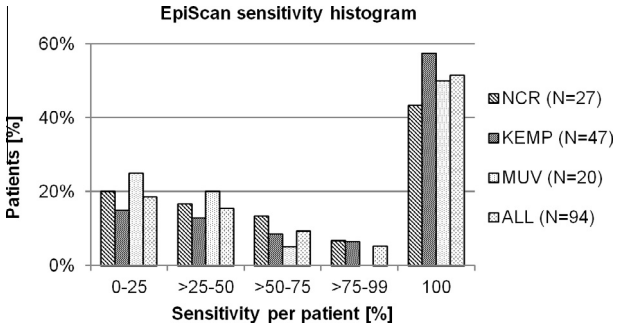

Fig. 3. Histogram of EpiScan detection sensitivities in the prospective study using all seizures ($C+E$). The normalized histograms of the three individual centers (NCR, KEMP, MUV) shows that more than half of the patients were detected with 100% sensitivities.

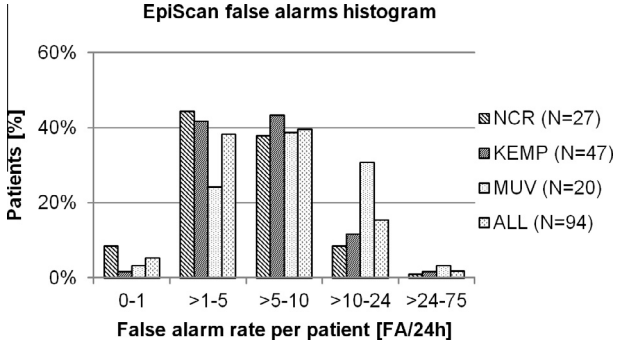

Fig. 4. Histogram of EpiScan false alarms of the three individual centers (NCR, KEMP, and MUV) and combined (ALL) normalized to percent. The mean values are: NCR = 5.1, KEMP = 6.9, MUV = 9.8, ALL = 7.05 FA/24 h.

3.2. Comparison with Persyst seizure detection

In addition to the analysis of the seizure detection system EpiScan we also performed an analysis of the Persyst seizure detector. Performance of EpiScan and Persyst 12 are presented for the data of the prospective study using all patients with seizures ($N = 94$). A comparison of the detection performance for different seizure perception values is given in Fig. 5. The results show an increase of the performance with increasing visibility of the seizure in the EEG for both detectors. For the clearly visible electrographic seizures in the E75 seizure group we found a sensitivity of 81% for EpiScan compared to 75% for the Persyst seizure detection. Similar increases in sensitivity for EpiScan were found for all other groups (+3.7% sensitivity for $C+E$, +6.2% sensitivity for E25, +5.5% sensitivity for E50, $p < 0.76$ for all groups). In addition to the higher sensitivity we found lower false alarm rate for the EpiScan seizure detector (−27% or −2.68 FA/24 h) compared to the Persyst seizure detector for all detection groups. The differences in sensitivity and false alarm rate are not statistically significant using a paired sample t-test.

3.3. Results on retrospective datasets

3.3.1. Results of the development dataset

The results on the development dataset, which contains an extraordinary high number of patients (Table 3), will be shown to further raise the reliability of the EpiScan detection performance. Here, correlation on patient diagnosis will be presented that is not yet available for the prospective study data.

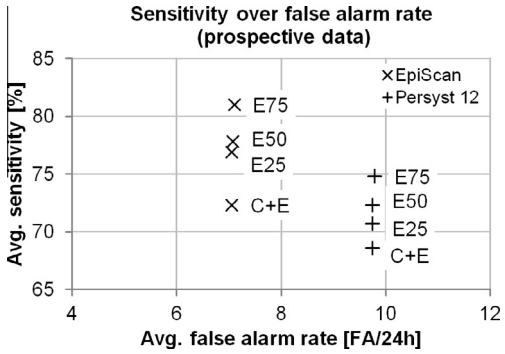

Fig. 5. Comparison of EpiScan and Persyst 12 detection performance using the prospective dataset showing the superior performance of EpiScan compared to Persyst 12. The performance increases with increasing level of perception value which is true for both seizure detectors. An average sensitivity of 81% is reached by EpiScan for unequivocal electrographic seizures (E75), compared to 75% reached by Persyst.

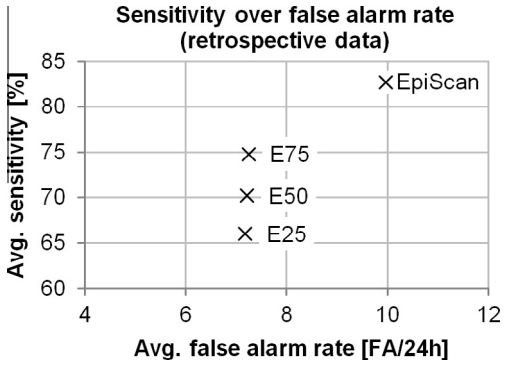

Fig. 6. EpiScan detection performance on the development dataset. The variation of the seizure perception value shows that detection performance increases if subtle electrographic seizures are removed from evaluation.

(Fig. 6) depicts the detection performance of EpiScan for this dataset. Results for the three perception value groups E25, E50 and E75 are shown. As for the prospective dataset, an increased seizure perception value (defined in Table 1) results in better detection performance. For perception value E75 the detection performance is 75% sensitivity with a false alarm rate of 7.2 FA/24 h.

We looked a possible correlation between the type of epilepsy and the detection performance of an automatic system. Table 4 compares the results for patients suffering from mesial temporal lobe epilepsy (mTLE), temporal lobe epilepsy (TLE), extra temporal lobe epilepsy (XTLE) and frontal lobe epilepsy (FLE). The best results of an average sensitivity of 87% were found for the subgroup with mTLE because of many regular rhythmic patterns during seizures. The TLE subgroup achieved an average sensitivity of 83% which is a very good result. The missed seizures are due to a few patients with neocortical seizure onset zone (SOZ) which often exhibited unique seizure patterns for each patient that were not always detected. A similar explanation for the lower detection sensitivity applies to the XTLE and FLE group, which also include some patients with neocortical SOZ. An additional problem of the FLE group was that a typical seizure shows high amplitude muscle artifacts but some seizures from certain patients lack this property and show only average amplitude artifact and no other obvious seizure activity in the EEG.

3.3.2. Results on pediatric EEG data

In addition to the large dataset recorded from adults a small pediatric dataset was analyzed. The detection performance of EpiScan on the MIT dataset had an average sensitivity of 67% (95% confidence interval from 53% to 79%) with 7.7 FA/24 h on average. No conclusions about the expected performance of EpiScan on pediatric data can be drawn as the amount of patients with seizures is too small.

Table 4
Average detection performance of EpiScan for patient groups with different diagnoses. The development dataset with unequivocal electrographic seizures (E75) were used. Mesial temporal lobe epilepsy (mTLE) showed the best results.

Diagnosis	Sensitivity [%]	False alarm rate [FA/24 h]
Average detection performance of EpiScan compared between patient diagnoses		
mTLE ($N = 11$)	87	6.6
TLE ($N = 52$)	83	6.7
XTLE ($N = 50$)	64	7.3
FLE ($N = 11$)	54	7.2
No epilepsy	–	7.2

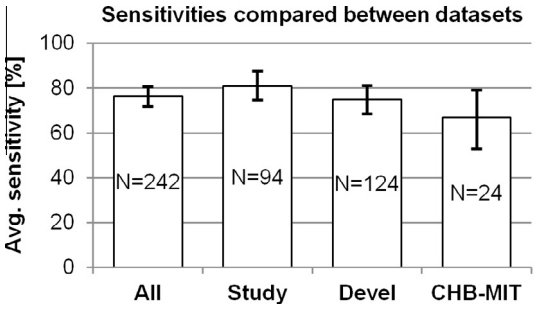

Fig. 7. Average detection sensitivities of EpiScan compared between different patients groups with different sample sizes. All sensitivities use the E75 seizure definition. At $N = 242$ the 95% confidence interval reduces to 8%, whereas a sample size of $N = 24$ will result in a range from 53% to 79%.

3.4. Cumulative meta-analysis of all available data

The prospective and all available retrospective data was used to assess statistical variables with high confidence level. The recording length, number of patients and number of seizures of all patients are listed in Table 3. Fig. 7 shows the nearly invariant detection performance of the different datasets. When combining all data into a large dataset the confidence interval reduces to 8.6% for all 242 patients with seizures. The difference in sensitivity between retrospective dataset (Devel) and prospective dataset (study) is small and statistically insignificant ($p = 0.91$). The average false alarm rate of all 539 patients was 7.2 false alarms in 24 h with a 95% confidence interval of 6.7–7.8 FA/24 h. No statistical significant difference between the different datasets could be found ($p > 0.38$ for all combinations).

4. Discussion

Seizure detection is an eagerly awaited feature in clinical practice of EMUs. We presented a multi-center study for the EpiScan online seizure detection system. EpiScan was tested as an online device in three different EMUs and its sensitivity and false alarm rate was calculated.

We were able to show that a sensitivity of 81% can be reached for seizures that are clearly visible in the EEG. We compared these results to the Persyst software, the most widely used seizure detection system. For the Persyst system we found a performance of about 75% showing that the new detection system EpiScan performs at least as good as the Persyst system.

The average sensitivity of EpiScan of about 81% was achieved for adult patients of age 18 and above. A small pediatric dataset

showed that EpiScan achieved comparable results also for EEG of children. No significant differences compared to the results of the development or prospective dataset were found. We concluded that an application of EpiScan on pediatric patients is also possible. However, the size of the pediatric dataset of $N = 24$ was too small to achieve a statistically valid comparison. More data are necessary to draw a reliable conclusion about the EpiScan performance for pediatric datasets.

A low false alarm rate is an essential feature of online seizure detection system. An alarm rate of several alarms per hour would render a seizure detection system useless regardless of the achieved sensitivity. We found an average false alarm rate of the EpiScan system of 7.1 false alarms per day. The comparison with the Persyst software showed that the Persyst system has a false alarm rate which is about 27% higher than that of EpiScan. As pointed out in the Methods section, false alarms within 30 s were counted as single false positive. An alarm occurs on a specific time point without having time duration. The concatenation of several false alarms within 30 s to a single false positive therefore corresponds to a false positive with maximal length of 30 s. The rational for this definition of a maximum length for a false alarm was that a reviewer should be able to determine if an EpiScan marker is a false alarm by looking at one single page of EEG. By restricting the maximum length of an artifact to 30 s we ensured that it would not be necessary to scroll through the EEG when classifying a marker as a false alarm. Variation of this maximal length parameter had only little impact on the false alarm rate. Increasing the time range for false alarms to 3 min will reduce the false alarm rate by 15%. We found a small increase in false alarm rate of less than 1% when the minimal perception value of the seizures was increased. This effect relates to the fact that EpiScan detected numerous seizures with a low perception value. When increasing the minimal perception value these seizure epochs do not longer count as true positive but add to the false alarms. Thus the marker converts to a false alarm according to the definitions in Fig. 1.

We found non-significant differences in sensitivity between the different centers that participated in the study. Only the center MUV showed a significantly higher false alarm rate of more than 10 false alarms per day due to artifacts appearing as epileptiform discharges during many of the recordings. The artifact contamination in the data from center KEMP was also high but did not raise the false alarm rate as in center MUV. A detailed analysis showed that most of the KEMP artifacts consisted of movement related electrode artifacts which did not trigger one of the seizure detection methods in EpiScan. Although some other artifacts like tooth brushing and movements during cycling did raise false alarms the incidence rate was too low to have a significate effect.

It is a question of debate if only a prospective clinical study of a seizure detection system can give a reliable proof of the performance. The results of our clinical study showed no statistically significant differences compared to the performance that was determined in offline experiments on a large development dataset. We did find a slightly increased performance in the study dataset that was due to the lower number of patients in the clinical study as compared to the offline-dataset. While usability issues can only be addressed in a clinical setting we argue that the use of large offline-datasets will give information about the performance of a seizure detection system that is as good as a prospective clinical study. The amount of data in an extensive offline analysis can be much larger than in any reasonable clinical trial of an automatic seizure detection system. It is of course necessary to ensure that the dataset used for offline-evaluation reflects the data found in an EMU. Taking for instance data from TLE-patients only, from seizure patients only, taking patients that only show a low level of artifacts or taking subsets of the datasets based on different inclusion criteria might strongly bias the results of an offline analysis. In addition it is important that large periods without seizures are included in every dataset in order to get a reliable estimation of both sensitivity and false alarm rate. Studies on many patients but very few hours of EEG only show the sensitivity but not the specificity of the method. In addition care must be taken not to over-fit the detection system to a given dataset by using the same small dataset for development and subsequent testing. Ideally the dataset for development and testing will be different. In practice a sufficient large dataset of several 10,000 h of data will also reduce this problem of over-fitting. In addition it is important to include complete 24 h recordings of the EEG in order to analyze the complete day/night cycle of a patient during the validation of a detection system.

We believe that patient safety in EMUs can be increased even by an imperfect seizure detection system. It is often assumed that medical staff will reach near 100% surveillance. This however is not always the case. A study by Atkinson et al. (2012) claimed that only 40% of the seizure showed a staff response. Although this number seems low, the general statement that a certain number of seizures in EMUs do not get immediate staff response is in line with our findings. Even at the large centers that participated in this study the automatic seizure detection system did find additional seizures that were not detected by the manual review procedures. The amount of missed seizures during recording depends heavily on the available staff in EMUs, their training and the time when a seizure occurs. Centers that do have highly trained staff available in their EMUs 24 h a day and that show a staff-patient-ratio of up to 1 technician for three patients during the day will be less likely to miss seizures. However, smaller hospitals often cannot afford this number of staff at their EMUs leading to an increased number of missed or un-responded seizures, especially during the night when even less staff is available. Here, an automated seizure detection system would be of great value by providing additional safety to the patient, since it will not depend on human factors, staff availability or time of day but on the quality of the visible EEG pattern only.

It is evident from all performed studies so far that an automatic seizure detection system will never replace the human EEG analysis. It will always be an additional source of information. Especially on highly specialized EMUs for pre-surgical evaluation such a system would be used in addition to today's procedures in order to find possible additional seizures that were missed by the visual analysis. However, even the larger centers are under constant pressure to cut costs. The available staff for EEG analysis generally decreases. Here, an automatic seizure detection system provides reliable monitoring of the EEG that will help to ensure the level of patient safety.

5. Conclusion

An automatic seizure detection and alerting system was validated in a prospective multi-center study and on retrospective data. In total 42,000 h of uncut long term EEG recordings were used to assess detection performance by means of sensitivity and false alarm rate on a high statistical confidence level. The results showed 81% sensitivity for seizures with high perception value at 7.1 false alarms per day in the prospective study. The analysis of 539 patient recordings showed no significant difference between prospective and retrospective detection results.

The multi-center study of EpiScan proved the ability of the system to work as seizure alarm device in the clinical long-term video EEG monitoring. The evaluation of the seizure detection performance on patient data with various diagnoses and ages showed the universal applicability of EpiScan. The comparison to the currently most prevalent seizure detection system from Persyst

showed that EpiScan reaches a lower false alarm rate on the collected prospective dataset but the difference was not of statistical significance. We conclude that the application of the EpiScan seizure detection system in EMUs could increase the efficiency and the safety level for patients.

Disclosure

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

Acknowledgements

Many thanks for the support of our work to all people involved and especially to Roy Krijn and to all staff members of the IT departments at AKH Vienna and the General Hospital Hietzing. Special thanks for the comprehensive and reliable EEG analysis to the technicians involved in the study, namely Harrie Geeris, Michaela Demel, and Julia Tarra.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.clinph.2014.09.023.

References

Atkinson M, Hari K, Schaefer K, Shah A. Improving safety outcomes in the epilepsy monitoring unit. Seizure J Br Epilepsy Assoc 2012;21:124–7.

Beniczky S, Polster T, Kjaer TW, Hjalgrim H. Detection of generalized tonic–clonic seizures by a wireless wrist accelerometer: a prospective, multicenter study. Epilepsia 2013;54:e58–61.

Carlson C. First do no harm: safety in the epilepsy monitoring unit. Epilepsy Curr 2009;9:162–3.

DiCiccio TJ, Efron B. Bootstrap confidence intervals. Stat Sci 1996;11:189–228.

Dollfuss P, Hartmann MM, Skupch A, Fürbass F, Kluge T. Automatic optimization of parameters for seizure detection systems. In: 35th Annu Int Conf IEEE Eng Med Biol Soc EMBC. 2013. p. 1976–9.

Fisher RS, van Emde Boas W, Blume W, Elger C, Genton P, Lee P, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia 2005;46:470–2.

Fürbass 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: 2012 Annu Int Conf IEEE Eng Med Biol Soc EMBC. 2012. p. 1020–3.

Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. Circulation 2000;101:e215–20.

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.

Gotman J, Gloor P. Automatic recognition and quantification of interictal epileptic activity in the human scalp EEG. Electroencephalogr Clin Neurophysiol 1976;41:513–29.

Hartmann MM, Fürbass F, Perko H, Skupch A, Lackmayer K, Baumgartner C, et al. EpiScan: online seizure detection for epilepsy monitoring units. Eng Med Biol Soc 2011 Annu Int Conf IEEE. 2011. p. 6096–9.

Hunyadi B, Signoretto M, Van Paesschen W, Suykens JAK, Van Huffel S, De Vos M. Incorporating structural information from the multichannel EEG improves patient-specific seizure detection. Clin Neurophysiol 2012;123:2352–61.

Kluge T, Hartmann M, Baumgartner C, Perko H. Automatic detection of epileptic seizures in scalp EEG-recordings based on subspace projections. Epilepsia [Internet] 2009 [cited 2011 Mar 15]; 26–27. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/j.1528-1167.2009.02377_1.x/abstract>.

Lee JW, Shah A. Safety in the EMU: reaching consensus. Epilepsy Curr 2013;13:107–9.

Perko H, Baumgartner C, Kluge T. Online algorithm for epileptic seizure detection. Lecture: annual meeting of swiss society of biomedical engineering. Neuchâtel, Schweiz; 2007.

Schachinger D, Baumgartner Ch, Kluge T. Bivariate phase space divergence: a measure for the automatic detection and anticipation of epileptic seizures in ECoG data. Lecture: international special topic conference on information technology in biomedicine. Ioannina, Greece; 2006.

Shoeb AH. Application of machine learning to epileptic seizure onset detection and treatment [Internet, thesis]. Massachusetts Institute of Technology; 2009 [cited 2013, July 9]. Available from: <http://dspace.mit.edu/handle/1721.1/54669>.

Skupch AM, Dollfuss P, Fürbass F, Gritsch G, Hartmann MM, Perko H, et al. Spatial correlation based artifact detection for automatic seizure detection in EEG. In: 2013 35th Annu Int Conf IEEE Eng Med Biol Soc EMBC. 2013. p. 1972–5.

Smith S. EEG in the diagnosis, classification, and management of patients with epilepsy. J Neurol Neurosurg Psychiatry 2005;76(Suppl. 2):ii2–7.

Spanaki MV, McCloskey C, Remedio V, Budzyn D, Guanio J, Monroe T, et al. Developing a culture of safety in the epilepsy monitoring unit: a retrospective study of safety outcomes. Epilepsy Behav 2012;25:185–8.

Todorov AB, Lesser RP, Uematsu SS, Yankov YA, Todorov AA. Distribution in time of seizures during presurgical EEG monitoring. Neurology 1994;44:1060–4.

6.3 Paper A3: Rhythmic and periodic pattern detection

Title: Automatic detection of rhythmic and periodic patterns in critical care EEG based on American Clinical Neurophysiology Society (ACNS) standardized terminology

Authors: F. Fürbass, M.M. Hartmann, J.J. Halford, J. Koren, J. Herta, A. Gruber, C. Baumgartner, T. Kluge

Published in: Neurophysiologie Clinique/Clinical Neurophysiology

Year: 2015

Authors' contribution: Fürbass Franz developed the algorithm for detection of rhythmic and periodic patterns including the time series segmentation algorithm. In addition data evaluation and statistical analysis was done by Franz Fürbass. Initial writing of the manuscript, implementation of corrections from coauthors and reviewers was done by Franz Fürbass. Data annotations, EEG recording, corrections, and the clinical evaluation were done by coauthors.

Neurophysiologie Clinique/Clinical Neurophysiology (2015) 45, 203—213

Disponible en ligne sur ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com/en

ORIGINAL ARTICLE/ARTICLE ORIGINAL

Automatic detection of rhythmic and periodic patterns in critical care EEG based on American Clinical Neurophysiology Society (ACNS) standardized terminology

Détection automatique de patterns rythmiques et périodiques dans l'EEG de soins intensifs basée sur la terminologie standardisée de l'American Clinical Neurophysiology Society (ACNS)

F. Fürbass[a,*], M.M. Hartmann[a], J.J. Halford[b], J. Koren[d], J. Herta[c], A. Gruber[c], C. Baumgartner[d], T. Kluge[a]

[a] *Austrian Institute of Technology GmbH (AIT), Safety & Security Department, Vienna, Austria*
[b] *Medical University of South Carolina, Comprehensive Epilepsy Center, Charleston, SC, USA*
[c] *Medical University of Vienna, Department of Neurosurgery, Vienna, Austria*
[d] *General Hospital Hietzing with Neurological Center Rosenhuegel, 2nd Neurological Department, Vienna, Austria*

Received 18 September 2014; accepted 5 August 2015
Available online 9 September 2015

KEYWORDS
ACNS ICU terminology;
EEG;
Automatic detection;
Rhythmic and periodic patterns;
Critical care

Summary
Aims of the study. — Continuous EEG from critical care patients needs to be evaluated time efficiently to maximize the treatment effect. A computational method will be presented that detects rhythmic and periodic patterns according to the critical care EEG terminology (CCET) of the American Clinical Neurophysiology Society (ACNS). The aim is to show that these detected patterns support EEG experts in writing neurophysiological reports.
Materials and methods. — First of all, three case reports exemplify the evaluation procedure using graphically presented detections. Second, 187 hours of EEG from 10 critical care patients were used in a comparative trial study. For each patient the result of a review session using the EEG and the visualized pattern detections was compared to the original neurophysiology report.

* Corresponding author at: Austrian Institute of Technology (AIT), Donau-City-Straße 1, 1220 Vienna, Austria. Tel.: +43(0) 50550 4230; fax: +43(0) 50550 4125.
E-mail address: franz.fuerbass@ait.ac.at (F. Fürbass).

http://dx.doi.org/10.1016/j.neucli.2015.08.001

Results. — In three out of five patients with reported seizures, all seizures were reported correctly. In two patients, several subtle clinical seizures with unclear EEG correlation were missed. Lateralized periodic patterns (LPD) were correctly found in 2/2 patients and EEG slowing was correctly found in 7/9 patients. In 8/10 patients, additional EEG features were found including LPDs, EEG slowing, and seizures.
Conclusion. — The use of automatic pattern detection will assist in review of EEG and increase efficiency. The implementation of bedside surveillance devices using our detection algorithm appears to be feasible and remains to be confirmed in further multicenter studies.

MOTS CLÉS
Terminologie ACNS
USI ;
EEG ;
Détection automatique ;
Motifs rythmiques et périodiques ;
Soins intensifs

Résumé
Buts de l'étude. — L'EEG continu (cEEG) des patients en unité de soins intensifs doit être évalué plus efficacement pour optimiser le traitement. Nous présentons une méthode informatique de détection de patterns rythmiques et périodiques. Celle-ci est basée sur la terminologie de soins intensifs (CCET) de l'American Clinical Neurophysiology Society (ACNS). Le but est de montrer que la détection de ces patterns permet aux experts d'écrire plus facilement des rapports neurophysiologiques.
Méthodes et matériaux. — Dans un premier temps, trois études de cas illustrent la procédure d'évaluation en utilisant des détections présentées graphiquement. Ensuite, 187 heures d'EEG venant de dix patients d'unités de soins intensifs ont été introduites dans une étude comparative. Pour chaque patient, le résultat d'une session de révision utilisant l'EEG et la détection des patterns a été comparé avec le rapport neurophysiologique original.
Résultats. — Parmi les cinq patients ayant eu des crises épileptiques, les crises de trois patients ont été reconnues correctement. Les deux autres patients avaient des crises cliniques très subtiles et sans corrélation claire dans l'EEG. Les patterns périodiques latéralisés (LPD) ont été correctement reconnus chez les 2 patients concernés et un ralentissement du EEG a été correctement reconnu dans 7/9 cas. Pour 8/10 patients des caractéristiques additionnelles ont été identifiées, incluant des patterns périodiques latéralisés, un ralentissement de l'EEG et des crises.
Conclusion. — L'utilisation d'algorithmes de détection automatique basés sur la CCET assisteront dans la révision de l'EEG et augmenteront son efficacité. L'implémentation de dispositifs de surveillance utilisant notre algorithme sera possible et sera montré dans de futures études multicentriques.

Introduction

Over the past decade, a considerable amount of research effort has been expended to study the prevalence in EEG of nonconvulsive seizures (NCS) or nonconvulsive status epilepticus (NCSE) in acutely ill patients. In 1999, a publication by Kaplan [14,15] showed that the extended use of continuous EEG (cEEG) revealed many patients with NCS/NCSE that would have been undiagnosed without cEEG. Several years later, Claassen et al. reported that the percentage of patients in the intensive care unit (ICU) undergoing cEEG monitoring who were found to have seizures was 19% [4] with a very high percentage (92%) of these seizures being nonconvulsive. A recent cohort study at 11 North American sites showed that 30% of pediatric ICU patients had seizures and 11% of the patients had NCSE [1]. Continuous EEG remains the gold standard for diagnosis of NCS/NCSE. CEEG is beginning to be used in ICU seizure treatment studies [12] and has been shown to be favorably associated with good outcome [18]. Recently, it has also been reported that not only patients with primary neurological diseases but also medical/surgical ICU patients with secondary neurological complications benefit from cEEG monitoring [3,13].

There is significant cost associated with cEEG monitoring. EEG recording equipment, network connections, MRI and CT compatible electrodes [5], and 24-hour EEG technologist support for connecting and maintaining electrodes are needed. However, another significant source of cost is the physician effort needed to review the cEEG signal, which is recorded by approximately 20 sensors over a time period of hours to days. Optimal diagnosis would involve continuous analysis of this signal to detect seizures, but this is unfeasible for conventional ICU staffing models. In clinical practice, manual analysis of cEEG recordings is done by reviewing pages showing 10 to 20 seconds of EEG. In order to review this much data, the physician reviewer often has to view the cEEG recording very rapidly, which makes it easy to miss brief seizures. An automated detection system could evaluate the cEEG continuously and present results in real-time. Detailed analysis of EEG segments labeled by an automated detection system could replace continuous evaluation of the full EEG and avoid an error-prone accelerated review of long-term EEGs.

Quantitative EEG (QEEG) was a first step towards an automatic and objective interpretation of the EEG signal to assist in evaluation and decision-making. QEEG allows

Table 1 Summary of the critical care EEG terminology (CCET).

Main term 1	Main term 2	Plus (+) modifier
G: generalized	PD: periodic discharges	No +
L: lateralized	RDA: rhythmic delta activity	+F: superimposed fast activity
BI: bilateral independent	SW: rhythmic spike-and-wave OR Rhythmic sharp and slow wave OR Rhythmic polyspike and wave	+R: superimposed rhythmic activity
Mf: multifocal		+S: superimposed sharp waves or spikes, or sharply contoured - applies to RDA only +FR: if both subtypes apply - applies to PD only +FS: if both subtypes apply - applies to RDA only

a time-compressed view at a scale of a few hours using measures like median amplitude or asymmetry that are supposed to capture clinical important information from the EEG. Numerous applications and assessments of QEEG have been reported in the literature [6—8,15,16,19,21]. A common problem is artifacts and physiological EEG patterns that contribute to the QEEG measure in the same way as pathological EEG patterns [20]. For example, the mean amplitude of the EEG cannot differentiate between high amplitude artifact and seizure activity, for which reason the interpretation of this single measure is highly ambiguous.

There is a need to standardized EEG patterns in order to avoid misinterpretation between staff members and different clinical sites. The standardized critical care EEG terminology (CCET) of the ACNS [11] lays a foundation for a common nomenclature for ICU EEG by defining clinically relevant EEG patterns. Table 1 summarizes the available codes to describe rhythmic and periodic EEG patterns. The codes for pattern localization (main term 1) and pattern type (main term 2) are concatenated to give a single pattern code, e.g. ''LRDA'' for lateralized rhythmic delta activity. These CCET codes showed a high interrater agreement [9,17]. In this work, we present a computational method for automatic detection of clinically significant EEG patterns in cEEG recordings from ICU patients based on the CCET criteria.

Materials and methods

We developed a computational method to detect rhythmic and periodic patterns according to ACNS CCET. Pattern locations and pattern types are defined for all detections in a time interval of a few seconds. In addition to EEG patterns defined in CCET, rhythmic patterns with frequencies of more than 4 Hertz are detected.

The automatic detection algorithm was developed by utilizing long-term EEG recordings from a neurological intermediate care unit and a neurological ICU as development dataset. This development dataset was recorded at the 2nd Neurological Department of the Neurological Center Rosenhuegel (NCR) and the Department of Neurosurgery at the Medical University of Vienna (MUV) using the international 10—20 electrode placement system at a sampling rate of 256 Hz.

The goal of this computational approach is to transcribe automatic detections to clinically established wording. This contrasts with conventional QEEG methodology, which evaluates the EEG to report technical measurements. The link between technical measurements and EEG patterns of interest is not always so clear. For example periodic patterns (PD) and spike wave activity (SW) at 1 Hertz have a similar rhythm in a time-frequency plot but they are very different clinical patterns and have different implications for the patient. Another advantage of a tool such a method is a strong data reduction property since it provides high-level data extraction. A periodic pattern (PD) can stretch over minutes or hours without changing frequency or amplitude. A computational method that recognizes PDs can define the start and the end of the pattern instead of reporting the same detection at short fixed intervals. This data compression property of the detection algorithm allows the graphical representation of several days of EEG on a few pages.

Implementation of automatic pattern detections based on CCET

The definitions of rhythmic and periodic EEG patterns in the CCET [11] are given using clinical wording that specifies only limited technical details in terms of signal morphology. Because of this, a detailed description of the way CCET was implemented by our detection algorithm will be given in this section.

Main term 1 of the CCET defines several categories for the localization of EEG patterns. First of all, rhythmic and periodic patterns can be generalized (G) or lateralized (L). In addition predominant areas can be specified. The implementation of recognition of the main term 1 in our detection algorithm supports generalized (G) and lateralized (L) pattern. A pattern qualifies as lateralized (L), if the maximum amplitude of the pattern in one hemisphere is at least 50% higher than in the contralateral hemisphere, based on bipolar transverse and longitudinal montages only. If not qualified as lateralized, the pattern is generalized. If a generalized pattern has 50% higher amplitude in the frontal, midline, or occipital area compared to a contralateral or bilateral reference area, the pattern is predominant. Frontally, midline, and occipital predominant patterns are therefore detected according to the CCET definition except that again only bipolar longitudinal and bipolar transverse montages are used.

The CCET defines various types of rhythmic and periodic patterns in main term 2 (see Table 1). These pattern types

are periodic discharges (PD), rhythmic delta activity (RDA), and repetitive spike-and-wave or sharp-and-wave (SW) patterns. PDs are defined as uniform discharges repeating in regular intervals with a clear inter-discharge interval (IDI). Only discharges with waveforms having less than 4 crossings of the baseline are allowed in order to distinguish periodic from burst suppression patterns. The average frequency allowed for PDs ranges from 0.2 to 3 Hz. The so-called ''relative amplitude'' is defined as the average discharge amplitude divided by the average amplitude between the discharges. The value of the relative amplitude of a periodic pattern has to be above 1.6 to be detected. Rhythmic delta activity (RDA) is defined as repeating discharges with uniform morphology without an inter-discharge interval. The computational detection of RDA conforms to this definition. Several modifiers to the main term 2 pattern types are defined in the standardized terminology that describes variants in the morphology. The modifier ''+S'' is defined as a pattern with frequent intermixed sharp waves/spikes or a sharply contoured pattern and is only applicable to patterns of type RDA. These ''RDA + S'' patterns are detected if at least one unequivocal spike is included in a RDA pattern or the pattern has a sharply contoured morphology. Spike-and-wave or sharp-and-wave (SW) patterns are defined as polyspike, spike or sharp wave consistently followed by a slow wave in a regularly repeating and alternating pattern. The automatic detection of SW is based on detected RDA + S patterns and requires in addition that 20% of the discharges in the RDA + S pattern coincide with unequivocal spike-and-waves or sharp-and-waves. This is a more relaxed condition compared to the CCET and increases the robustness of the automatic detection. Finally, a frequency up to 4 Hertz is allowed to qualify for SW in our method. The detection of rhythmic theta and rhythmic alpha patterns (RTA, RAA) is logically equivalent to the RDA detection but requires more than 6 successive discharges and frequencies between 4—7.5 or 7.5—12 Hertz, respectively. The modifier ''amplitude'' is defined as the average amplitude of all discharges in the pattern. The discharge amplitude is defined as the minimum of the two peak-to-peak voltages measured from the start of the discharge to the maximum and from the maximum to the end of the discharge. The modifier ''frequency'' is determined as the average distance between consecutive discharges in a pattern.

Our automatic detection method does not capture the full depth of the CCET. The main term 1 types bilateral independent (BI) and multifocal (Mf) are not implemented. This means that patterns of this main term 1 type are assigned to generalized or lateralized patterns, depending on the exact amplitude distribution over channels. The modifiers amplitude and frequency are implemented but other modifiers and EEG background are not evaluated in this version.

Calculation procedure

The following is a description of how the detection algorithm processes EEG data. At the beginning, EEG artifacts are removed using the PureEEG algorithm [10]. The PureEEG algorithm is based on a neurophysiological model and utilizes an iterative Bayesian estimation scheme to remove typical scalp EEG artifacts like movement, muscle, line noise, and loose electrode artifacts. The output of the PureEEG algorithm is a clean EEG signal that is solely used for further analysis. This approach assumes that all following pattern detections are of cerebral origin. Frequencies below 0.4 Hertz and above 70 Hertz are then removed by a finite impulse response filter. Bipolar longitudinal and transverse montages are created according to ACNS' proposal for clinical EEG montages [2]. The signal in each bipolar channel is divided into segments that represent spikes, waves, or any other discharge item with durations between 40 milliseconds and 1.5 seconds and amplitudes above 20 microvolts. The wave segmentation procedure scans the EEG signal in the time domain for arbitrary peaks with more than 20 mV. Each peak is then extended on both sides as long as the waveform lies above two projection lines that start at the borders of the starting peak and have 20% reduced slope value. All resulting wave segments below 40 milliseconds and above 1.5 seconds are dropped. These single-channel segments are then combined over several channels to build multi-channel segments. The spatial distribution of potentials within multi-channel segments is checked and segments with non-cerebral origin are discarded. All multi-channel segments are then marked as spike, sharp wave, or non-spike segment by a spike-detection algorithm. The multi-channel segments are also used to build groups of representing RDA, RDA + S, SW, RTA, RAA, and PD segments as described above. Whether a pattern meets minimal requirements for duration and the number of discharges is checked and patterns that do not meet criteria are discarded. The spatial location of all detected patterns is analyzed and a main term 1 definition is assigned. Finally, segments of equal pattern type are concatenated to groups with a maximal duration of 30 seconds. Fig. 1 summarizes the major calculation steps of the algorithm. These groups are displayed in the main term 1 and 2 plot of the graphical user interface.

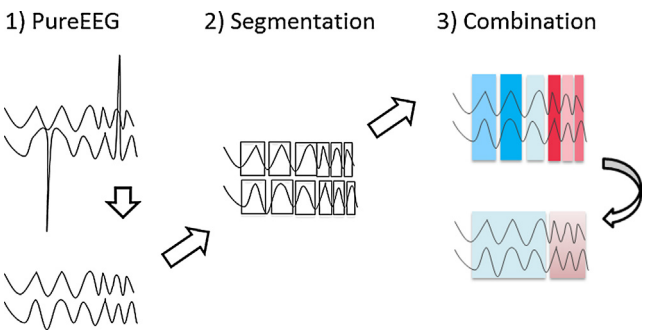

Figure 1 The three major steps of the detection algorithm: 1) Artifact removal using the PureEEG algorithm. 2) Channel wise discharge segmentation. 3) Combination of discharge segments over several channels followed by grouping in time. The final pattern groups represent EEG patterns with multiple discharges in a certain spatial area.

Graphical detection user interface

A graphical user interface (GUI) was created to present information of the detected patterns to a reviewer. This detection user interface simplifies review session by allowing visual recognition of clusters with similar information.

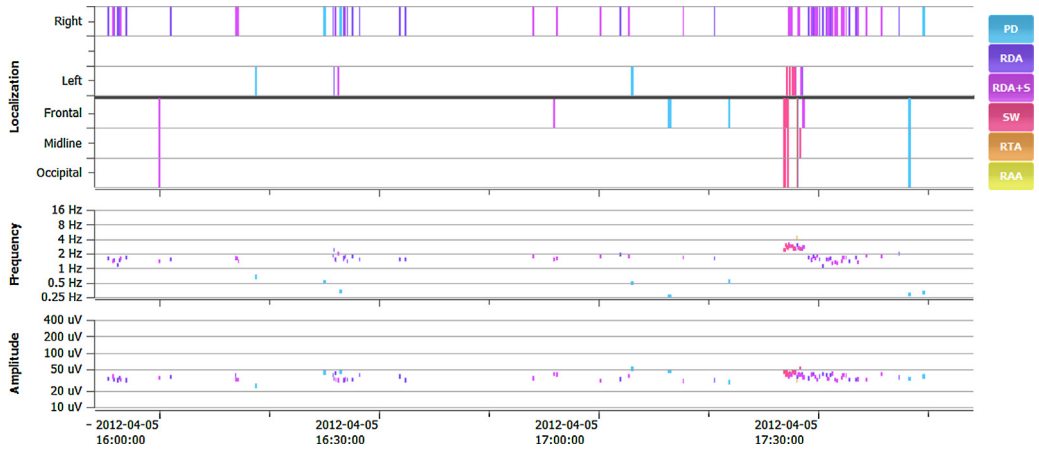

Figure 2 The detection user interface showing of 2 hours of EEG recording. The graphical interface displays rhythmic and periodic pattern detections with colors corresponding to main term 2 (PD, rhythmic delta activity [RDA], RDA + S, SW) or faster rhythmic activity (rhythmic theta activity [RTA], rhythmic alpha activity [RAA]). The ''localisation'' plot shows the predominant areas of the activity. Lateralized patterns are displayed using a box with small height on the correspondig lateralization line (right, left). Generalized patterns may have a predominance in a specific area (frontal, midline, occipital) and will be displayed using a box with small height. A generalized pattern without anterior-posterior predominance is shown as rectangle spanning over the positions of ''frontal'', ''midline'', and ''occipital''. The frequency and amplitude of all patterns are shown on the second and third plots, respectively. This zoomed detection plot of case 1 shows right lateralized rhythmic activity (LRDA, LRDA + S) at the beginning and near the 16:30 timepoint. An electrographic seizure with a spike-and-wave morphology (GSW and LSW, red colored markers) can be observed at 17:30. Post-ictal slowing is marked by LRDA and LRDA + S (violet and magenta markers).

Fig. 2 shows an example of the GUI from a 2-hour EEG recording. It displays periodic discharges (PD), rhythmic delta activity (RDA), and spike-and-wave patterns (SW), rhythmic theta activity (RTA), and rhythmic alpha activity (RAA) as color-coded bars. All detected patterns are presented on separated but time aligned plots showing localization, frequency, and amplitude.

The localization of the pattern (main term 1) can be observed on the upmost plot. Predominant patterns are plotted as rectangle in one of the five possible vertical positions, which are labeled as ''right'', ''left'', ''frontal'', ''midline'', and ''occipital''. A generalized pattern without predominance is shown as rectangle spanning over the positions of ''frontal'', ''midline'', and ''occipital''. This kind of visualization enables the observation of trends in the spatial distribution of pattern potentials.

Frequency and amplitude of all patterns are shown on the second and third plots, respectively. The vertical position indicates the frequency and amplitude of the patterns on logarithmic scales, respectively. Trends in frequency or amplitude might reveal additional information that can be uncovered on these plots. The underlying EEG can be viewed in an EEG viewer by clicking on a time position in the GUI.

Assessment methodology

We assessed the performance of our detection method in two parts. In the first part, three cases from neurological ICUs are presented that were retrospectively analyzed with our computer algorithm. We summarized the original EEG reports and compared them with the detections shown on the detection user interface, to exemplify the evaluation procedure and to show differences and additions between the neurophysiological report and automatic calculated detections.

In the second part, we present the results of a preliminary evaluation of the ability of our detection algorithm to capture ACNS-defined features of the cEEG from ICU patients. For this study, we randomly selected 10 ICU patients including 187 hours of EEG with an associated clinical EEG report from the Comprehensive Epilepsy Center, Medical University of South Carolina (MUSC) recorded between October 2011 and April 2012. All EEGs were recorded using the international 10—20 electrode placement system at a sampling rate of 256 Hz. The average age of the patients was 57 years (min. 25, max. 74) and the average recording duration was 19 hours (min. 11 h, max. 30 h). The clinical indication for long-term EEG recording included altered mental status, intercerebral hemorrhage, history of status epilepticus with new onset seizures, and stroke. The neurophysiology reports from MUSC included general EEG annotations and statements but no ratings according to CCET terminology. We asked a clinical neurophysiologist from the Neurological Center Rosenhuegel (NCR) blinded to the original EEG report and naive to these EEGs to write reports using our detection user interface (and the raw EEG, if needed) using only 10 minutes of time per patient. These detection-guided reports (DGR) were then compared to the original clinical reports (CR) from MUSC. All of the reports from MUSC were generated by academic clinical neurophysiologists board certified by the American Board of Clinical Neurophysiology. A comparison of the reports was done manually by matching keywords like ''slowing'', ''seizure'', and ''periodic discharge''. Interictal patterns with less than 6 cycles [like temporal intermittent rhythmic delta activity (TIRDA)] were excluded because they are not covered by the CCET nomenclature. If older terms were used, which predated the ACNS

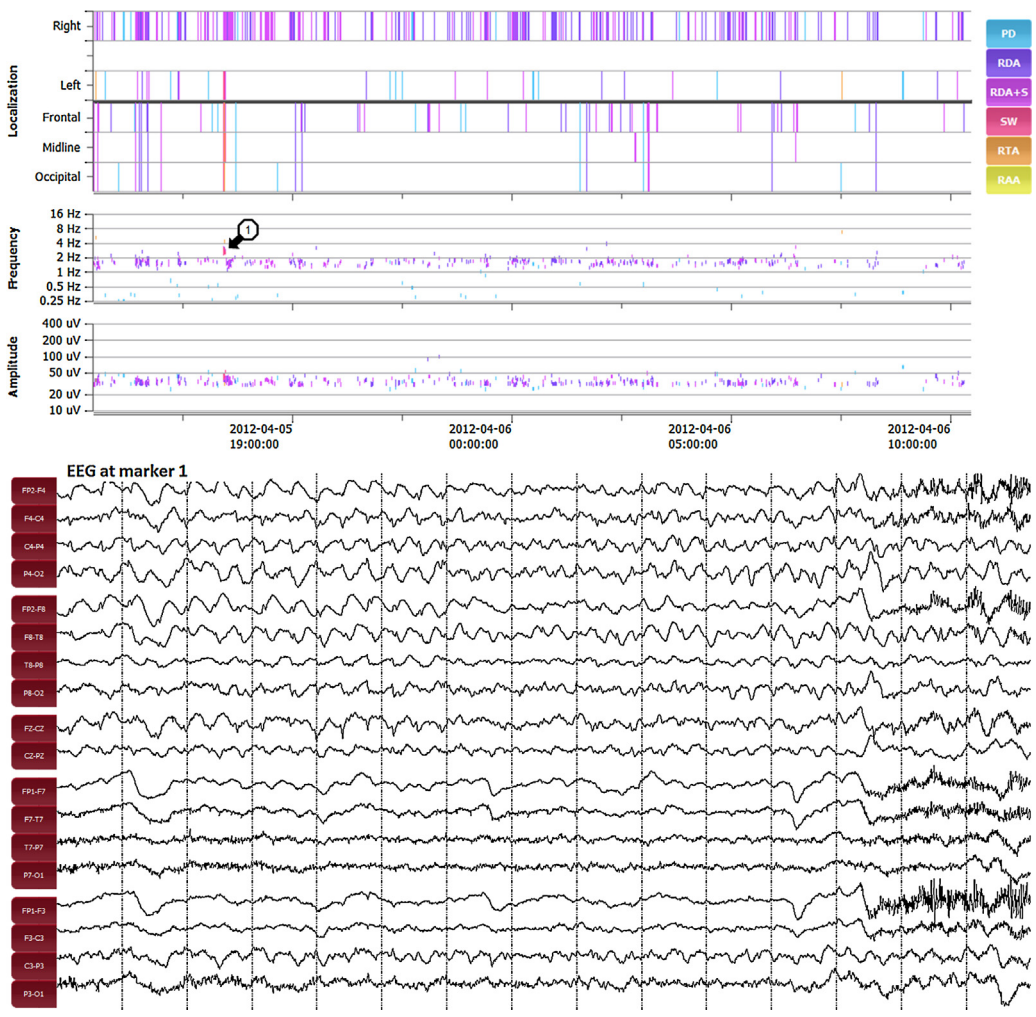

Figure 3 The detection user interface of case 1 shows the result of 20 hours of EEG from a 49-year-old female patient with a brain abscess. Marker 1 shows the time point of an electrographic seizure with repetitive spike-and-wave (SW) activity. The short-term seizure event stays clearly visible on a large time scale because of the color-coded pattern morphology.

terminology [such as ''periodic lateralized epileptiform discharges'' for lateralized periodic discharges (LPDs)], these were considered equivalent. The localization information in both reports was matched based on lateralization or generalization. More precise localization terms referring to a region (i.e. frontal-temporal) were excluded from comparison.

Results

Case reports

Case 1: patient with brain abscess

The patient is a 49-year-old female with history of a brain abscess. The video-EEG monitoring procedure revealed focal, right hemispheric delta slowing and a single electrographic seizure at 17:25:57. The screenshot on Fig. 3 shows detections of RDA and RDA + S from the beginning of the recording by displaying violet and magenta colored bars, respectively. The location of these patterns is mostly on the right hemisphere indicated by most boxes drawn at the label ''right''. The corresponding EEG segments show interictal delta waves at approximately 2c/s with amplitudes of 50 uV. The most prominent group of spike-and-wave detections (in red at marker 1) corresponds to the electrographic seizure mentioned in the original EEG report. The segments without detected patterns did not show any clinically relevant abnormal EEG activity. This case shows that the detection algorithm is able to pinpointing to clinically interesting EEG segments.

Case 2: patient with stroke and abnormal movements

The patient is a 59-year-old female with a history of stroke and repetitive movements of the right forearm. During video-EEG monitoring, left hemispheric slowing together with LPDs on the left parasagittal region were reported. No seizures occurred. The detection result on Fig. 4 confirms a continuous LPD in the left hemisphere (PD colored in light blue) for the first 4 hours of recording with a sudden decay in frequency in the evening at 20:10 (marker 1). In addition, occasional rhythmic delta activity with sharp morphology (RDA + S, colored in magenta) was found in the same time span on the same hemisphere. After a pause of two hours,

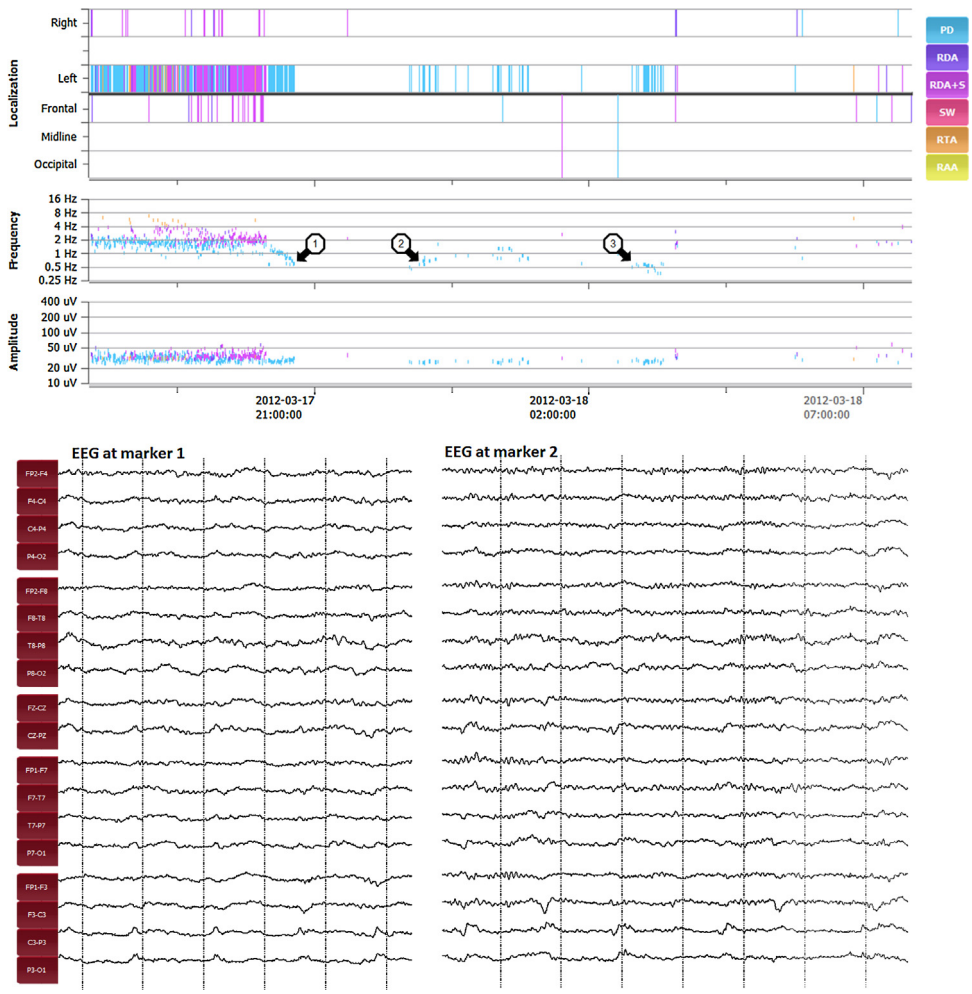

Figure 4 Case 2 is a 59-year-old female with a history of stroke and repetitive movements of right forearm. Left hemispheric lateralized periodic patterns with decaying frequency can be observed (marker 1). During the same time, rhythmic delta activity with sharp morphology (rhythmic delta activity + S, magenta color) can be seen. Periodic activity reappears after pauses of several hours (marker 2 and marker 3).

occasional LPD can be seen at 23:00 (marker 2) and at 03:00 in the morning of the next day (marker 3). It is interestingly to note the trend of the LPD frequency over time and the clusters of PD patterns shown in the GUI.

Case 3: patient with NCSE after ICH

The patient is a 74-year-old female with history of intracranial hemorrhage (ICH) and nonconvulsive status epilepticus (NCSE). The report from video-EEG monitoring at MUSC described diffuse arrhythmic delta slowing and frequent spikes from the left temporal region. At least nine seizures were identified from 21:22 to 22:23. The screenshot of the detection GUI on Fig. 5 shows an initial period with abundant spike-and-wave patterns in the left temporal region (LSW, shown in red) together with rhythmic theta activity (RTA, in orange). The time position of these very brief subclinical seizures can be observed by the SW markings (example marker 1). All seizure events are lateralized to the left side. In addition to the findings reported at MUSC a segment of LPD from the left hemisphere can be seen in the detection GUI analysis (marker 2). An interesting point is that a basic change in the EEG patterns can be observed on a large time scale.

Comparison between manual and detection-guided EEG review

The evaluation of the trial study comparing the detection-guided report (DGR) using 10-minute review time and the clinical report (CR) is summarized in Table 2. The CR was written using the evaluation of video, which allows a more reliable decision between seizure and artifact given an ambiguous EEG segment. Some patterns in the DGR were therefore described as ''episode of rhythmic theta activity at 6c/s'' instead of ''seizure with 6c/s'' but were considered to be equivalent.

The major items in both reports are described and additional findings in the DGR and detections missed in the DGR are outlined for each patient. A comparison between the DGR and the CR shows that in three out of five patients with reported seizures (patients P1, P3, and P6) all seizures from the original report were found. In patients P5 and P6,

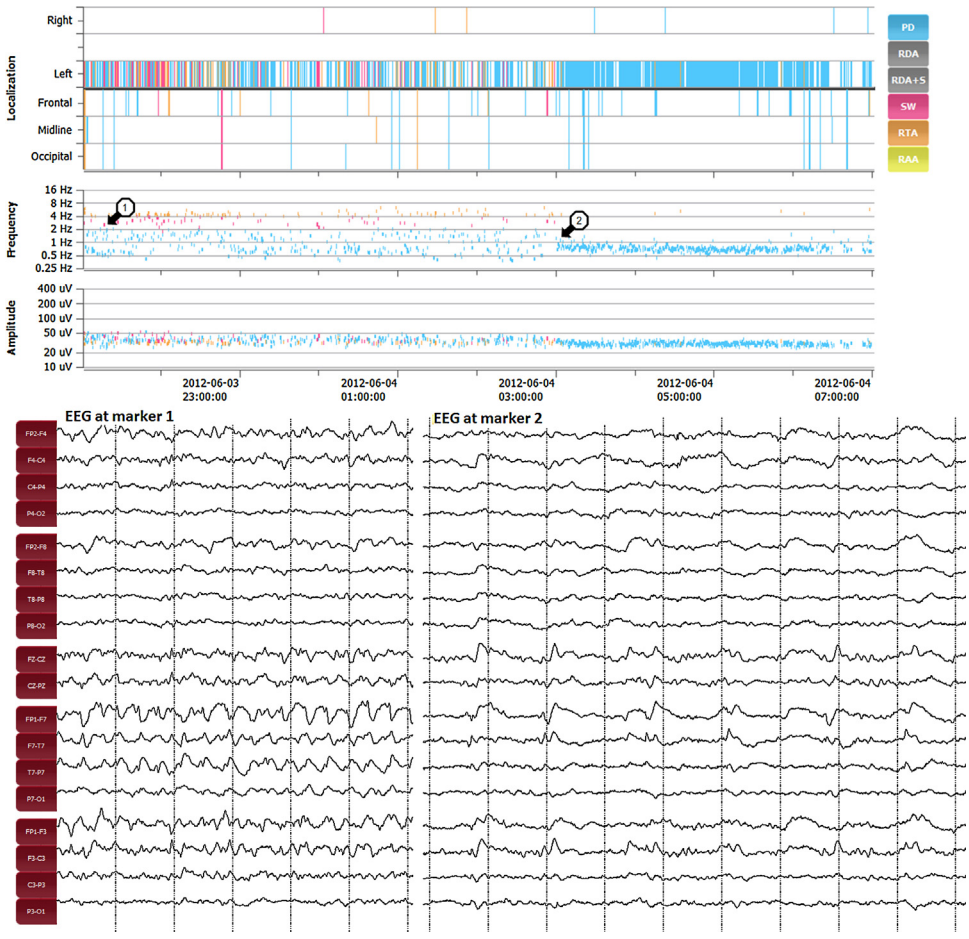

Figure 5 Case 3 is a 74-year-old female patient with history of intracranial hemorrhage and now with nonconvulsive status epilepticus. The detection user interface shows an initial period with many repetitive spike-and-wave patterns on the left temporal region (LSW shown in red, i.e. at marker 1) together with sharply contoured rhythmic delta activity (rhythmic delta activity + S, magenta). A period of lateralized periodic patterns (marker 2, light blue) can be observed over the same hemisphere, which continues for more than 3 hours.

additional seizures were found during the guided review. In one patient (P4), all 14 clinical seizures were missed because 12 seizures showed no clear EEG correlation and two showed only minor electrographic correlations. The seizures could only be picked up with extensive video monitoring showing a slight head movement and a movement of the right arm. The events were then annotated as complex partial seizures. Another patient (P7) had one seizure with poor electrographic correlation, which was missing in the DGR. LPDs were coincided in the reports of 2/2 patients (P1, P2). In three patients (P4, P5, and P10), LPDs were additionally found in the DGR. In patient P7, LPD was reported by the MUSC reviewer but GPD was reported in the DGR. The reports of EEG slowing by MUSC interpreters were in agreement in 7/9 patients (missing in DGR of P2 and P10). In two patients (P6 and P8), episodes of generalized rhythmic theta activity were reported in the DGR, which had not been mentioned in the MUSC clinical report. In one patient (P2), a focal slowing was not detected and reported in the DGR review because of the missing background evaluation capability in the detection method.

Discussion

In this article, we presented an automated computer algorithm that detects rhythmic and periodic patterns in ICU EEG recordings based on ACNS standardized critical care EEG terminology (CCET) [11]. This terminology was defined by a group of experienced neurologists using standardized clinical wording. Detection results are displayed on a graphical detection user interface to simplify review sessions. The interpretation of the detections on the user interface is then based on these clinically defined terms, therefore avoiding mathematical or technical nomenclature. A clinical application of such a detection user interface will allow a quick overview of several hours of EEG without overwhelming medical staff with technical information.

Due to the complexity of the EEG signal, in current clinical practice EEG can only by analyzed by highly trained EEG experts. These experts have limited time for EEG review and, due to the recent expansion in ICU EEG monitoring, they are currently be overwhelmed by the amount of EEG that needs to be reviewed at medical centers which are able to implement ICU EEG monitoring. Pre-analysis of EEG

Table 2 Comparison between 10 EEG reports written after time unlimited clinical video-EEG review (CR) and after a 10-minute EEG review using automatic detections of rhythmic and periodic patterns (DGR).

ID	Clinical video-EEG report (CR)	Detection-guided report (DGR)	Additional in DGR	Missed in DGR	Rating
P1	Right side PLEDS which improve over time; brief electrographic seizures. Slowing	Frequent 0.5—1/s right LPD, slowing, and many brief electrographic seizures	—	—	=
P2	Left hemisphere slowing, PLEDS in the left parasagittal region, epileptiform discharges occurring at 2 Hz, twitching of the right hand	Left fronto-temporal LPDs until 9 pm disappearing for 4 h, and reappearing at 1 am with increasing frequency. No seizures	Time course of LPD, frequency changes	Slowing	±
P3	Focal right frontal-temporal hemispheric delta slowing. One electrographic seizure at 5:25 pm on 4/5/12	Slowing on the right, seizure with 3/s GRDA/GSW at 05:25 pm on 4/5/12	—	—	=
P4	Right parasagittal delta. Multiple subtle seizures	Occasional right LPD starting at 00:00 h, intermittent slowing on the right	LPD right (matching right parasagittal delta)	14 clinical seizures, only 2 with EEG correlate	±
P5	Diffuse arrhythmic delta slowing, TIRDA, no seizures	Intermittent slowing, abundant 1/s left LPD until approx. 11:00 pm, one brief electrographic seizure	One brief electrographic seizure, LPD	TIRDA (too short)	+
P6	Independent left and right temporal spike waves, one electrographic seizures	Intermittent slowing, 3 electrographic seizures	1 additional electrographic seizure, slowing, increasing theta activity at the end	Spikes left temporal	±
P7	Intermittent right frontal-central bursts of 2—3 Hz delta slowing, PLEDS, intermittent right frontal sharp wave discharges, one electrographic seizure	Occasional GPD, rhythmic theta at 7:00 pm, generalized slowing, no seizures	—	1 focal right seizure	—
P8	Continuous bilateral, frontally predominant generalized sharp waves, 5—6 Hz theta slowing of the background bilaterally	Slowing, GPD 11:30 pm to 4:15 am, 3—4/s GSW from 09:20 pm to 10:30 pm	Theta pattern from beginning to 1:00 am. GPD, GSW	—	++
P9	Diffuse arrhythmic delta slowing, triphasic waves	Slowing, continuous 1—2/s GRDA/GRDA + S/GSW of 1—10 s duration with fluctuating morphology starting from 11:00 pm, which matches NCSE criteria	Time course, more detailed pattern description	—	++
P10	Periods of intermittent generalized delta slowing	Very few episodes of low amplitude left LPD at 0.5/s or less, no background activity	LPD	Intermittent slowing	±

LPD: lateralized periodic patterns; PLEDS: periodic lateralized epileptiform discharges; TIRDA: temporal intermittent rhythmic delta activity.

performed by computer algorithms may speed the process of this EEG review and perhaps allow personnel with less EEG training to accomplish thorough EEG review.

The results of the trial study provide some evidence that review of long EEG recordings using our computer algorithm is possible in a short time period. In the limited series of cases presented, most of the important elements in the EEG were automatically detected in a short review time. The detection-guided reviewer had only 10 minutes time to evaluate 11 to 28 hours of EEG. Given this substantial time restriction, we think that the high detection accuracy we found is promising. The detection accuracy for PD and RDA was almost 100% compared to the extensive video analysis of the EEG, in fact, additional PD patterns were detected in many patients. In addition, generalized rhythmic theta was detected in two patients that were missed by the MUSC reviewer. The detection-guided reviewer missed several subtle seizures in one patient that did not show clear EEG correlation. These seizures could only be picked up using video-EEG and could not be detected by an automated system that only examined EEG. On the other hand, the detection-guided reviewer found additional information about pattern trends over a long time period. Manual analysis of highly complex EEG tends to pick out sporadic events to describe the patient status. The results show that systematic digital analysis of patterns additionally enables to capture trends hidden in complex EEG data. This information is emphasized in the reports of the detection-guided review of patients P2, P6, P8, and P9.

We envision the use of our automatic detection system for monitoring of intensive care patients, where EEG recording equipment would send the digitized EEG data in real-time during the recording. Sudden changes in the EEG could then be observed through periodic checks of the detection user interface by the ICU staff. The visual combination of detection representing rhythmic and periodic patterns in combination with other neuromonitoring parameters such as intracerebral pressure (ICP) on the same time scale could potentially reveal additional information.

Our computer algorithm detects patterns according to CCET nomenclature only. Information about normal patterns present in the EEG is not reported. In the future, we would like to add additional features that would evaluate and report features of the normal background EEG such as continuity, dominant frequency, and background amplitude. In such an approach, every EEG segment would be evaluated without exception.

We are aware that the study design has several significant weaknesses. First, there were only a limited number of cEEG recordings studied and these recordings were not collected prospectively. Second, the standard EEG reports to which the detection-guided report was compared, were based on only one MUSC reviewer and this MUSC reviewer was different for many of the EEG recordings. Third, only a single expert reviewer performed the detection-guided review. However, the high agreement between MUSC and the detection-guided reports and the high percentage of detection-guided reports with additional information encourage the initiation of a comprehensive prospective study.

Conclusion

A computational method to detect rhythmic and periodic patterns based on the ACNS' standardized critical care EEG terminology was presented. The results of three patient cases showed the potential of this system to review EEGs of critical care patients. An objective comparison in a preliminary trial study of 10 long-term EEG recordings showed that the utilization of a detection-guided review system could possibly assist with clinical EEG analysis. Further multicenter studies including larger prospectively acquired EEG recordings and validation using multiple expert EEG reviewers are needed to validate our computational method.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgements

This work was partly done within the DeNeCoR project which received funding from the ECSEL Joint Undertaking under grant agreement No. 324257 and from the national funding authorities of Austria, Czech Republic, Germany, Italy, the Netherlands, Spain and the United Kingdom.

References

[1] Abend NS, Arndt DH, Carpenter JL, Chapman KE, Cornett KM, Gallentine WB, et al. Electrographic seizures in pediatric ICU patients: cohort study of risk factors and mortality. Neurology 2013;81(4):383—91.
[2] American Clinical Neurophysiology Society. Guideline 6: a proposal for standard montages to be used in clinical EEG. Am J Electroneurodiagnostic Technol 2006;46(3):226—30.
[3] Claassen J, Mayer SA, Hirsch LJ. Continuous EEG monitoring in patients with subarachnoid hemorrhage. J Clin Neurophysiol 2005;22(2):92—8.
[4] 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(10):1743—8.
[5] Das RR, Lucey BP, Chou SH-Y, Espinosa PS, Zamani AA, Dworetzky BA, et al. The utility of conductive plastic electrodes in prolonged ICU EEG monitoring. Neurocrit Care 2009;10(3):368—72.
[6] Finnigan S, van Putten MJAM. EEG in ischaemic stroke: quantitative EEG can uniquely inform (sub-)acute prognoses and clinical management. Clin Neurophysiol 2013;124(1):10—9.
[7] Foreman B, Claassen J. Quantitative EEG for the detection of brain ischemia. Crit Care 2012;16(2):216.
[8] Friedman D, Claassen J, Hirsch LJ. Continuous electroencephalogram monitoring in the intensive care unit. Anesth Analg 2009;109(2):506—23.
[9] 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(9):1366—73.
[10] Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: automatic EEG artifact removal for epilepsy monitoring. Neurophysiol Clin Neurophysiol 2014;44(5):479—90.
[11] 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):1—27.
[12] Husain AM. Treatment of recurrent electrographic nonconvulsive seizures (TRENdS) study. Epilepsia 2013;54(Suppl. 6):84—8.
[13] Kamel H, Betjemann JP, Navi BB, Hegde M, Meisel K, Douglas VC, et al. Diagnostic yield of electroencephalography in the medical and surgical intensive care unit. Neurocrit Care 2013;19(3):336—41.
[14] Kaplan PW. Assessing the outcomes in patients with nonconvulsive status epilepticus: nonconvulsive status epilepticus is underdiagnosed, potentially overtreated, and confounded by comorbidity. J Clin Neurophysiol 1999;16(4):341—52 [discussion 353].
[15] LaRoche SM, editor. Handbook of ICU EEG monitoring. 1st ed. New York (NY): Demos Medical; 2012.
[16] Machado C, Estévez M, Carrick F, Mellilo R, Leisman G. qEEG may increase the reliability of diagnostic and prognostic procedures in cerebral arterial gas embolism. Clin Neurophysiol 2012;123(2):225—6.
[17] Mani R, Arif H, Hirsch LJ, Gerard EE, LaRoche SM. Interrater reliability of ICU EEG research terminology. J Clin Neurophysiol 2012;29(3):203—12.
[18] Ney JP, van der Goes DN, Nuwer MR, Nelson L, Eccher MA. Continuous and routine EEG in intensive care: utilization and outcomes, United States 2005—2009. Neurology 2013;81(23):2002—8.
[19] Nuwer M. Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society. Neurology 1997;49(1):277—92.
[20] Nuwer MR. Assessing digital and quantitative EEG in clinical settings. J Clin Neurophysiol 1998;15(6):458—63.
[21] Salinsky MC, Binder LM, Oken BS, Storzbach D, Aron CR, Dodrill CB. Drug exposure and EEG/qEEG findings. Epilepsia 2002;43(5):482—90.

6.4 Paper A4: Burst suppression detection

Title: Monitoring burst suppression in critically ill patients: Multi-centric evaluation of a novel method

Authors: Franz Fürbass, Johannes Herta, Johannes Koren, M. Brandon Westover, Manfred M. Hartmann, Andreas Gruber, Christoph Baumgartner, Tilmann Kluge

Published in: Clinical neurophysiology

Year: 2016

Authors' contribution: Fürbass Franz developed the algorithm for detection of burst suppression patterns. In addition data evaluation and statistical analysis was done by Franz Fürbass. Initial writing of the manuscript, implementation of corrections from coauthors and reviewers was done by Franz Fürbass. Data annotations, EEG recording, parts of the corrections, and the clinical work were done by coauthors.

Clinical Neurophysiology 127 (2016) 2038–2046

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Monitoring burst suppression in critically ill patients: Multi-centric evaluation of a novel method

Franz Fürbass [a,*], Johannes Herta [b], Johannes Koren [c], M. Brandon Westover [d,e], Manfred M. Hartmann [a], Andreas Gruber [b], Christoph Baumgartner [c], Tilmann Kluge [a]

[a] AIT Austrian Institute of Technology, Safety & Security Department, Vienna, Austria
[b] Medical University of Vienna, Department of Neurosurgery, Vienna, Austria
[c] General Hospital Hietzing with Neurological Center Rosenhuegel, 2nd Neurological Department, Vienna, Austria
[d] Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
[e] Department of Neurology, Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Article history:
Accepted 3 February 2016
Available online 9 February 2016

Keywords:
Automatic detection
Burst suppression pattern
EEG
Real-time monitoring
Periodic pattern

HIGHLIGHTS

- Fully automatic computational method to detect burst suppression patterns in critical care EEG.
- Insensitivity to EEG artifacts and periodic patterns makes the system suitable for clinical use in real-time patient monitoring.
- Multi-centric evaluation including the EEG of 88 patients showed high sensitivity and specificity.

ABSTRACT

Objective: To develop a computational method to detect and quantify burst suppression patterns (BSP) in the EEGs of critical care patients. A multi-center validation study was performed to assess the detection performance of the method.
Methods: The fully automatic method scans the EEG for discontinuous patterns and shows detected BSP and quantitative information on a trending display in real-time. The method is designed to work without setting any patient specific parameters and to be insensitive to EEG artifacts and periodic patterns. For validation a total of 3982 h of EEG from 88 patients were analyzed from three centers. Each EEG was annotated by two reviewers to assess the detection performance and the inter-rater agreement.
Results: Average inter-rater agreement between pairs of reviewers was $\kappa = 0.69$. On average 22% of the review segments included BSP. An average sensitivity of 90% and a specificity of 84% were measured on the consensus annotations of two reviewers. More than 95% of the periodic patterns in the EEGs were correctly suppressed.
Conclusion: A fully automatic method to detect burst suppression patterns was assessed in a multi-center study. The method showed high sensitivity and specificity.
Significance: Clinically applicable burst suppression detection method validated in a large multi-center study.

1. Introduction

Burst suppression is an electroencephalogram (EEG) pattern consisting of intermittent periods of very low voltage brain electrical activity (“suppression”), alternating in a quasi-periodic fashion with periods of higher amplitude activity (“bursts”). Burst suppression patterns (BSP) are found in a wide range of pathological and clinically-induced conditions, including anesthetic-induced coma, hypothermia (Pagni and Courjon, 1964; Nakashima et al., 1995) deep (Ching et al., 2012; Westover et al., 2015), or arising spontaneously as a result of anoxic brain injury (Niedermeyer et al., 1999; Rossetti et al., 2012). The definition for burst durations and for suppression amplitudes varies depending on patient age and clinical context, ranging from 0.5 to 30 s for the duration of a burst and

* Corresponding author at: AIT Austrian Institute of Technology, Donau-City-Straße 1, 1220 Vienna, Austria. Tel.: +43 50550 4230; fax: +43 50550 4125.
E-mail address: franz.fuerbass@ait.ac.at (F. Fürbass).

http://dx.doi.org/10.1016/j.clinph.2016.02.001
1388-2457/

from 5 to 20 μV for suppression amplitudes (Shellhaas et al., 2011; Zschocke and Hansen, 2011; Hirsch et al., 2013). Although commonly described as a generalized phenomenon, BSP can be asynchronous across the cortex and can occur in limited cortical regions. Local cortical dynamics of BSP were analyzed in Lewis et al. (2013) and are reported in Sperling et al. (1986), Lazar et al. (1999) and Mader et al. (2014).

Manual evaluation of BSP in the EEG is a widely used but impractical approach. Manual evaluation lacks objectivity, and is not feasible for continuous monitoring over multiple hours. Several automatic or semi-automatic detection methods exist in the literature. The recent work of Murphy analyzed burst and suppression segments of pre-term infants using various mathematical features (Murphy et al., 2015). The method was validated using preselected EEG segments and resulted in high agreement compared to three reviewers. A detection method based on the line length feature using the EEG of 10 pre-term infants was presented in Koolen et al. (2014). An automatic classification method for burst and suppression events was validated in (Westover et al., 2013) on 20 critical care EEG recordings that were selected based on clinical EEG reports. The detection algorithm was trained on these 20 EEGs and showed high agreement compared to human annotations. Numerous other methods exist in literature that use various mathematical features to detect BSP (Thomsen et al., 1991; Lipping et al., 1995; Bruhn et al., 2000, 2006; Jaggi et al., 2003; Liang et al., 2014) but include a limited number of patients.

This work will present a fully automated detection method to find burst suppression patterns in multi-channel EEG. The method is insensitive to EEG artifacts and periodic patterns and can be calculated in real-time. We present detection performance results from an evaluation of continuous EEG recordings from 88 adult patients from three intensive care units.

2. Methods

2.1. Automatic detection method

A computational method is presented that automatically detects burst suppression patterns (BSP) in digital multi-channel electroencephalograms (EEGs). The method works fully automatically without the use of training data and without estimation of patient-specific parameters. Data is analyzed in real-time to allow continuous patient monitoring. The goal is to graphically visualize the detection results over large time scales of up to several days in a quantitative EEG interface similar to the approach shown in (Fürbass et al., 2015a). Fig. 1 shows examples of burst suppression and periodic pattern detections of a 20 h EEG recording.

The major steps in the whole detection procedure are outlined in Fig. 2. First, the EEG is segmented into consecutive and non-overlapping detection segments of 15 s. All further processing is based on these detection segments. Scalp EEG artifacts are removed using the PureEEG method (Hartmann et al., 2014). The PureEEG method is based on a neurophysiological model and utilizes an iterative Bayesian estimation scheme to remove artifacts like movement, muscle, line noise, and loose electrode artifacts. Further analysis is based solely on the output of the PureEEG module. All subsequent detection and classification steps therefore assume that the activity measured in the EEG channels are of cerebral origin. The EEG channels are converted to bipolar longitudinal and transversal montages following ACNS recommendations (American Clinical Neurophysiology Society, 2006).

Next, a channel-wise detection of burst suppression events is performed. In each EEG channel x_t the peak-to-peak amplitude is measured by subtracting the minimum from the maximum digital value in non-overlapping chunks of 0.4 s. Only EEG samples of the current detection segment are used. The peak-to-peak time series of channel x_t is smoothed by a moving average filter resulting in $y_t^S = \frac{1}{n}\sum_{i=1}^{n}|x_{t+i}|$. The length of the averaging window n is chosen so that the minimum time for a suppression event is covered. Here, a minimum duration of 1.5 s for suppression events is assumed. The same procedure but with a window length of 0.5 s is repeated resulting in the time series y_t^B. The samples of the time series y_t^S and y_t^B are then used to detect suppression events in the channel. An event may include several chunks of 0.4 s. A chunk is defined as part of a suppression event if either a chunk with double amplitude follows in 1.5 s ($y_{t+1.5}^B / y_t^S > 2$) or if a chunk with double amplitude precedes with 1.5 s distance ($y_{t-1.5}^B / y_t^S > 2$). All remaining chunks in the detection segment are part of a suppression event if their amplitude is below the amplitude of the initially detected suppression chunk. All chunks that are not marked as part of a suppression event at this processing step are part of a burst event if the peak-to-peak amplitude is higher than double amplitude of the surrounding suppression chunks. Fig. 3 shows the processing steps of the channel-wise detection procedure.

The channel-wise detection information is then used as input to a hierarchical cluster algorithm to find spatial groups of the same activity type. The $k \times k$ distance matrix M_S includes the time distance between the middle points of k suppression chunks. The variable k is the total number of suppression chunks in the detection segment. Chunks that were neither marked as suppression nor burst do not contribute to the distance matrix and are also not considered further. The distance matrix is then used to create a hierarchical cluster tree. The Euclidean distance between two chunk positions $a = M_S^{i,j}$ and $b = M_S^{i,j}$ defined as $d(a,b) = \sqrt{\sum_i (a_i - b_j)^2}$ is used to measure the distance between two chunks. The unweighted average distance algorithm using the cluster linkage criteria $\frac{1}{|A||B|}\sum_{a\in A}\sum_{b\in B} d(a,b)$ defines the dissimilarity between two groups of suppression chunks A and B. The same procedure is repeated for chunks of burst activity. The normalized cluster tree is cut with a constant cutoff factor to create burst and suppression clusters. By solely utilizing the middle point as distance metric an influence of the spatial location of the suppression or burst activity is avoided. This also means that channels used to build up a cluster do not have to be spatially adjacent (e.g. cluster C_{SUPP}^4 in Fig. 2). In a next step the best fitting cluster for each time point is determined. Clusters are sorted descending according to their duration. Starting with the longest cluster and by elaborating each cluster in the sorted list, the first cluster that covers a time point is accepted. Subsequent overlapping clusters are reduced in time to be non-overlapping with accepted clusters. Clusters with durations less than the minimum requirement for burst or suppression will be discarded. This approach will discharge parts of the suppression or burst chunks that are not time aligned with the majority of the other chunks in the cluster. This also means that there is no need for a single channel to fully cover the time span of the cluster. All channels are treated equally, the method do not exploit the spatial location of the involved channels. The resulting clusters represent burst or suppression detections that span several EEG channels and extend over a certain time period. In this method clusters need to span at least 40% of the cortical area covered by electrodes to be further used in the detection procedure. The minimum coverage value of 40% was determined empirically and serves as a sensitivity parameter of the method (see Section 4).

An important task in automatic detection of BSP is to avoid false detections of other EEG patterns that consist of discontinuous waveforms. A defining feature of periodic patterns is that they contain regularly repeating waveforms of duration less than 0.5 s. The inter discharge interval of PDs range from a fraction of a second to several seconds and can therefore share some features of burst

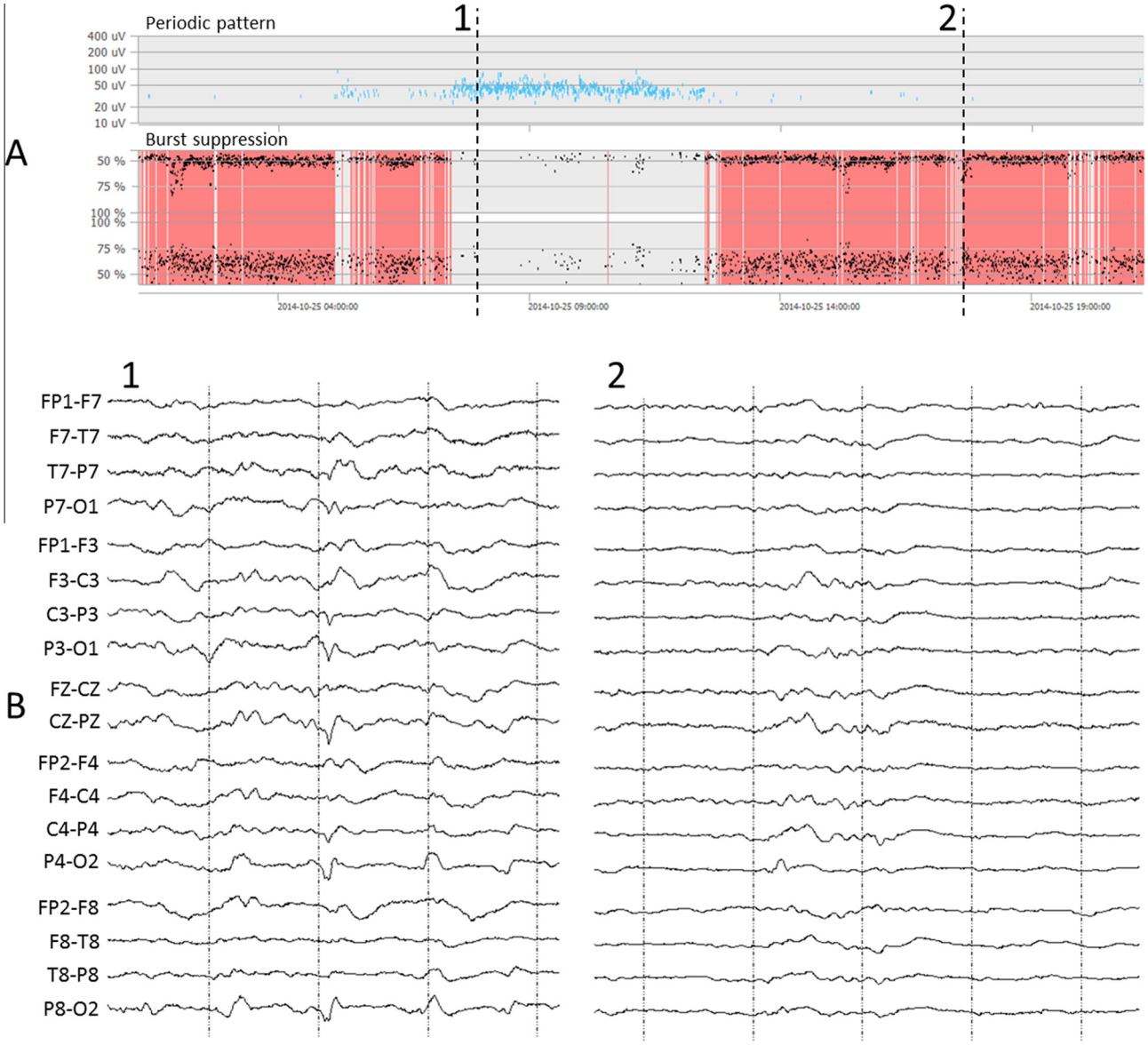

Fig. 1. Quantitative EEG interface (NeuroTrend, www.encevis.com) showing the detection results of a 20 h EEG recording registered in the intensive care unit. (A) The upper plot show periodic pattern detections (Fürbass et al., 2015a) which are continuously present for approx. 5 h in this patient. The lower plot labeled "burst suppression" represents the automatically detected burst suppression patterns. (B) The EEG at time point 1 shows an example of a periodic pattern; the EEG at time point 2 gives an example of a burst suppression pattern.

suppression patterns. When repeating EEG waveforms occur on a low voltage (<10 μV) background that last no longer than 0.5 s or exhibit no more than 3 baseline crossings, the critical care EEG terminology of the ACNS (Hirsch et al., 2013) defines the pattern as one of periodic discharges rather than burst suppression. By contrast, bursts in burst suppression patterns need to have durations of at least 0.5 s and at least 4 baseline crossings. Fig. 4 outlines the differences between burst suppression and periodic patterns and also shows some borderline examples. In this work we apply the definitions of the ACNS critical care EEG terminology by counting the number of waveform crossings of the baseline in each EEG channel of the burst cluster and by measuring the length of the bursts. All burst clusters with a length of less or equal 0.5 s and less than 4 phases are considered as periodic patterns and are dropped. This behavior is also in concordance with the method for automatic detection of periodic patterns presented in (Fürbass et al., 2015a; Herta et al., 2015).

The average length of the bursts τ_{BURST} and the length of the inter burst intervals τ_{IBI} as well as the average suppression and burst amplitude A_{IBI} and A_{BURST} are calculated. These values are stored in the detection result and can be used to characterize the BS patterns. For example the ratio of burst to suppression length is commonly used to measure the depth of pharmacologically induced coma sedation.

Detection segments are marked as a burst suppression pattern if two conditions apply: first, at least one suppression cluster was detected with $A_{IBI} \leqslant 10\ \mu V$ and $\tau_{IBI} \geqslant 1.5$ s. Second, one burst cluster with A_{BURST} of more than two times the lowest suppression cluster amplitude and $\tau_{BURST} \geqslant 0.5$ s was found in the detection segment. The quantity of these parameters follow the definitions in the ACNS' critical care EEG terminology (Hirsch et al., 2013). An exception is the value of 1.5 s for τ_{IBI} which was found empirically through extensive manual evaluation of BSP during algorithm development.

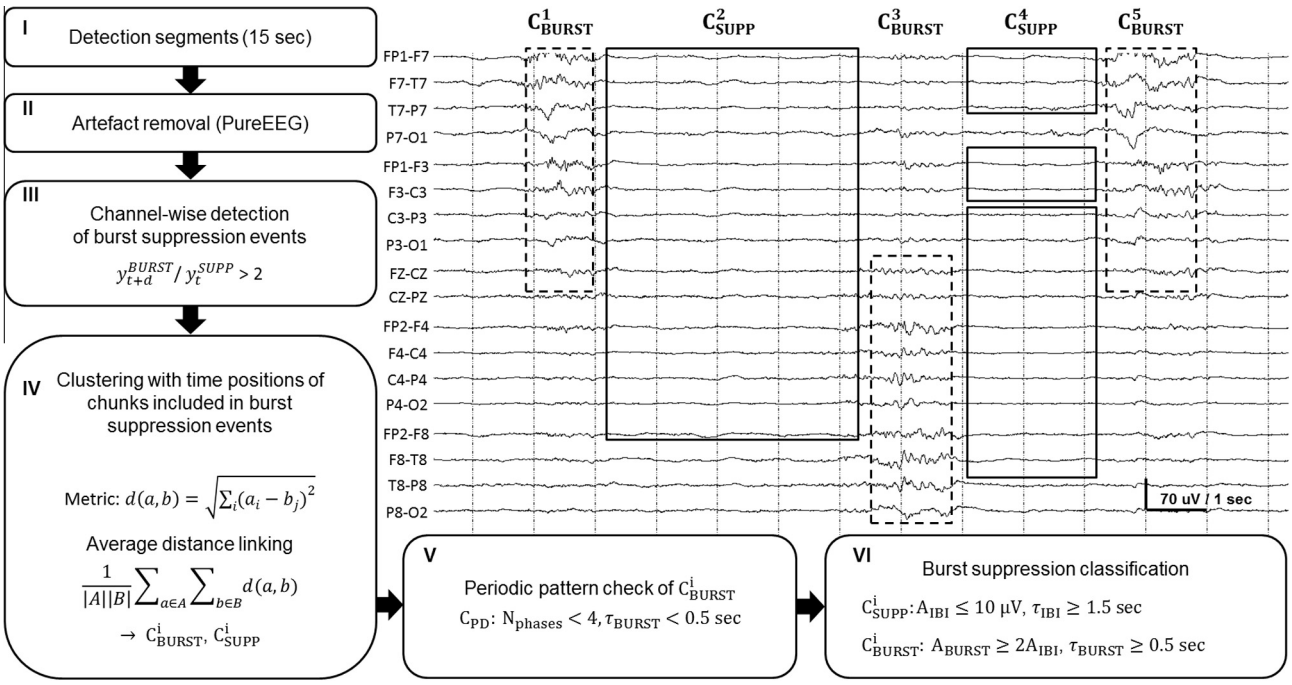

Fig. 2. Block diagram of the automatic burst suppression detection method with the resulting clusters visualized in an EEG segment. The burst and suppression clusters C_{BURST} (dashed line) and C_{SUPP} (solid line) do not overlap in time and may include several channels. The EEG sample shows a bihemispheric and asynchronous burst suppression pattern where each burst covers approximately 50% of the EEG channels. The detection segment was correctly classified as a burst suppression pattern. The calculation procedure involves: (I) segmentation of the EEG, (II) artifact removal using the PureEEG module, (III) channel-wise detection of burst suppression chunks, (IV) building of spatial clusters using the time position of the detected burst and suppression chunks, (V) detection and removal of periodic patterns, (VI) burst suppression classification.

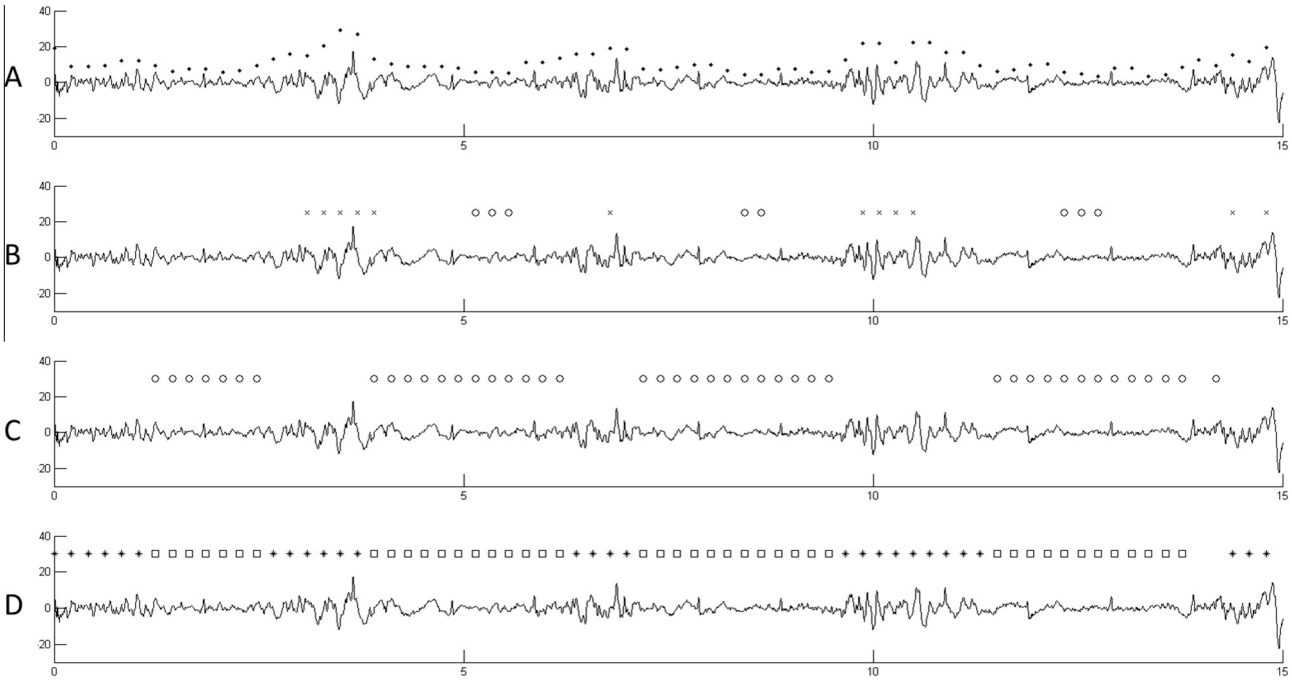

Fig. 3. Processing steps of the channel-wise burst suppression event detection. (A) The peak-to-peak amplitudes measured for each chunk of 0.4 s length are shown as dots. (B) Based on a smoothed time series of these chunk amplitudes each chunk with a preceding double amplitude chunk is marked as part of a suppression event (circle markers on the suppression chunks). The same is done for chunks followed by a double amplitude chunk (cross markers showing the double amplitude chunks). (C) All suppression events are expanded to include all chunks with amplitudes below the initially detected suppression chunks. (D) The final channel wise detection of bursts chunks (star markers) and suppression chunks (square markers) after expansion of the burst.

The method was implemented in the programming language C++ to allow fast calculation and integration in the detection user interface shown in Fig. 1. The software module is able to analyze 24 h of EEG in 20 min on a standard PC hardware and is therefore 72 times faster as the recording speed. The method uses consecutive detection segments of 15 s length, each detection segment can

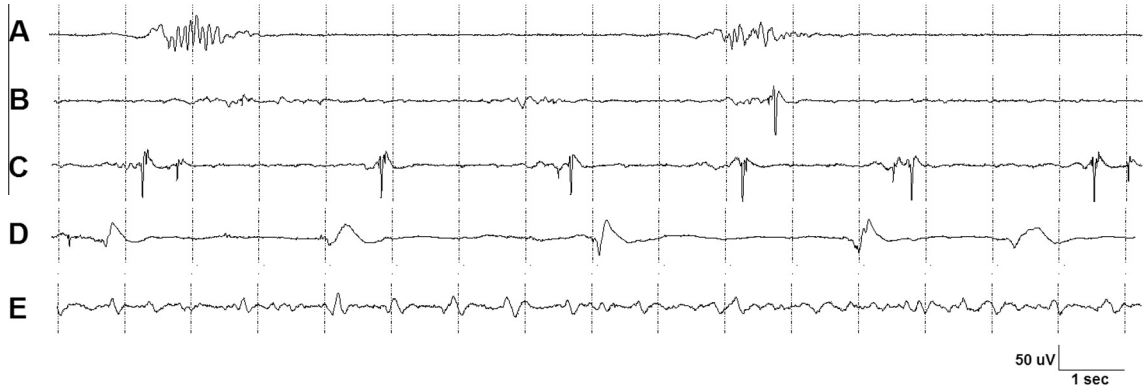

Fig. 4. EEG examples showing morphology differences between burst suppression and periodic pattern detections. (A) Burst suppression pattern; both bursts have a length of one second and more than 3 phases. (B) Burst suppression pattern; the very low burst amplitude requires increased amplitude sensitivity for visual inspection. The last burst includes a single discharge of higher amplitude. (C) Periodic pattern; repetitive high amplitude waveforms of less than 0.5 s length with a surrounding low amplitude burst suppression pattern. As the amplitude of the surrounding burst suppression waveform is negligible, the EEG segment was detected as periodic pattern. (D) Periodic pattern; although the length of the discharges sometimes reaches the limit of burst suppression patterns (0.5 s) all discharges have less than 3 phases and are therefore periodic patterns. The pattern could be misinterpreted as burst suppression by an automatic detection system if the number of phases is not evaluated. (E) Periodic pattern; the low amplitude discharges repeat with an inter discharge interval of less than 1 s.

be analyzed in about 200 ms. The output of the method is based on a single detection segment without using any future information or other detection segments. Together with the delay to wait for 15 s of EEG data the overall processing delay sums up to 15.2 s. Hence, "realtime" monitoring of patients with a constant time delay of 15.2 s is possible.

2.2. Clinical validation

We determined the detection performance by comparing detection results of the presented computational method with EEG annotations of several reviewers. Sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) are measured as defined in Eqs. (1)–(4) of Table 1. The inter-rater agreement (IRA) between human annotations and detection results was also quantified (see below).

EEG data of adult critical care patients from three different centers was used for evaluation and is summarized in Table 2. Video-EEGs from the neurological ICU of the Neurological Center Rosenhuegel Vienna and the neurosurgical ICU of the General Hospital Vienna were recorded using a Micromed EEG system (SystemPLUS Evolution 1.04.95) between March 1, 2013 and September 1, 2014. Data was recorded with a sampling rate of 256 Hz using the international 10–20 electrode system. The initial purpose of the recordings was the validation of a method for detection of rhythmic and periodic patterns (Fürbass et al., 2015a; Herta et al., 2015). The data of these two centers was combined for the dataset named VIEN. The EEG data of the third center was recorded at the Massachusetts General Hospital (MGH) between August 2010 and March 2012. The EEGs from critically ill neurological patients were identified by retrospective review of clinical EEG reports. All of these EEGs included burst suppression activity and were used to validate a real-time burst suppression segmentation method (Westover et al., 2013). All these EEGs were recorded at 256 Hz using XLTEK clinical EEG equipment (Natus Medical Inc., Oakville, Canada) with silver/silver chloride electrodes in the international 10–20 electrode system. In this work the dataset was named MGH.

The EEGs of the dataset VIEN were independently annotated by two clinical neurophysiologists (JH, JK). To reduce the workload for annotation of long-term recordings the first minute of each hour was annotated resulting in 3969 annotation segments. The EEG software package encevis (www.encevis.com) was used to annotate these one-minute EEG segments. The reviewers were able to choose between the choices "EEG with burst suppression patterns" and "EEG without burst suppression patterns" for each segment.

The EEGs of the dataset MGH were likewise independently annotated by two experienced clinical neurophysiologists (BW, MS). They were asked to mark the beginning and end of all suppression events; all non-suppression segments were defined as bursts (Westover et al., 2013). For this work the result of the annotation procedure at MGH was available as time series defining one of three states for each time point: (1) BW and MS agree on suppression, (2) BW and MS agree on burst, (3) disagreement between BW and MS. The EEGs from MGH where then split into consecutive

Table 1
Statistical equations.

Sensitivity	$SE = \frac{TP}{TP+FN}$	(1)
Specificity	$SP = \frac{TN}{TN+FP}$	(2)
Positive predictive value	$PPV = \frac{TP}{TP+FP}$	(3)
Negative predictive value	$NPV = \frac{TN}{TN+FN}$	(4)
95% CI for probabilities $\hat{p}$ (SE, SP, PPV, NPV)	$CI_{95\%,\hat{p}} = \hat{p} \pm \left(\frac{1}{2n} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$	(5)
Cohen's κ value	$\kappa = \frac{(p_o - p_e)}{(1-p_e)}$	(6)
Standard deviation of Cohen's κ value	$SD_\kappa = \sqrt{\frac{p_o(1-p_o)}{n(1-p_e)^2}}$	(7)
95% CI for Cohen's κ value	$CI_{95\%,\kappa} = \kappa \pm 1.96 SD_\kappa$	(8)

The numbers of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) events are used to calculate sensitivity, etc. The confidence interval for point estimates of probabilities like the sensitivity (SE) involves the number of samples (n) that were used to calculate the parameter (i.e. TP + FN for SE). The 95% confidence interval of the Cohen's κ value is given by the approximated standard deviation SD_κ.

Table 2
Summary of EEG data used for validation.

Recording site	Dataset name	Patients (n)	Hours of EEG monitoring (h) (min, mean, max)	Annotation segments (n)	Segments with consensus annotations (n) (%)
Neurological Center Rosenhuegel Vienna, General Hospital Vienna	VIEN	68	3969 (4, 74, 388)	3969	440 (11%)
Massachusetts General Hospital Boston	MGH	20	12.9 (0.34, 0.63, 1.26)	774	597 (77%)
Σ	VIEN + MGH	88	3982	4743	1037 (22%)

The recording sites with the resulting datasets and the number of annotation segments that were each reviewed by two EEG experts are shown. The number of segments with burst suppression patterns (BSP) is low (11%) for centers that prospectively collected data without using an exclusion criteria on the content of the EEG. The high number of segments including BSP of EEGs from MGH (77%) is based on retrospective review and selection of patients with BSP. The dataset combining all three recording centers is called VIEN + MGH and includes EEGs of 88 patients having BSP in 22% of the annotation segments.

segments of one-minute where the annotation "EEG with burst suppression patterns" was assigned to segments including at least one event of type 1 (BW and MS agree on suppression) and one event of type 2 (BW and MS agree on burst). For all other one-minute segments the annotation "EEG without burst suppression patterns" was assigned.

The burst suppression detection method was applied to all EEGs using a computer cluster that processed the digital EEG data using the detection software module. The results are stored in an SQLite (www.sqlite.org) database format with one detection result for each non-overlapping EEG segment of 15 s length. The results of the annotation sessions and the results of the computational analysis were read by an evaluation script written in Matlab (Natus, MA). Statistical formulas were calculated in Matlab.

2.3. Statistical analysis

Statistical analysis of detection performance was done by comparing the annotations in the one-minute annotation segments to the detection results of the computational method. Each EEG segment annotated as "EEG with burst suppression" with overlapping any burst suppression detection of 15 s length was defined as a true positive (TP) event. EEG segments annotated as "EEG with burst suppression" without any overlapping burst suppression detection were defined as false negatives (FN). Segments annotated as "EEG without burst suppression patterns" and with overlapping burst suppression detection were defined as false positives (FP). All other segments were defined as true negatives (TN).

The statistical parameters SE, SP, PPV, and NPV were calculated including the events of all annotation segments of the respective dataset (Eqs. (1–4) of Table 1). The utilization of an arithmetic mean over patient wise results to estimate the expected value was avoided (see Section 3.2). To define the 95% confidence interval for these measures the equation for confidence interval calculation of probabilities (Weiß and Rzany, 2013) is used (Eq. (5)).

The inter rater agreement (IRA) was evaluated by matching the human annotations segments with the detection segments of the same kind calculated by the computational method. The Cohen's κ value was used to quantify the IRA, which is calculated by comparing the difference of the agreement observed, p_o, and the estimate of the expected percent agreement, p_e, divided by the normalization value $(1 - p_e)$ (Eq. (6) of Table 1). The confidence interval for the Cohen's κ value uses an approximation formula for the standard deviation (Cohen, 1960) and is given in Eqs. (7) and (8) of Table 1.

2.4. Analysis of periodic pattern rejection ratio

Periodic patterns represent another important type of electrographic activity that is frequently found in the EEG of critically ill patients. The morphology of periodic pattern can show similarities to BSP but is categorized separately in the ACNS critical care EEG terminology (see Section 1). Periodic patterns were annotated in the dataset VIEN by two EEG reviewers in the previous work of Herta (Herta et al., 2015). These annotations were used to investigate how sensitively the burst suppression detection method reacts to periodic patterns. Annotation segments that were concordantly annotated by two reviewers as "EEG without burst suppression patterns" and were concordantly annotated in our previous work as EEG with a periodic pattern are compared to the results of the automatic burst suppression method. The number of these segments without detection divided by the number of all such segments defines the periodic pattern rejection ratio. High values imply robustness of the method against confusing periodic patterns with burst suppression patterns.

3. Results

3.1. Inter-rater agreement of annotations

The EEGs in dataset VIEN were annotated by two reviewers (JH, JK) that were able to choose between BS (EEG with burst suppression) and $\overline{BS}$ (EEG without burst suppression). The inter-rater agreement shows substantial agreement with a κ value of 0.71 (0.68–0.74). Table 3 shows the detailed results.

The EEGs of the dataset MGH were annotated by the two reviewers (BW, MS) which had to mark the start and end time points of burst and suppression events. The inter-rater agreement of the two reviewers was analyzed in (Westover et al., 2013) and showed an average κ value of 0.57 (min 0.05, max 0.89).

By weighting the 3969 review segments from dataset VIEN with $\kappa = 0.71$ and the 774 review segments from dataset MGH with $\kappa = 0.57$ an average agreement of all review segments can be calculated with the equation: $\bar{\kappa} = 0.71\frac{3969}{4743} + 0.57\frac{774}{4743} = 0.687$. The average agreement of the burst suppression annotations of two reviewers on 4743 one-minute review segments is therefore $\bar{\kappa} = 0.69$.

3.2. Suitable statistical analysis for burst suppression patterns

The kind of statistical analysis that is suitable for a problem depends on the distribution and prevalence of the events under investigation. To gain more insight into the prevalence of burst suppression patterns we analyzed the percentage of segments with

Table 3
Inter-rater agreement of annotations in dataset VIEN.

		JK	
		BS	$\overline{BS}$
JH	BS	440	199
	$\overline{BS}$	91	3239
		$\kappa = 0.71$ (0.68–0.74)	

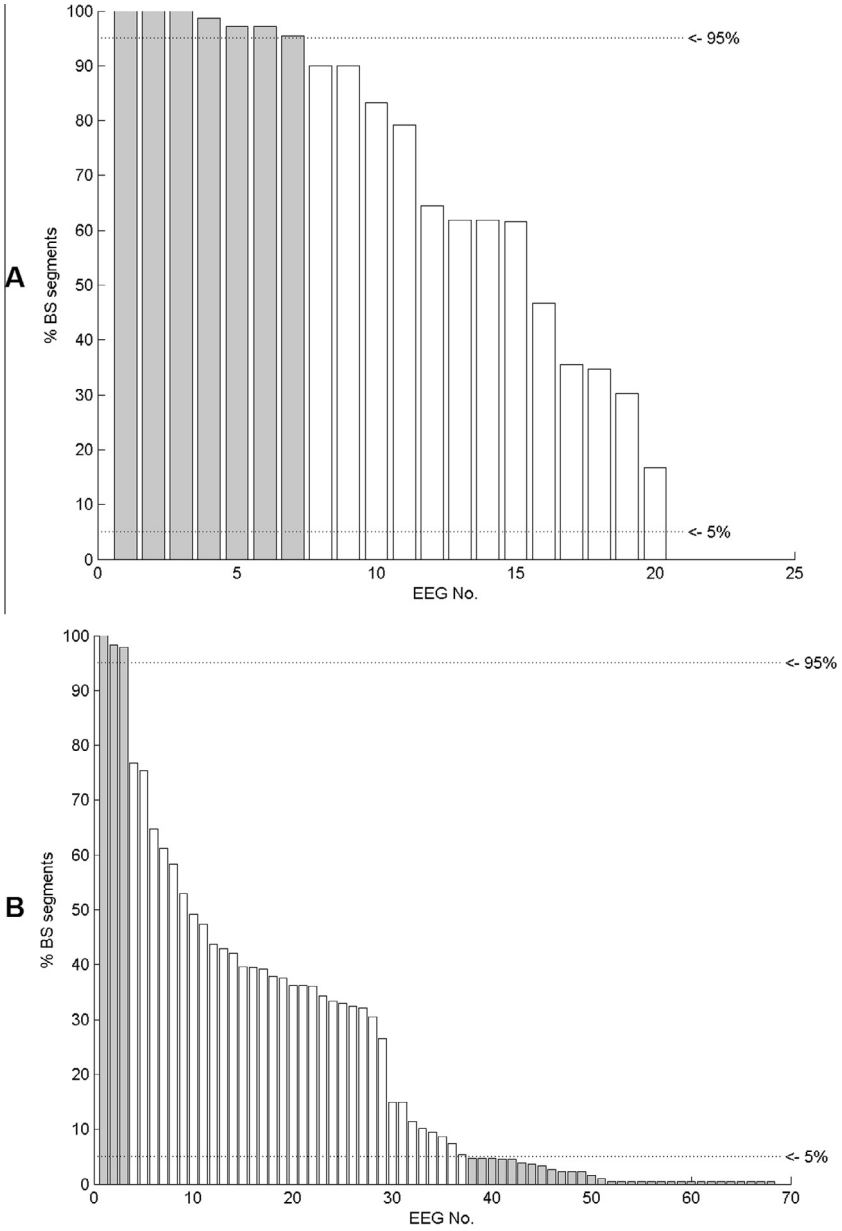

Fig. 5. Prevalence of segments with burst suppression patterns (BS) for each patient in dataset MGH (A) and dataset VIEN (B). Patients with less than 5% or more than 95% of burst suppression segments are marked in grey color.

burst suppression patterns per patient. We used all segments with concordant annotations of both reviewers to define the number of segments with and without burst suppression patterns. Fig. 5 shows the percentage of burst suppression segments for each patient in dataset MGH and VIEN. The use of statistical values based on a very small number of samples is problematic and has to be avoided. The calculation of the sensitivity solely uses segments with burst suppression annotations which are marginally represented in 30 patients (34%) of the study (shown in Fig. 5). A similar situation arises for the calculation of the specificity which is based on annotation segments without burst suppression patterns. In this study 10 patients (11%) included only a marginal number of annotation segments without burst suppression patterns. Overall, the patient-wise statistic of 41 patients (47%) would be based on very small number of samples. The detection performance is therefore analyzed including all annotation segments of all EEGs without using patient wise statistics.

3.3. Performance of the automatic detection method

The results of the automatic burst suppression detection method were compared to the manual annotations of the reviewers. Table 4 summarizes the results of the measured detection performance of the automatic method. The detection performance was analyzed for each reviewer in the dataset VIEN (reviewer JH and JK) and for their consensus annotations (JH + JK). The consensus annotations only include annotation segments with agreement. The results are quite similar for annotations of rater JH and JK with sensitivities of 89% and 88% and specificities of 84% and 81% respectively. The consensus annotations JH + JK of dataset VIEN result in a higher sensitivity of 92% and a specificity 85% as some segments with more difficult patterns have no agreement and are dropped. The detection performance measured on the annotations of the dataset MGH showed a similar sensitivity as the VIEN dataset but a lower specificity of 68%. The positive predictive value

Table 4
Performance of the automatic burst suppression detection method.

Performance measures					Reviewers and EEG data		
SE (%) ($CI_{95\%}$)	SP (%) ($CI_{95\%}$)	PPV (%) ($CI_{95\%}$)	NPV (%) ($CI_{95\%}$)	κ (%) ($CI_{95\%}$)	Rev. (n)	Rev. IDs	Dataset
89 (86–91)	84 (83–85)	51 (48–54)	97 (97–98)	56 (53–59)	1	JH	VIEN
88 (85–91)	81 (80–83)	42 (39–45)	98 (97–98)	47 (44–51)	1	JK	VIEN
92 (89–95)	85 (84–87)	46 (43–49)	99 (98–99)	54 (50–58)	2	JH + JK	VIEN
88 (85–91)	68 (61–75)	90 (88–93)	63 (55–70)	62 (55–68)	2	BW + MS	MGH
90 (88–92)	84 (83–86)	64 (61–66)	96 (96–97)	65 (63–68)	2	JH + JK, BW + MS	VIEN + MGH

Detection performance and agreement between the detection algorithm and the EEG reviewers is shown. Sensitivity (SE), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV) are calculated based on annotations defined by one or two reviewers. The Cohen's κ value measures the level of agreement between the reviewer and the result of the detection algorithm. The number of reviewers (Rev.) of each EEG sample, their IDs and the annotated datasets are shown.

of 90% shows the percentage of correct detections in this dataset. The result of the combined dataset VIEN + MGH using the consensus annotations of JH + JK for the EEGs of VIEN and BW + MS for the EEGs of MGH show a sensitivity of 90% and a specificity of 84%.

3.4. Rejection ratio of periodic patterns

The dataset VIEN was used to evaluate the periodic pattern rejection ratio of the burst suppression detection method. The results of the burst suppression annotation session from two reviewers (JH, JK) were compared to consensus annotations of periodic and rhythmic patterns created in Herta et al. (2015). Of 3969 annotations segments only 17 (0.43%) were annotated as burst suppression and as periodic pattern EEG simultaneously. This shows that periodic patterns and burst suppression patterns are well established terms that are differentiated in clinical practice. We found 230 annotation segments that were concordantly annotated as EEG with periodic patterns and without burst suppression patterns by human reviewers. Only 11 of these 230 segments (5%) included burst suppression detections. 95% included no burst suppression detection. The periodic pattern rejection ratio of the method was therefore 95%.

4. Discussion

Long-term EEG monitoring of critically ill neurological patients has recently received increased attention in the scientific community and in clinical practice. Automatic evaluation of the EEG by computer methods can reduce the burden of visual evaluation and can further raise acceptance of long-term EEG in the critical care unit but needs to be validated in studies with clinical relevance. Burst suppression patterns (BSP) are commonly found in EEGs of anesthetized patients or during pharmacologically induced coma in the treatment of status epilepticus. In this work we evaluated an automatic burst suppression method that was designed to work in the clinical setting.

The initial objective of this work was to develop a robust and universally applicable method for automatic detection and quantification of BSP. The presented computer method and clinical validation methodology contribute in several ways to ongoing work in the field of automatic EEG evaluation.

First, the number of EEG recordings used in this clinical validation study of a BSP detection method exceeds the number used in previous works. We believe that the utilization of EEGs recorded under various clinical and technical conditions contributes to the generalizability of the results. The small confidence intervals of the statistical performance measures confirm that the number of patients was large enough of this detection problem (see supplementary data of Fürbass et al. (2015b)). The annotation of the data was done by different reviewers for the EEGs in dataset VIEN and dataset MGH. Although this may be criticized as problematic, the diverse educational backgrounds of reviewers acts as additional randomization which is generally considered a positive feature. Morphologies of BSP are an especially widely discussed topic in literature as different anesthetics agents and different pathological conditions lead to wide variations in the duration of inter burst intervals and amplitudes (see Section 1), and in the character of activity within bursts. Despite the quite general definition of BSP in the critical care EEG terminology (Hirsch et al., 2013) Zschocke and Hansen (2011) defines three basic types of BSP based on clinical observations. The inter-rater agreement of $\kappa = 0.71$ for a commonly generalized and prolonged EEG pattern like BSP confirms uncertainties in the visual analysis of these patterns.

The EEGs of dataset VIEN were recorded prospectively for the work presented in Herta et al. (2015). The detection performance shows high values for sensitivity and specificity of 92% and 85% which we interpret as an excellent result for a fully automatic detection method. The large percentage of EEG segments without burst suppression help to reduce the confidence interval of the specificity (4% for SE vs. 3% for SP in dataset VIEN + MGH). The result of the κ agreement between human reviewers and the automatic method is more diverse. The highest agreement could be measured between reviewers JH and JK with 0.71 where the high number of segments without burst suppression have a strong bias on this value. Comparing the results of the automatic method to these reviewers resulted in κ values of 0.56 and 0.47 which is significantly lower ($p < 0.05$) but with an acceptable absolute value. The patient wise κ agreement between the reviewers BW and MS was 0.57 with outliers of 0.05 and 0.95. We like to emphasize that the κ value of the human annotations in dataset MGH is based on annotations of separate burst and suppression events which holds more detailed information then the annotations of one-minute segments for dataset VIEN. The lower κ value of annotations in dataset MGH compared to the κ value of annotations in dataset VIEN is therefore based on differences in the time resolution during review. The agreement of the automatic method in dataset MGH could only be measured on the combined annotations of BW and MS and showed a κ value 0.62. This data supports the thesis that automatic detection of BSP can be done with high sensitivity and specificity and at a level of agreement similar to that of two human experts.

This work thoroughly investigates the spatial coverage of BSP with respect to the utilized electrode system. EEG settings with few channels are commonly used to monitor sedation depth and do not require solutions to multi-channel issues of automatic evaluation. The full 10–20 electrode system is used to monitor critically ill patients with suspected seizures to increase the detection sensitivity for focal patterns of brain activity in the EEG. Lateralized burst suppression or bilateral asynchronous burst suppression can only be properly analyzed by computer methods that allow for BSP with a reduced spatial profile. Another frequently observed issue is that human reviewers tend to recog-

nize patterns based on clear cut activity seen in very short time intervals or in a few channels and by extrapolation of this subjective opinion to a more stretched time–channel area. EEG activity like BSP can show altered level of amplitudes over channels that forces the computer method to detect the activity based on the more pronounced EEG channels alone. We are convinced that both reasons explain the fact that our presented computer method needed to allow detections of BSP with an electrode coverage as low as 40%. Further experiments on these settings have shown that by increasing this value to 50% the sensitivity decreased by approximately 10%.

In contrast, automatic detection of spatially limited patterns will reduce the specificity of the detection result in general, by causing a reduced signal-to-noise ratio. As we pointed out in Section 1, the a priori reduction of EEG artifacts with the PureEEG method is used by our detection system. The experimental deactivation of this pre-processing step resulted in decreased specificity and sensitivity, which is explained by the raised level of artifacts that trigger detection as well as artifacts that resemble physiological EEG patterns.

Another contribution of this work is the ability to distinguish between burst suppression patterns and periodic patterns automatically. The critical care EEG terminology of the ACNS (Hirsch et al., 2013) clearly defines these two types of patterns. Periodic and burst suppression patterns may occur in the same patient as reported for some patients with coma following cardiac arrest (Hofmeijer et al., 2014). At least three EEGs recorded in this study exhibited both patterns, as Fig. 1 exemplifies. The clustering approach of this work is able to combine all EEG activity belonging to one burst into a single information unit which leads to simple and robust classification of periodic discharges. Previous reports of automated burst suppression analysis are based on channels-wise analysis that results in a feature time series following the methodology of "single-channel classification with late integration" which differs from this work that is based on early integration of multiple channels (Hunyadi et al., 2011). In summary, the channel-wise detections of burst suppression events, the combination with a spatial clustering algorithm and the use of rejection algorithms for artifacts and periodic patterns represent the key innovation of this work.

5. Conclusion

We presented a fully automated method for detection of burst suppression EEG patterns. The detection performance on EEGs from 88 adult patients from three independent recording sites showed high sensitivity and specificity, comparable to expert–expert levels of inter-rater agreement. The method is able to detect burst suppression patterns even when occurring over limited cortical regions, and is insensitive to EEG artifacts and periodic patterns. In addition the method quantifies the duration of burst and suppression events and works in real-time. The high detection performance on prospectively collected data without the need for patient-specific parameter tuning shows that utilization for clinical patient monitoring of burst suppression patterns is feasible.

Acknowledgements

This work was partly done within the DeNeCoR project which received funding from the ECSEL Joint Undertaking under grant agreement No. 324257 and from the national funding authorities of Austria, Czech Republic, Germany, Italy, the Netherlands, Spain and the United Kingdom.

Conflict of interest: The authors declare that they have no conflicts of interest concerning this article.

References

American Clinical Neurophysiology Society. Guideline 6: a proposal for standard montages to be used in clinical EEG. Am J Electroneurodiagnostic Technol 2006;46:226–30.

Bruhn J, Bouillon TW, Shafer SL. Bispectral index (BIS) and burst suppression: revealing a part of the BIS algorithm. J Clin Monit Comput 2000;16:593–6.

Bruhn J, Myles PS, Sneyd R, Struys MMRF. Depth of anaesthesia monitoring: what's available, what's validated and what's next? Br J Anaesth 2006;97:85–94.

Ching S, Purdon PL, Vijayan S, Kopell NJ, Brown EN. A neurophysiological-metabolic model for burst suppression. Proc Natl Acad Sci USA 2012;109:3095–100.

Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20 (46):37.

Fürbass F, Hartmann MM, Halford JJ, Koren J, Herta J, Gruber A, et al. Automatic detection of rhythmic and periodic patterns in critical care EEG based on American Clinical Neurophysiology Society (ACNS) standardized terminology. Neurophysiol Clin 2015a;45:203–13.

Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. Clin Neurophysiol 2015b;126:1124–31.

Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: automatic EEG artifact removal for epilepsy monitoring. Neurophysiol Clin 2014;44:479–90.

Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. Epilepsy Behav 2015;49:273–9.

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.

Hofmeijer J, Tjepkema-Cloostermans MC, van Putten MJAM. Burst-suppression with identical bursts: a distinct EEG pattern with poor outcome in postanoxic coma. Clin Neurophysiol 2014;125:947–54.

Hunyadi B, Vos MD, Signoretto M, Suykens JAK, Paesschen WV, Huffel SV. Automatic seizure detection incorporating structural information. In: Honkela T, Duch W, Girolami M, Kaski S, editors. Artif neural netw mach learn – ICANN 2011. Springer Berlin Heidelberg; 2011. p. 233–40 [cited 2015 Aug 14]. Available from: http://link.springer.com/chapter/10.1007/978-3-642-21735-7_29.

Jaggi P, Schwabe MJ, Gill K, Horowitz IN. Use of an anesthesia cerebral monitor bispectral index to assess burst-suppression in pentobarbital coma. Pediatr Neurol 2003;28:219–22.

Koolen N, Jansen K, Vervisch J, Matic V, De Vos M, Naulaers G, et al. Line length as a robust method to detect high-activity events: automated burst detection in premature EEG recordings. Clin Neurophysiol 2014;125:1985–94.

Lazar LM, Milrod LM, Solomon GE, Labar DR. Asynchronous pentobarbital-induced burst suppression with corpus callosum hemorrhage. Clin Neurophysiol 1999;110:1036–40.

Lewis LD, Ching S, Weiner VS, Peterfreund RA, Eskandar EN, Cash SS, et al. Local cortical dynamics of burst suppression in the anaesthetized brain. Brain 2013;136:2727–37.

Liang Z, Wang Y, Ren Y, Li D, Voss L, Sleigh J, et al. Detection of burst suppression patterns in EEG using recurrence rate. Sci World J 2014;2014:295070.

Lipping T, Jäntti V, Yli-Hankala A, Hartikainen K. Adaptive segmentation of burst-suppression pattern in isoflurane and enflurane anesthesia. Int J Clin Monit Comput 1995;12:161–7.

Mader EC, Villemarette-Pittman NR, Rogers CT, Torres-Delgado F, Olejniczak PW, England JD. Unihemispheric burst suppression. Neurol Int 2014;6:5487.

Murphy K, Stevenson NJ, Goulding RM, Lloyd RO, Korotchikova I, Livingstone V, et al. Automated analysis of multi-channel EEG in preterm infants. Clin Neurophysiol 2015;126:1692–702.

Nakashima K, Todd MM, Warner DS. The relation between cerebral metabolic rate and ischemic depolarization. A comparison of the effects of hypothermia, pentobarbital, and isoflurane. Anesthesiology 1995;82:1199–208.

Niedermeyer E, Sherman DL, Geocadin RJ, Hansen HC, Hanley DF. The burst-suppression electroencephalogram. Clin Electroencephalogr 1999;30:99–105.

Pagni CA, Courjon J. Electroencephalographic modifications induced by moderate and deep hypothermia in man. Acta Neurochir Suppl 1964;14(Suppl. 13):35–49.

Rossetti AO, Carrera E, Oddo M. Early EEG correlates of neuronal injury after brain anoxia. Neurology 2012;78:796–802.

Shellhaas RA, Chang T, Tsuchida T, Scher MS, Riviello JJ, Abend NS, et al. The American Clinical Neurophysiology Society's guideline on continuous electroencephalography monitoring in neonates. J Clin Neurophysiol 2011;28:611–7.

Sperling MR, Brown WJ, Crandall PH. Focal burst-suppression induced by thiopental. Electroencephalogr Clin Neurophysiol 1986;63:203–8.

Thomsen CE, Rosenfalck A, Nørregaard Christensen K. Assessment of anaesthetic depth by clustering analysis and autoregressive modelling of electroencephalograms. Comput Methods Programs Biomed 1991;34:125–38.

Weiß C, Rzany B. Basiswissen medizinische statistik. 6 überarb. Aufl. Springer; 2013.

Westover MB, Shafi MM, Ching S, Chemali JJ, Purdon PL, Cash SS, et al. Real-time segmentation of burst suppression patterns in critical care EEG monitoring. J Neurosci Methods 2013;219:131–41.

Westover MB, Ching S, Kumaraswamy VM, Akeju O, Pierce E, Cash SS, et al. The human burst suppression electroencephalogram of deep hypothermia. Clin Neurophysiol 2015;126:1901–14.

Zschocke S, Hansen H-C. Klinische elektroenzephalographie. 3 aktualisierte und erweiterte Auflage. Berlin, Heidelberg: Springer; 2011.

6.5 Paper A5: Multimodal seizure detection

Title: Automatic multimodal detection for long-term seizure documentation in epilepsy

Authors: Fürbass F, S. Kampusch, E. Kaniusas, J. Koren, S. Pirker, R. Hopfengärtner, H. Stefan, T. Kluge, C. Baumgartner

Published in: Clinical neurophysiology

Year: 2017

Authors' contribution: Fürbass Franz developed the algorithm for detection of epileptic seizures. In addition data evaluation and statistical analysis was done by Franz Fürbass. Initial writing of the manuscript, implementation of corrections from coauthors and reviewers was done by Franz Fürbass. Kampusch S. participated in the development of ECG based methods which is part of the detection algorithm. In addition Kampusch S. and Kaniusas E worked on corrections of the manuscript. Data annotations, EEG recording, parts of the corrections, and the clinical evaluation were done by other coauthors.

Clinical Neurophysiology 128 (2017) 1466–1472

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Automatic multimodal detection for long-term seizure documentation in epilepsy

F. Fürbass [a,*], S. Kampusch [b], E. Kaniusas [b], J. Koren [c,e], S. Pirker [c,e], R. Hopfengärtner [f], H. Stefan [f], T. Kluge [a], C. Baumgartner [c,d,e]

[a] Center for Health & Bioresources, AIT Austrian Institute of Technology GmbH, Vienna, Austria
[b] Institute of Electrodynamics, Microwave and Circuit Engineering, TU Wien, Vienna, Austria
[c] Department of Neurology, General Hospital Hietzing with Neurological Center Rosenhuegel, Vienna, Austria
[d] Department of Epileptology and Clinical Neurophysiology, Sigmund Freud University, Vienna, Austria
[e] Karl Landsteiner Institute for Clinical Epilepsy Research and Cognitive Neurology, Vienna, Austria
[f] Epilepsy Center, Department of Neurology, University Hospital Erlangen, Germany

ARTICLE INFO

Article history:
Accepted 15 May 2017
Available online 25 May 2017

Keywords:
Multimodal
Automatic
Seizure detection
Algorithm
ECG
EMG
EEG

HIGHLIGHTS

- Automatic seizure detection assessing efficacy of EEG/ECG/EMG signals for seizure documentation.
- Multi-center evaluation including 92 patients with 494 seizures comparing full to reduced montages.
- Using 8 frontal and temporal electrodes will significantly improve conventional seizure reporting.

ABSTRACT

Objective: This study investigated sensitivity and false detection rate of a multimodal automatic seizure detection algorithm and the applicability to reduced electrode montages for long-term seizure documentation in epilepsy patients.
Methods: An automatic seizure detection algorithm based on EEG, EMG, and ECG signals was developed. EEG/ECG recordings of 92 patients from two epilepsy monitoring units including 494 seizures were used to assess detection performance. EMG data were extracted by bandpass filtering of EEG signals. Sensitivity and false detection rate were evaluated for each signal modality and for reduced electrode montages.
Results: All focal seizures evolving to bilateral tonic-clonic (BTCS, $n = 50$) and 89% of focal seizures (FS, $n = 139$) were detected. Average sensitivity in temporal lobe epilepsy (TLE) patients was 94% and 74% in extratemporal lobe epilepsy (XTLE) patients. Overall detection sensitivity was 86%. Average false detection rate was 12.8 false detections in 24 h (FD/24 h) for TLE and 22 FD/24 h in XTLE patients. Utilization of 8 frontal and temporal electrodes reduced average sensitivity from 86% to 81%.
Conclusion: Our automatic multimodal seizure detection algorithm shows high sensitivity with full and reduced electrode montages.
Significance: Evaluation of different signal modalities and electrode montages paces the way for semi-automatic seizure documentation systems.

1. Introduction

Seizure documentation and quantification represents the primary outcome measure of epilepsy therapy including antiepileptic drug treatment, epilepsy surgery, and neurostimulation. Presently, patients document their seizures using seizure diaries without systematic and objective validation approach by physicians. Recent publications showed that manual seizure counting suffers from underreporting with sensitivities of 50% during day and as low as 30% during night and can therefore be considered as highly unreliable (Blachut et al., 2015). This inaccuracy represents a major issue for the assessment of treatment efficacy including drug trials.

* Corresponding author at: AIT Austrian Institute of Technology, Donau-City-Straße 1, 1220 Vienna, Austria.
E-mail address: franz.fuerbass@ait.ac.at (F. Fürbass).

http://dx.doi.org/10.1016/j.clinph.2017.05.013
1388-2457/

We propose a semi-automatic system for seizure documentation and quantification based on computer methods to scan biomedical signals for epileptic seizures followed by a manual evaluation of these detections by trained staff. For this application a low number of sensors should be used to assure patient compliance and to simplify hardware design. On the other hand, data from ictal events needs to be recorded with a reasonable number of sensors to allow post-hoc analysis for correct seizure identification. A prerequisite for this approach is a wearable electrophysiological hardware setup that can be utilized over long time periods. Secondly, and with utterly importance, a clinically validated computer based detection method has to be used. This method has to ensure high sensitivity and low false detection rates, to pay off additional efforts of neurophysiological measurements with numerous EEG electrodes and other sensors.

EEG represents the gold standard in epilepsy diagnosis and to prove the epileptic nature of seizures which makes it the primary modality for automatic seizure documentation. Automatic seizure detection methods based on surface EEG recorded during inpatient epilepsy monitoring showed high sensitivity in multi-center studies (Fürbass et al., 2015a; Hopfengärtner et al., 2014). Reduced EEG electrode sets showed a rapid drop in detection sensitivity for rhythmic patterns (Herta et al., 2017) which has to be considered for wearable documentation devices.

ECG can be utilized as another modality for seizure detection. Epileptic seizures cause an activation of the central autonomic network (CAN) resulting in changes in heart rhythm at seizure onset. Ictal tachycardia (ITC) represents the most frequent change in heart rhythm and can be observed in 65–86% of seizures (Eggleston et al., 2014; Leutmezer et al., 2003). Furthermore, a larger affected brain area was reported to define the degree and rate of ITC (Stefanidou et al., 2015). ITC occurs early during seizure evolution and often even precedes EEG changes visible on scalp-EEG (Leutmezer et al., 2003). The high sensitivity of ITC, its early occurrence, and the easy technical setup for ECG measurement makes this biomarker highly promising for automatic seizure detection devices.

Other modalities for automatic seizure detection were investigated recently, including methods based on surface EMG (Beniczky et al., 2016) and motion sensors (Conradsen et al., 2012) as well as gyroscopic sensors and dermal skin conductance sensors (Banks et al., 2014).

In this study we present a multimodal automatic seizure detection method using information from EEG, ECG assessing ictal tachycardia and EMG measuring ictal tonic muscle activity. We investigated this method both with a full 10–20 electrode set as well as a reduced number of EEG electrodes suitable for ambulatory settings. We assessed strengths and weaknesses of this approach in patients with specific seizure and epilepsy types.

2. Methods

2.1. Data

We retrospectively analyzed 92 long-term EEG/ECG/EMG recordings from two epilepsy monitoring units including at least 21 EEG electrodes and at least one ECG channel. Signed informed consent was obtained from all patients. We included all available EEG recordings with one or more epileptic seizure during the recording period resulting in a total of 11,978 h of data with 494 epileptic seizures of various types (min per patient = 23 h, max per patient = 547 h). From 92 patients included in our study 55 patients had temporal lobe epilepsy (TLE) and 37 patients had extratemporal lobe epilepsy (XTLE). Data were recorded with a Micromed (Veneto, SpA) and an ITmed (Natus Medical Incorporated) system at a 256 Hz sampling rate using gold-disc electrodes placed according to the international 10–20 system with additional temporal electrodes. To mimic the behaviour of prospective data, digital EEGs were analyzed without manual pre-processing, data selection or data cutting.

The effect of reduced scalp electrode montages was simulated by removing electrodes from the digital EEG file before further analysis. Two different montages with reduced number of electrodes were assessed: the **8 electrode forehead montage** including electrodes FP1, F7, T7, FP2, F8, T8, FZ, ECG and the **7 electrode posterior montage** including electrodes T7, P7, O1, T8, P8, O2, ECG. Fig. 1 shows standard electrode positions (circles) as well as electrodes of forehead montage (dashed circles) and electrodes of posterior montage (shaded circles).

2.2. Performance evaluation methodology

Seizures were annotated following standard protocols of the two epilepsy monitoring units using both clinical and EEG information. The first three seizures of each patient were categorized according to the ILAE operational seizure classification (http://www.ilae.org/Visitors/Centre/documents/ClassificationSeizureILAE-2016.pdf) in order to facilitate performance evaluation according to seizure type. Seizure markers were set based on standard EMU review procedure using video, EEG, and other clinical information including manual validation of seizures by an experienced clinical epileptologists (HS, SP, or CB). Only validated seizure markers were used to define seizure epochs as basis for assessing detected and undetected seizures. Each seizure epoch ranged from 30 s before the clinical seizure marker to 180 s after this marker resulting in a total 210 s intervals of single seizure epochs.

Our seizure detection algorithm provided both time points and modality of detection. Time points of detected events were compared to the visually identified seizure epochs. Seizure epochs were defined as true positive (TP) if at least one detection occurred within the epoch time range. Detections outside of seizure epochs were defined as false positives (FP). Seizure epochs without a matching detection were defined as false negative (FN). For assessment of detection performance according to seizure types we distinguished between focal seizures (FS group) and focal seizures evolving to bilateral tonic-clonic (BTCS group). The first three seizure epochs including seizure type annotations in each patient were evaluated, consecutive seizure epochs and detections overlapping these epochs were ignored. Patients with at least one seizure of a certain type were included in the corresponding seizure type group. Patients having two different seizure types were included in both seizure type groups.

Sensitivity (**SE**) was defined as the ratio between the number of true positives (#TP) and the number of all seizures (#TP + #FN) and was calculated for each patient. False detection rate was defined as the number of false detections per 24 h (**FD/24 h**).

A paired *t*-test was used as test statistic between performance results of two detector types or electrode sets.

2.3. Computer algorithm

The computer algorithm detects seizures using EEG, surface EMG, and ECG signals that were recorded using scalp EEG and chest ECG electrodes. Fig. 1 gives an overview of the detection system.

EEG is able to pick up pathologic brain activity by showing rhythmic signal components, but patient movements and loose electrode contacts can cause signal artefacts with similar morphology. Before applying the EEG seizure detection algorithm artefacts were removed applying PureEEG, a fully automatic artefact

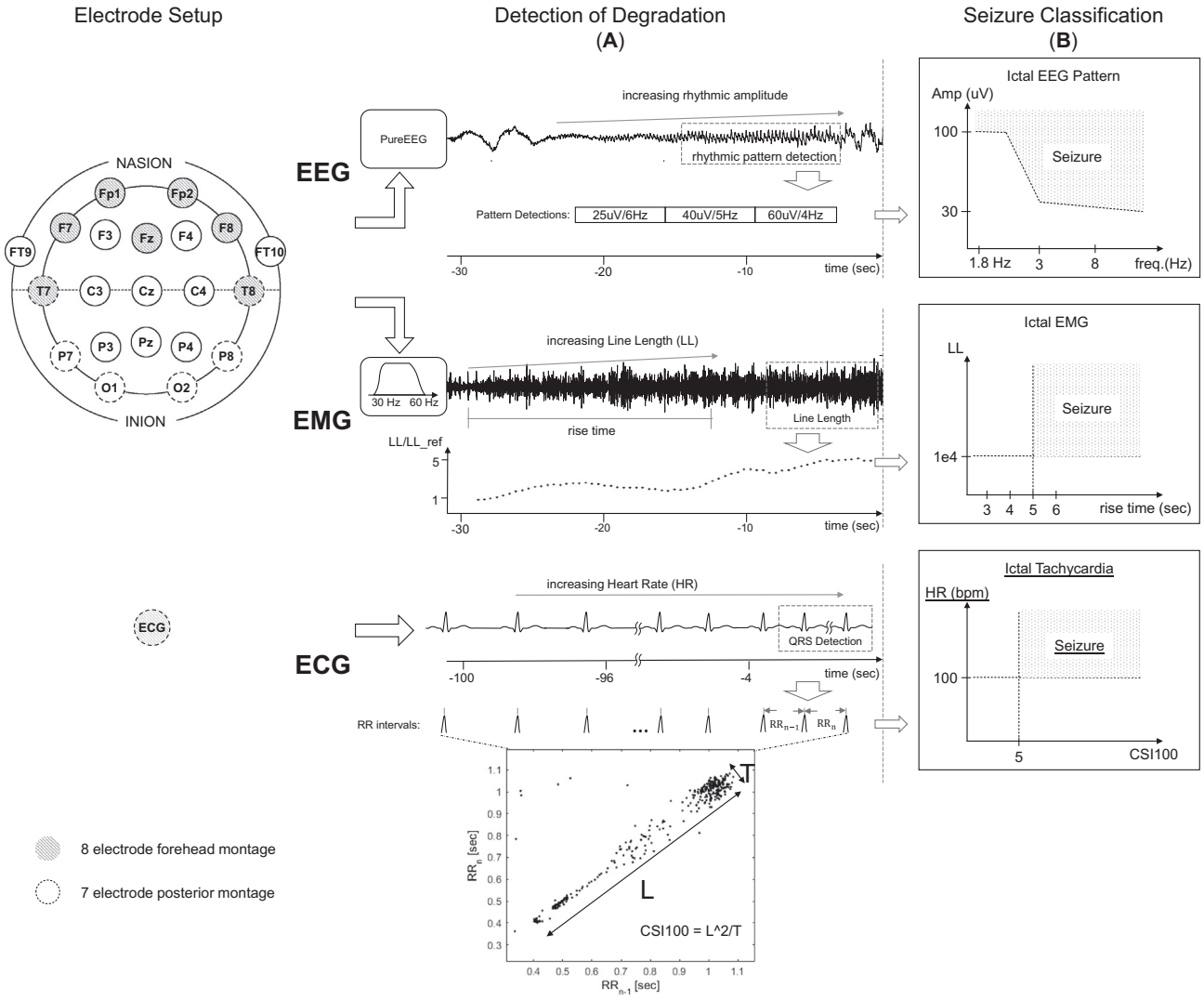

Fig. 1. Multimodal Seizure Detection System: Signal modalities EEG, surface EMG, and ECG are derived from scalp EEG and chest ECG leads. For each modality, seizure specific features of the current time point are compared to past values to detect an increasing seizure likelihood called degradation (A) and for real time seizure classification (B). Detection events were defined as logical AND of conditions A and B. For **EEG** the increasing rhythmic signal amplitude (A) with a high absolute amplitude compared to an average EEG spectrum (B) triggered detections; for **EMG** an elevated Line Length (LL) compared to baseline (LL_ref) (A) and steady increasing tonic activity for more than 5 s with high absolute values (B) triggered detections; for **ECG** elevated heart rate (HR) compared to baseline (A) and increased heart rate above 100 beats per minute (bpm) with a high cardiac sympathetic index of 100 beats (CSI100) (B) triggered detections.

removal method (Hartmann et al., 2014). A rhythmic pattern detection algorithm described previously (Fürbass et al., 2015b; Herta et al., 2015) was then used to detect rhythmic activity between 1.8 and 12.5 Hz and to measure amplitude and frequency of these patterns. Amplitude baseline was estimated using the 50% percentile of all rhythmic patterns that occurred in the previous four minutes. Rhythmic patterns that were classified as ictal EEG patterns (Fig. 1, (B)) and that had a 40% higher amplitude compared to EEG baseline were defined as seizure detections.

Automatic seizure detection on EMG was based on the occurrence of sustained and excessive EMG activity. EMG signals were extracted from data recorded on EEG electrodes by bandpass filtering the signals between 30 and 60 Hz. Signal strength was quantified using the line length method defined as the sum of distances between each consecutive data sample in non-overlapping 0.5 s windows. Seizure events were defined as high absolute line length values (LL), a steady increase over 5 s, and a 500% increase compared to maximum line length in a four minute baseline window (LL_ref).

ECG signals from a single chest electrode were used for measuring heart rate and for automatic detection of ictal tachycardia. We defined ictal tachycardia as a heart rate above 100 beats per minute (bpm). The detection algorithm first resampled ECG signals to 500 Hz and then high pass filtered the signal with a cut off frequency of 8 Hz to remove T wave components. Then a detection algorithm designed to find periodic patterns scanned for QRS complexes (Fürbass et al., 2015b). The exact time position of R peak was defined at the maximum of the QRS complex. Consecutive R to R time intervals (RR intervals) of the last 10 s were used to define the average bpm at each time point. Cardiac baseline activity was defined as average heart rate during four minutes before the current time point. To differentiate physiologic from ictal activity the modified cardiac sympathetic index based on previous 100 RR intervals (CSI100) was calculated (Jeppesen et al., 2014) as follows: given the Lorenz plot of RR intervals the longitudinal length (L) and the transversal length (T) was estimated as four times the standard deviation. The value of CSI100 was then calculated by L^2 divided by T. An elevated heart rate of more than 100 bpm

and a minimum increase of 30% compared to baseline as well as a CSI100 value above 5 defined a seizure event.

The logical OR combination of the three detection modalities was used for evaluation of the overall system performance. Detection results can be obtained in real time with less than 10% CPU usage of a standard PC; or equivalently a 24 h EEG recording needs approximately 2 h of calculation time. The algorithm will be part of the encevis EEG software package (www.encevis.com).

3. Results

3.1. Detection performance

Assessment of overall detection performance in 92 patients including 494 epileptic seizures resulted in 86% sensitivity (**SE**) and an average of 16.5 false detections per 24 h (**FD/24 h**). Evaluation of TLE patients including 284 epileptic seizures resulted in 94% SE and 12.8 FD/24 h, XTLE patients ($n = 37$) including 210 seizures showed a sensitivity of 74% and 22.2 FD/24 h. Evaluation according to seizure types involved a maximum of 3 epileptic seizures per patient (see Section 2.2). The focal seizure (**FS**) group included 64 patients with 139 seizures and resulted in 89% SE and 16.4 FD/24 h. On the other hand, evaluation of 35 patients with 50 focal seizures evolving to bilateral tonic-clonic seizures (**BTCS**) resulted in 100% SE and 14.1 FD/24 h.

ECG based seizure detection resulted in low sensitivities (TLE: 40%, XTLE: 8%, FS: 27%, BTCS: 43%). EMG based seizure detection reached high sensitivity for BTCS (93%) but low sensitivities for other patients and seizure types (TLE: 25%, XTLE: 35%, FS: 8%). EEG based seizure detection showed high sensitivities in general (TLE: 91%, XTLE: 74%, FS: 88%, BTCS: 97%).

Electrode set reduction using a **8 electrode forehead montage** including frontal and temporal as well as ECG electrodes (FP1, F7, T7, FP2, F8, T8, FZ, ECG) resulted only in statistically non-significant lower detection sensitivity ($p > 0.05$) for XTLE (−6% SE), FS (−5% SE), and BTCS (−3% SE). Significant reduction by −5% was found when all patients and seizures were used (group ALL) with $p = 0.02$ and for TLE (−6% SE) with $p = 0.01$.

Reduction to **7 electrode posterior montage** including temporal and occipital as well as ECG electrodes (T7, P7, O1, T8, P8, O2, ECG) showed even lower detection sensitivity which was non-significant ($p > 0.05$) only for patients with BTCS (−6% SE), whereas ALL (−12% SE), TLE (−10% SE), XTLE (−15% SE), and FS (−11% SE) showed a significant reduction ($p < 0.05$).

Table 1 summarizes the results separate for different detection modalities (EEG, ECG, EMG), combination of modalities (EEG + EMG, **COMB** defined as EEG + ECG + EMG) based on data of five different evaluation groups (ALL, TLE, XTLE, FS, BTCS). Detection performance using the full 10–20 electrode set including 21 EEG and 1 ECG electrode (22 electrode montage) as well as the 8 electrode forehead montage and the 7 electrode posterior montage are shown.

Fig. 2 visualizes detection performance by receiver operating characteristic plots (ROC) for full 22 electrode and the 8 electrode forehead montage. The 95% confidence intervals for sensitivity values are shown using vertical error bars of the COMB detector. Comparing COMB performance to EEG + EMG combination shows the added value of an ECG based detection system.

3.2. Detection delays

Time delay of seizure detections are of minor importance to our proposed semi-automatic seizure documentation approach but will be in focus of ambulatory seizure alarming devices. In this section we elaborate on detection delays to get more insights into this closely related and important topic. Comparing time delays of automatically calculated detections to visually selected seizure markers indicated a correlation of average delays with detection modalities. Fig. 3 shows boxplots of detection delays in seconds of all detected seizures based on the full 22 electrode montage. ECG based detections had a median delay of only 19 s (min = −22 s, max = 75 s) followed by EEG based detections (median = 26 s, min = −10 s, max = 165 s), and surface EMG based detections (median = 45 s, min = 6 s, max = 141 s). Negative delays indicate detection of seizures prior to visual identification on scalp-EEG or video, and were found in 16 seizures (ECG = 12, EEG = 4, EMG = 0 seizures). In this work the detection horizon prior to visual

Table 1
Detection performance of our automatic seizure detection algorithm based on data of a 22 electrode montage and two reduced electrode montages (rows). Average sensitivity in percent (**SE (%)**) and false detections per 24 h (**FD/24** h) are shown for five different evaluation groups (columns). Group **ALL** includes all patients, **TLE** and **XTLE** patients with respective epilepsy types, **FS** (focal seizure) and **BTCS** (focal seizures evolving to bilateral tonic-clonic) patients with respective seizure types. Number of patients (n) and number of seizures (nSz) are shown for each evaluation group. Combined detector (**COMB**) performance was defined as the combination of **EEG**, **ECG**, and surface **EMG** based detections.

Automatic seizure detection performance										
	ALL $n = 92$ nSz = 494		TLE $n = 55$ nSz = 284		XTLE $n = 37$ nSz = 210		FS $n = 64$ nSz = 139		BTCS $n = 35$ nSz = 50	
	SE (%)	FD/24 h	SE (%)	FD/24 h	SE (%)	FD/24 h	SE (%)	FD/24 h	SE (%)	FD/24 h
22 electrode montage										
EEG	84	15.6	91	11.9	74	21.2	88	15.5	97	13.3
ECG	27	0.6	40	0.6	8	0.6	27	0.7	43	0.4
EMG	29	0.4	25	0.3	35	0.6	8	0.4	93	0.5
EEG + EMG	84	16.0	92	12.2	74	21.7	88	15.9	100	13.8
EEG + ECG + EMG (COMB)	86	16.5	94	12.8	74	22.2	89	16.4	100	14.1
8 electrode forehead montage										
EEG	79	11.5	87	8.5	67	16	82	11	97	11.2
ECG	27	0.6	40	0.6	8	0.6	27	0.7	43	0.4
EMG	34	1.6	29	1.3	41	2.1	17	1.4	96	2.4
EEG + EMG	81	12.1	90	9.1	67	16.6	84	11.6	97	11.5
EEG + ECG + EMG (COMB)	81	13.5	90	10.3	68	18.3	84	12.8	97	13.7
7 electrode posterior montage										
EEG	68	4.2	76	3.4	55	5.4	72	4.0	94	3.8
ECG	27	1.2	40	0.6	8	0.6	27	0.7	43	0.4
EMG	28	1.2	23	1.2	35	1.5	8	1.1	93	1.7
EEG + EMG	69	5.4	77	4.5	58	6.9	73	5.1	94	5.4
EEG + ECG + EMG (COMB)	74	6.0	84	5.1	59	7.4	78	5.7	94	5.8

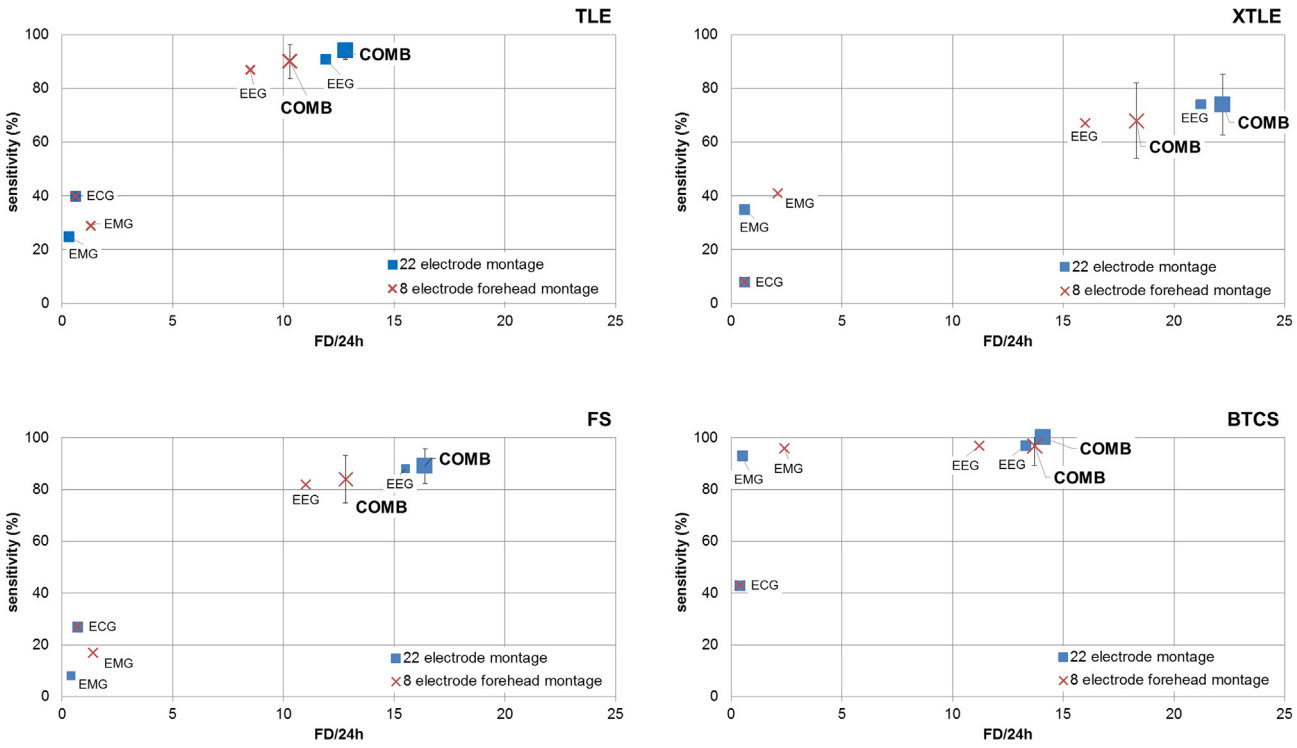

Fig. 2. Detection performance by means of sensitivity and false detections per 24 h (**FD/24** h) shown in ROC plots. The left upper corner of each plot defines the theoretical optimum point with 100% sensitivity and no false detections. Results for our seizure algorithm based on different modalities (EEG, EMG, ECG) and their logical OR combination (**COMB**) are shown. Data of the 22 electrode montage (boxes) and the 8 electrode forehead montage (crosses) is shown. Both montages include the same ECG data wherefore ECG based detection performance results in the same values (box overlaid with cross labelled ECG). Vertical error bars on the COMB values indicate the 95% confidence intervals of sensitivities. All focal seizures evolving to bilateral tonic-clonic (BTCS) are detected with the 22 electrode montage and 97% of BTCS using a reduced forehead montage. The ECG based detector stays below 43% sensitivity in all evaluation groups with false alarm rates below 0.7 FD/24 h.

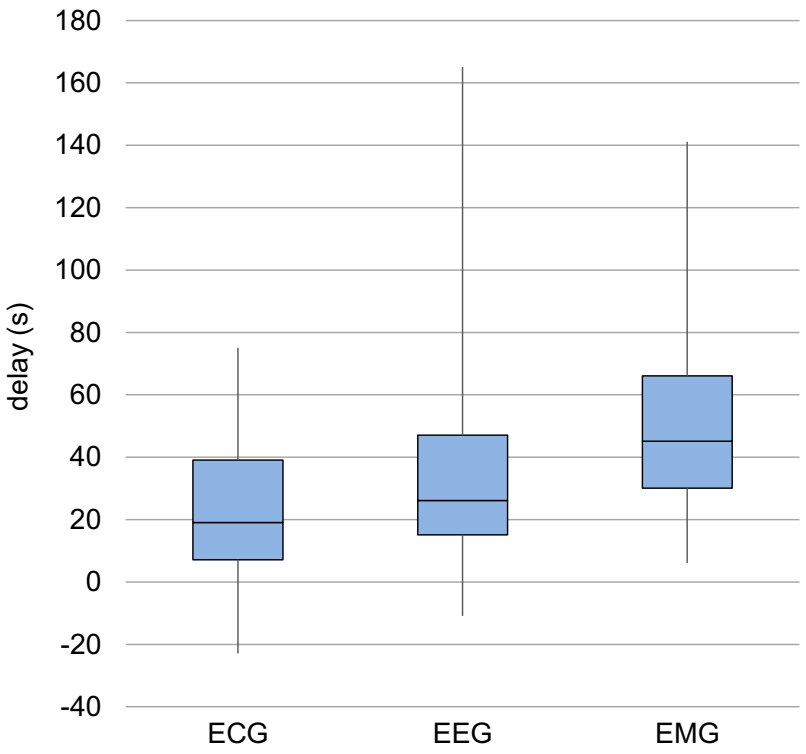

Fig. 3. Box plots represent time delay of seizure detections for different modalities compared to seizure markers set by clinicians. Whiskers of each box plot include the minimum and maximum value.

identification on scalp-EEG or video was limited to 30 s because of the definition of the seizure epoch (see Section 2.2).

Fig. 3 shows that some detections triggered by the ECG or EEG signal even occurred before clinical onset (whiskers include the minimum and maximum value). Median delays of ECG (19 s), EEG (26 s), and EMG (45 s) detections show that ictal ECG features appear earlier in time compared to EEG based features (although less frequent, see Fig. 2), and that surface EMG has the largest median delay.

4. Discussion

Automatic seizure documentation for outpatients has to proof high sensitivity and needs post-hoc manual evaluation for reliable seizure identification. Low false detection rates are mandatory to reduce workload of manual evaluation procedure. We present a multimodal seizure detection algorithm working in real time that is able to detect epileptic seizures with high sensitivity using EEG, EMG, and ECG signals.

Our results show very high detection sensitivity of 94% for TLE and overall detection sensitivity of 86% using the 22 electrode montage (21 EEG electrodes plus one ECG). Furthermore, the algorithm was able to detect all focal seizures evolving to bilateral tonic-clonic ($n = 50$). Therefore, our automatic seizure detection system potentially increases sensitivity of seizure documentation compared to manual procedures in all patient and seizure groups.

Reduced electrode montages for automatic seizure documentation assures patient compliance in long-term outpatient settings. E.g. omitting posterior electrodes will increase sleep comfort and therefore positively influences EEG quality of nocturnal events. Furthermore, setup time of ambulatory EEG is a major cost factor besides data evaluation which is reduced by a factor of three when using 7 EEG electrodes only. We found lower sensitivities compared to the full 10–20 electrode montage (forehead −5% SE, posterior −12% SE). Based on our results we conclude that the 8 electrode forehead montage is most beneficial for this application. Similar montages showed high sensitivities for emergency and prehospital care application (Jakab et al., 2014). Also our previous work using EEG from intensive care unit patients showed promising results of automatic pattern detection based on forehead EEG montages (Herta et al., 2017).

False detection rates of reduced electrode montages dropped down only by −3 FD/24 h for forehead but by −10 FD/24 h for posterior montages showing a positive correlation between the number of electrodes and false detection rate that is more pronounced for posterior electrodes.

Results therefore encourage the use of reduced electrode sets based on frontal and temporal electrodes for long-term seizure documentation. Even XTLE patients showing the lowest sensitivity in our study (68% SE), true seizures counts can be significantly improved as compared to manual seizure counting sensitivity of 50% (Blachut et al., 2015).

Reduction of EEG electrodes will negatively influence visual inspection and seizure validation. It is important to limit electrode reduction to maintain interpretability of the EEG. We therefore avoid electrode reduction below 6 EEG electrodes or sole use of non-EEG signals which would not allow seizure validation at all.

We found significantly lower sensitivities for XTLE patients as compared to TLE patients which can be explained by differences in ictal EEG patterns. Visual analysis of false negatives in XTLE showed that these seizures include low amplitude beta, gamma activity, or high amplitude muscle artefacts but only marginal rhythmic activity. The high rate of interictal abnormal EEG activity in these patients is the reason for the high false detection rate of 22 FD/24 h in this evaluation group.

In our study absolute values of ECG based seizure detections were low (TLE: 40%, XTLE: 8%, FS: 27%, BTCS: 43%). These results are in good agreement with previous publications (Eggleston et al., 2014). Discrepancies as compared to other studies (Leutmezer et al., 2003) can be explained by differences in the definition of ictal tachycardia (Eggleston et al., 2014).

Added value of ECG based seizure detection is marginal when full 10–20 EEG is available. Data shows that sensitivity increases by only a few percent or not at all when adding ECG based detections to 21 electrode EEG based detections (TLE: +2%, XTLE: 0%, FS: +1%, SGTC: 0%). Similar results were found for detections based on 8 electrode forehead montage. ECG based detections gain importance only for the less sensitive 7 electrode posterior montage (ALL: +5%, TLE: +7%, XTLE: +1%, FS: +5%, GTCS: 0%). This shows that low sensitive EEG setups can partly recover sensitivity by using other signal modalities like ECG.

EMG signals were extracted from EEG data via a bandpass filter. An important point of this work was to reduce effort of the electrophysiological setup. Furthermore, dedicated EMG signals were not available in the data of this work. Detection sensitivity for focal seizures evolving to bilateral tonic-clonic solely using derived EMG signals was very high (93%) and an additional surface EMG sensor is therefore avoidable. A further advantage of EEG electrode based EMG detection is that recorded EEG data can be used to validate the EMG based seizure alarms which is impossible with accelerometer data alone.

Accelerometer sensors are able to detect clonic or tonic-clonic seizures with high sensitivity. Detection performance of 66% sensitivity and 1.1 false detections per night was reached (Van de Vel et al., 2016). Combination of accelerometer data and electrodermal activity (EDA) reached 89% sensitivity and 93% specificity on data of 8 patients (Heldberg et al., 2015). In this work EMG based seizure detections reached 93% for SGTC and showed a shorter detection delay than accelerometer based detectors.

Comparing detection performance of the presented multimodal seizure detection algorithm to the online seizure detection method EpiScan (Fürbass et al., 2015a) shows that a higher sensitivity for TLE (94–83%) and XTLE patients (74–64%) but also a higher false detection rate (TLE 12.8–6.7, XTLE: 22.2–7.3 FD/24 h) can be reached. On the other hand comparing results of our study to results of a study published by Hopfengärtner et al. (2014) shows higher sensitivity for TLE (94–89%) but lower sensitivity for XTLE (74–77%). False alarm rate reported in (Hopfengärtner et al., 2014) is lower compared to results of our study (12.8–4.5 FD/24 h). The higher sensitivity of the multimodal seizure detection algorithm and low CPU calculation time fit the use case of semi-automatic seizure documentation. In turn, the very low false alarm rate of EpiScan and similar algorithms is well suited for triggering acoustic alarms for patient surveillance.

In this work we presented an automatic seizure detection algorithm and results of retrospective data analysis. We are fully aware that a complete seizure documentation infrastructure has to include wearable electrode systems for monitoring, storing detected seizure periods, and IT infrastructure as well as software for transmitting and reviewing stored seizure data for final validation.

5. Conclusion

We presented an automatic multimodal seizure detection algorithm for long-term seizure documentation. Evaluation of detection performance on 92 long-term EEG/ECG/EMG recordings from two epilepsy monitoring units including 11,978 h of data and 494 seizures resulted in high detection sensitivity. The effect of different signal modalities on detection performance and detection

delay was analyzed in detail. The effect of reduced electrode montages on detection performance showed the superiority of frontal and temporal EEG electrodes for automatic seizure detection. The work showed that improved long-term seizure documentation is possible using automatic seizure detection algorithms based on only 8 frontal and temporal as well as one ECG electrode. We conclude that using semi-automatic seizure documentation will improve seizure documentation in general and justifies the additional electrophysiological effort.

Acknowledgements

This work was partly done within the EEG-Serv project which received funding from the EUROSTARS programme grant agreement No. E! 10145. Many thanks for the support of our work to all people involved and especially to Dr. Katharina Kottmel and Dr. Lejla Elezi.

Conflict of interest: The authors declare that they have no conflicts of interest concerning this article.

References

Banks SJ, Bellerose J, Douglas D, Jones-Gotman M. The insular cortex: relationship to skin conductance responses to facial expression of emotion in temporal lobe epilepsy. Appl Psychophysiol Biofeedback 2014;39:1–8. http://dx.doi.org/10.1007/s10484-013-9236-3.

Beniczky S, Conradsen I, Pressler R, Wolf P. Quantitative analysis of surface electromyography: Biomarkers for convulsive seizures. Clin Neurophysiol Off J Int Fed Clin Neurophysiol 2016;127:2900–7. http://dx.doi.org/10.1016/j.clinph.2016.04.017.

Blachut B, Hoppe C, Surges R, Stahl J, Elger CE, Helmstaedter C. Counting seizures: the primary outcome measure in epileptology from the patients' perspective. Seizure 2015;29:97–103. http://dx.doi.org/10.1016/j.seizure.2015.03.004.

Conradsen I, Beniczky S, Wolf P, Kjaer TW, Sams T, Sorensen HBD. Automatic multimodal intelligent seizure acquisition (MISA) system for detection of motor seizures from electromyographic data and motion data. Comput Methods Programs Biomed 2012;107:97–110. http://dx.doi.org/10.1016/j.cmpb.2011.06.005.

Eggleston KS, Olin BD, Fisher RS. Ictal tachycardia: the head-heart connection. Seizure 2014;23:496–505. http://dx.doi.org/10.1016/j.seizure.2014.02.012.

Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. Clin Neurophysiol 2015a;126:1124–31. http://dx.doi.org/10.1016/j.clinph.2014.09.023.

Fürbass F, Hartmann MM, Halford JJ, Koren J, Herta J, Gruber A, et al. Automatic detection of rhythmic and periodic patterns in critical care EEG based on American Clinical Neurophysiology Society (ACNS) standardized terminology. Neurophysiol Clin Neurophysiol 2015b;45:203–13. http://dx.doi.org/10.1016/j.neucli.2015.08.001.

Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: automatic EEG artifact removal for epilepsy monitoring. Neurophysiol Clin Clin Neurophysiol 2014;44:479–90. http://dx.doi.org/10.1016/j.neucli.2014.09.001.

Heldberg BE, Kautz T, Leutheuser H, Hopfengartner R, Kasper BS, Eskofier BM. Using wearable sensors for semiology-independent seizure detection – towards ambulatory monitoring of epilepsy. In: Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf 2015. p. 5593–6. http://dx.doi.org/10.1109/EMBC.2015.7319660.

Herta J, Koren J, Fürbass F, Hartmann M, Gruber A, Baumgartner C. Reduced electrode arrays for the automated detection of rhythmic and periodic patterns in the intensive care unit: frequently tried, frequently failed? Clin Neurophysiol 2017. http://dx.doi.org/10.1016/j.clinph.2017.04.012.

Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. Epilepsy Behav 2015. http://dx.doi.org/10.1016/j.yebeh.2015.04.064.

Hopfengärtner R, Kasper BS, Graf W, Gollwitzer S, Kreiselmeyer G, Stefan H, et al. Automatic seizure detection in long-term scalp EEG using an adaptive thresholding technique: a validation study for clinical routine. Clin Neurophysiol Off J Int Fed Clin Neurophysiol 2014;125:1346–52. http://dx.doi.org/10.1016/j.clinph.2013.12.104.

Jakab A, Kulkas A, Salpavaara T, Kauppinen P, Verho J, Heikkilä H, et al. Novel wireless electroencephalography system with a minimal preparation time for use in emergencies and prehospital care. Biomed Eng Online 2014;13:60. http://dx.doi.org/10.1186/1475-925X-13-60.

Jeppesen J, Beniczky S, Johansen P, Sidenius P, Fuglsang-Frederiksen A. Using Lorenz plot and Cardiac Sympathetic Index of heart rate variability for detecting seizures for patients with epilepsy. Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf 2014;2014:4563–6. http://dx.doi.org/10.1109/EMBC.2014.6944639.

Leutmezer F, Schernthaner C, Lurger S, Pötzelberger K, Baumgartner C. Electrocardiographic changes at the onset of epileptic seizures. Epilepsia 2003;44:348–54.

Stefanidou M, Carlson C, Friedman D. The relationship between seizure onset zone and ictal tachycardia: an intracranial EEG study. Clin Neurophysiol 2015;126:2255–60. http://dx.doi.org/10.1016/j.clinph.2015.01.020.

Van de Vel A, Milosevic M, Bonroy B, Cuppens K, Lagae L, Vanrumste B, et al. Long-term accelerometry-triggered video monitoring and detection of tonic-clonic and clonic seizures in a home environment: Pilot study. Epilepsy Behav Case Rep 2016;5:66–71. http://dx.doi.org/10.1016/j.ebcr.2016.03.005.

6.6 Paper A6: NeuroTrend as bedside monitor

Title: Applicability of NeuroTrend as a bedside monitor in the neuro ICU

Authors: J. Herta, J. Koren, F. Fürbass, A. Zöchmeister, M. Hartmann, A. Hosmann, C. Baumgartner, A. Gruber

Published in: Clinical neurophysiology

Year: 2017

Authors' contribution: Fürbass Franz developed the computer algorithms that were used in this study. Franz Fürbass worked on corrections and assisted in evaluation of study data. Clinical work, study design, collection of study data and setup, as well as initial manuscript editing was done by first and other coauthors.

Clinical Neurophysiology 128 (2017) 1000–1007

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Applicability of NeuroTrend as a bedside monitor in the neuro ICU

J. Herta [a,*], J. Koren [b], F. Fürbass [c], A. Zöchmeister [a], M. Hartmann [c], A. Hosmann [a], C. Baumgartner [b,d], A. Gruber [a]

[a] Department of Neurosurgery, Medical University of Vienna, Vienna, Austria
[b] Karl Landsteiner Institute for Clinical Epilepsy Research and Cognitive Neurology, 2nd Neurological Department, General Hospital Hietzing with Neurological Center Rosenhuegel, Vienna, Austria
[c] AIT Austrian Institute of Technology GmbH, Digital Safety & Security Department, Vienna, Austria
[d] Department of Epileptology and Clinical Neurophysiology, Sigmund Freud University, Vienna, Austria

ARTICLE INFO

Article history:
Accepted 2 April 2017
Available online 11 April 2017

Keywords:
Epileptic seizure detection
Nursing
Interrater agreement
Continuous EEG
Intensive care unit
Screening device
Monitoring
Periodic discharge
Rhythmic and periodic patterns

HIGHLIGHTS

- Proposal and guidance on how a computer algorithm may be used by ICU staff as a cEEG bedside monitor.
- High interrater agreement among nurses for EEG patterns that may indicate subclinical seizures.
- Large amount of prospectively recorded, randomized long-term video EEG data from two neuro ICUs.

ABSTRACT

Objective: To assess whether ICU caregivers can correctly read and interpret continuous EEG (cEEG) data displayed with the computer algorithm NeuroTrend (NT) with the main attention on seizure detection and determination of sedation depth.
Methods: 120 screenshots of NT (480 h of cEEG) were rated by 18 briefly trained nurses and biomedical analysts. Multirater agreements (MRA) as well as interrater agreements (IRA) compared to an expert opinion (EXO) were calculated for items such as pattern type, pattern location, interruption of recording, seizure suspicion, consistency of frequency, seizure tendency and level of sedation.
Results: MRA as well as IRA were almost perfect (80–100%) for interruption of recording, spike-and-waves, rhythmic delta activity and burst suppression. A substantial agreement (60–80%) was found for electrographic seizure patterns, periodic discharges and seizure suspicion. Except for pattern localization (70.83–92.26%), items requiring a precondition and especially those who needed interpretation like consistency of frequency (47.47–79.15%) or level of sedation (41.10%) showed lower agreements.
Conclusions: The present study demonstrates that NT might be a useful bedside monitor in cases of subclinical seizures. Determination of correct sedation depth by ICU caregivers requires a more detailed training.
Significance: Computer algorithms may reduce the workload of cEEG analysis in ICU patients.

1. Introduction

Continuous electroencephalography (cEEG) is used in the intensive care unit (ICU) to detect subclinical seizures and to monitor sedation depth in cases of refractory seizures or elevated intracranial pressure (Eisenberg et al., 1988; Friedman et al., 2009; Sutter et al., 2013). Previous studies showed that subclinical seizures occur more often than anticipated in the ICU (Kaplan, 1999) and frequently develop at an early stage of acute brain injury (Claassen et al., 2004). Since mortality increases exponentially with seizure duration in critical care patients, proper application and instant interpretation of cEEG is crucial in this setting (Young

Abbreviations: AC1, Gwet's multirater agreement coefficient of first-order; AIT, Austrian Institute of Technology; aEEG, amplitude integrated electroencephalography; BMA, biomedical analyst; BS, burst suppression; CCET, American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology; cEEG, continuous electroencephalography; ESP, electrographic seizure pattern; EXO, expert opinion; FIRDA, frontal intermittent rhythmic delta activity; GCS, Glasgow coma scale; ICU, intensive care unit; IRA, interrater agreement; MRA, multirater agreement; NT, NeuroTrend; PD, periodic discharge; qEEG, quantitative electroencephalography; RAA, rhythmic activity in the alpha range, "rhythmic alpha activity"; RDA, rhythmic delta activity; RDA+S, rhythmic delta activity plus seizures; RTA, rhythmic activity in the alpha range, "rhythmic theta activity"; SIRPIDs, stimulus-induced rhythmic, periodic, or ictal discharges; SW, spike wave.

* Corresponding author at: Department of Neurosurgery, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria. Fax: +43 01 40400 45660.
E-mail address: johannes.herta@meduniwien.ac.at (J. Herta).

http://dx.doi.org/10.1016/j.clinph.2017.04.002
1388-2457/

et al., 1996; Vespa et al., 1999). Its use has been associated with a favorable outcome in the critically ill (Ney et al., 2013). But cEEG monitoring is not available in the majority of hospitals as it requires a lot of resources.

To minimize the diagnostic effort of visually screening hours of cEEG the Austrian Institute of Technology (AIT) has developed a computer algorithm called NeuroTrend (NT) with a strong ability to visualize rhythmic and periodic patterns in a time compressed fashion (Fürbass et al., 2015, 2016). A possible field of application lies in the use of NT as a bedside monitor. However, in contrast to other ICU monitors, an automatic alarm system for seizures would be ineffective, as false alarms would be too frequent in an environment that contains a plurality of possible EEG artefacts. In addition, NT data needs interpretation as it also displays trend data of patterns that are not clearly ictal. Therefore, trained nurses, taking care of the same patient over several hours would be best suited to use and interpret the computer results.

The present study investigated whether briefly trained ICU caregivers can read and interpret NT cEEG data correctly. To test this hypothesis, 15 ICU nurses and 3 biomedical analysts (BMA) not familiar with EEG, had to evaluate NT cEEG data from patients with acute brain injury. The evaluations were then compared between the respondents as well as with an expert opinion (EXO) and tested for their consistency. The main parameters tested for consistency were: (1) Identification of seizures occurrence and seizure progression (2) Assessment of the sedation depth.

2. Methods

2.1. Dataset

A dataset of 83 prospectively recorded continuous video-EEGs (6733 h, mean 73 h) from a neurological (Neurological Center Rosenhuegel) and a neurosurgical ICU (General Hospital Vienna) was used. All recordings were obtained from patients older than 18 years with a median age of 58.5 years. EEGs were recorded using a Micromed EEG system (SystemPLUS Evolution 1.04.95, Micromed S.p.A., Veneto, Italy) with a sampling rate of 256 Hz, placing 21 electrodes according to the international 10–20 system. Only video-EEGs with a duration of more than 24 h and a sufficient EEG signal quality over the whole recording period were used in this study. Patients were selected using the NeuroTrend (NT) Analysis Database. This database was established in 2011 with its main focus of investigating rhythmic and periodic EEG patterns of 'ictal-interictal uncertainty' as well as subclinical seizures and status epilepticus (Koren et al., 2015). All cEEGs registered in the database were reviewed by board certified neurophysiologists and screened for electrographic seizure patterns (ESP), spike wave (SW), rhythmic delta activity (RDA), periodic discharges (PD), burst suppression (BS) patterns and patterns mimicking artefacts as described elsewhere (Herta et al., 2015). EEG changes in frequency, prevalence, localization and morphology were reevaluated every 24 h according to the guidelines of the American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology (CCET) (Hirsch et al., 2013). Additional information included treatment protocols, patient characteristics, certain neurologic scores and follow up data (Glasgow Outcome Score after six month). From this database 20 patients were randomly selected with a predefined split into the following six groups: PD ($n = 3$), ESP ($n = 3$), SW ($n = 3$), RDA ($n = 3$), BS ($n = 3$) and none of the above-mentioned patterns ($n = 5$). If a patient showed more than one pattern type he or she could be randomized into multiple groups but overall could not be selected more than once.

2.2. NeuroTrend

Included cEEGs were analyzed by the computer algorithm NT. This algorithm detects and visualizes rhythmic and periodic EEG patterns with a strong emphasis on data and time compression as well as artefact rejection (Hartmann et al., 2014; Fürbass et al., 2015, 2016). A color code displays the following patterns: periodic discharges (PD), rhythmic delta activity (RDA), rhythmic delta activity plus superimposed sharp waves or spikes (RDA + S), rhythmic activity in the theta range (RTA), rhythmic activity in the alpha range (RAA) and spike wave (SW). Pattern localization, pattern frequencies, frequency bands (beta, alpha, theta and delta range) and amplitude integrated EEG (aEEG) are calculated and displayed on a graphical user interface. The definition of rhythmic and periodic EEG patterns follows the guidelines of the CCET adding unequivocal electrographic seizures including generalized spike-wave discharges at 3 Hz or faster, evolving discharges that reach frequencies of more than 4 Hz as well as BS patterns (Hirsch et al., 2013). A validation of NT was recently published elsewhere (Herta et al., 2015). NT is part of the encevis software package. Version V1.3 of encevis was used in this study (http://www.encevis.com).

2.3. NT training and data preparation

Nurses and BMAs (in the following referred to as "respondents") from a neurosurgical ICU were asked to volunteer for the study. A total of 18 respondents, including 3 BMAs and 15 nurses, then compiled a questionnaire where they were asked about work experience, experience with EEG, computer skills, experience in playing computer games and presence or absence of color blindness. All personal data were anonymized for further analysis. All respondents underwent a brief educational course of approximately one hour. Features of NTs graphical user interface and the study design were explained. Samples of cEEG and NT data were presented. The presentation files as well as a short rating manual were handed out to the respondents (available as Supplementary material).

Shortly thereafter each respondent was asked to rate the preselected NT data. The rating manual could be used during the evaluation process. During the assessment, the time needed to evaluate the NT data of 20 patients was measured. NT data were presented to the respondents by an editable Microsoft® PowerPoint slideshow (Fig. 1). A brief patient history was given, including information about admission diagnosis, operative procedures undertaken and their time course, seizures prior to EEG, anesthetics and antiepileptic drugs administered, clinical features that may indicate subclinical seizures (Husain et al., 2003) and Glasgow Coma Scale (GCS) at cEEG start. Subsequently, results of NT analysis were presented in a mask that allowed simultaneous ratings of each slide. Slides displayed 6 consecutive NT screenshots for each patient with a length of 4-h each, giving in total 120 screenshots or 480 h of cEEG.

2.4. NeuroTrend review scheme

Rating possibilities were grouped into 4 categories and could be selected by check boxes. In category one the patterns recognized by NT (PD, RDA, RDA + S, RTA, RAA, SW) had to be identified (Fig. 1a). The selection of multiple patterns for each 4-h segment was possible. For each pattern the principal location had to be defined. If the pattern was not clearly localized to the left or right hemisphere, generalized had to be selected. Furthermore, for each pattern the consistency of frequency had to be indicated (Fig. 1b). A consistent frequency was assumed if the frequency remained the same or increased/decreased continuously over a longer recording period of at least 30 min.

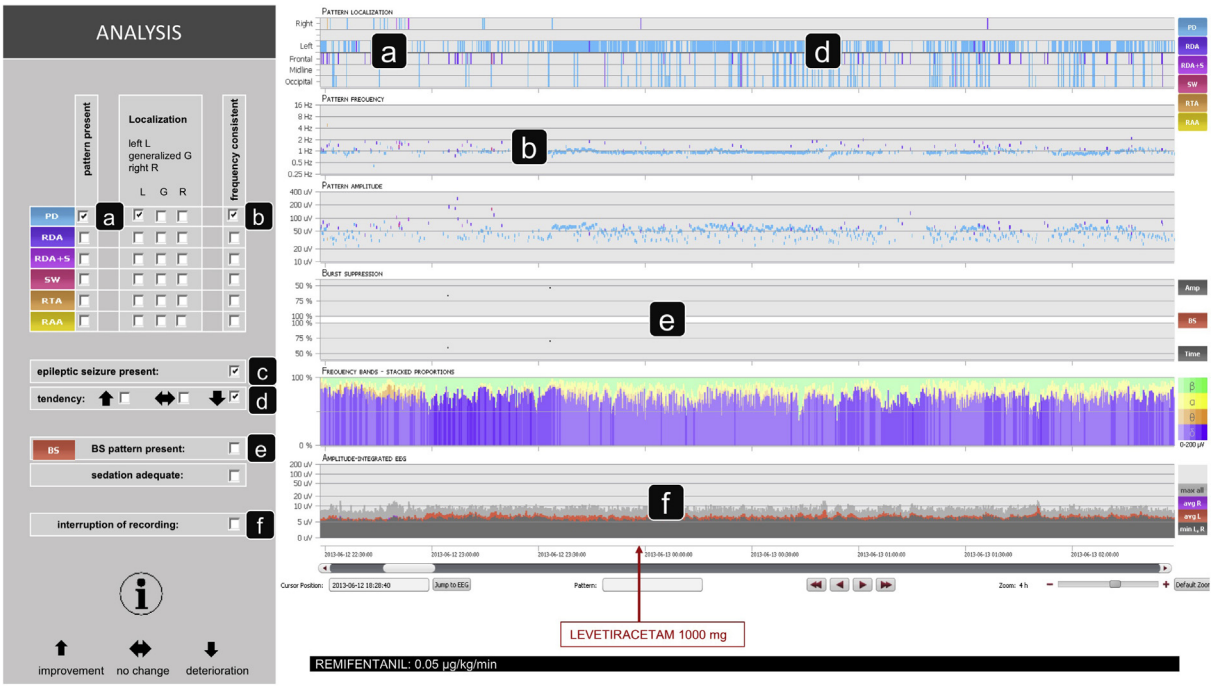

Fig. 1. Example of a rating slide. The slide is divided into two parts. On the right side a 4-h sample of NeuroTrend (NT) with additional information concerning anesthetics and antiepileptic drugs is presented to the respondent. On the left side a rating matrix which is divided into four categories was implemented (white boxes). The information sign at the left bottom gives the respondent a short summary about the medical record of the patient. In this case a comatose patient suffered from subclinical seizures due to a large left parietal metastasis of a lung carcinoma. EEG was used during the weaning process after resection of the metastasis. Even after resection the NT screenshot displays left sided periodic discharges (PD) (a) with a stable frequency (b) which was interpreted as ongoing seizure activity (c). Because of an increase in pattern occurrence (PD) the tendency was rated as "seizure deterioration" (d). No burst suppression (BS) occurred during the recording period (e). EEG was continuously recorded as there is no discontinuity in the aEEG (f).

In the second category respondents had to check if they suspected a seizure (Fig. 1c). Seizures were defined by presence of SW, RTA and/or RAA. More equivocal patterns like PD, RDA and RDA + S had to be rated as possible seizures if these patterns occurred continuously over a long-time period with a consistent frequency. This definition was established in order not to oversee the third non-convulsive seizure criterion introduced by Chong et al. which includes sequential rhythmic, periodic, or quasi-periodic waves at 1/sec with an unequivocal evolution in frequency, morphology or location (Chong and Hirsch, 2005). If seizures were present the tendency (unchanged, improvement, deterioration) between the current and the consecutive slide of a patient had to be specified (Fig. 1d). Improvement was defined by absence of seizures or a decline in seizure frequency. Deterioration was defined by occurrence of new seizures or increase of seizure frequency.

BS patterns could be selected in category three (Fig. 1e). In case of BS it had to be specified if anesthesia depth was adequate. Anesthesia depth was defined as adequate if: (1) the BS pattern occurred with a nearly simultaneous change in the anesthetic regimen after seizures were treated. (2) If the BS pattern were present while the patient suffered from elevated ICP. Administered anesthetics were displayed under the timeline of the NeuroTrend interface. Elevated ICP was mentioned in the patient history.

Finally, in category four, respondents had to choose if an interruption of the recording had occurred, best seen in an abrupt discontinuation of the aEEG (Fig. 1f).

The same editable slideshow was evaluated by a board-certified neurophysiologist who was familiar with NT. Not only NT results but also the raw EEG and chart reviews were accessible to the neurophysiologist during the rating process to establish an "expert opinion" (EXO).

For further analysis, the NT detections of type RDA and RDA + S were merged into RDA and RTA and RAA into ESP. There was no reason to differentiate these patterns for the present study.

2.5. Statistical evaluation

For statistical evaluation, an interrater agreement (IRA) was calculated. Seven independent, as well as twenty dependent items, were assessed by the IRA. Independent items included occurrence of specific patterns (SW, PD, RDA, ESP, BS), seizure suspicion and interruption of recording. Dependent items included seizure tendency, pattern localization and adequacy of BS. The terms "dependent items" and "independent items" refer to the fact that dependent items serve to further describe and sub-classify independent items. For example, if we choose a specific pattern like "SW" as the independent item, it can be further sub-classified by the dependent item "pattern localization" into "lateralized" or "generalized".

To categorize the IRA, ranges of agreement have been defined as 0.01–0.20 for a slight agreement, 0.20–0.40 for a fair agreement, 0.40–0.60 for a moderate agreement, 0.60–0.80 for a substantial agreement and 0.80–1.00 for a perfect agreement (Blood and Spratt, 2007). To quantify the rating agreement of the respondents Gwet's multirater agreement coefficient of first-order (MRA AC1) was used (Gwet, 2014). The same method was used to measure the rating agreement in each category between the respondents with the EXO (IRA AC1). All IRA AC1 were then averaged over the respondents to allow further comparisons that consider the whole study population. The same approach was used for dependent items if the precondition (independent item) of the respondent met the precondition of the EXO. Because of a given EXO, sensitivity and specificity of all rating items were calculated for every

respondent. The respondents' answer was counted as a true positive or a true negative if there was an agreement with the EXO. Furthermore, for each respondent individual performance scores were calculated. The individual performance score was composed by the averaged IRA between a respondent and the EXO of all independent items excluding the item "disruption of recording". Only independent items were used for this analysis because the number of ratings was equal for all respondents. The individual performance score was used to determine differences and correlations in the rating behavior of the respondents. Sex, profession, preexisting experience with EEG and computer gaming experience were evaluated by an independent two-sided t-test. Computer skills were assessed by a one-way analysis of variance. Pearson correlation coefficient was used for age and work experience (in years). Significance levels of 0.05 were applied for all statistical tests.

The two non-binary items "seizure tendency" and "pattern localization" were displayed using confusion matrices. For each item the percentage of agreement (true positive and true negative ratings to all ratings) for every available category was presented as a heat map and compared with the EXO ratings.

Statistical calculations were performed by using MATLAB (The MathWorks, Natick, MA, U.S.A.) and its Statistics toolbox as well as IBM SPSS Statistics V23.

2.6. Ethics approval and consent

The study protocol was approved by the institutional ethics commission. Informed consent was given by all nurses and BMAs that volunteered for the study. Patients included in the NT database were mainly not able to give consent during EEG recordings. Therefore, the ethics commission requested that all patients that were not able to give consent and their relatives receive a written patient information and/or were informed about the study and the possibility to withdraw their personal data in the future.

3. Results

The mean time for the evaluation of the 120 screenshots was 1 h and 22 min. In Table 1 the IRA of dependent as well as independent items is shown. We observed the following multirater agreement coefficients (MRA AC1) for independent items: interruption of recording (92.72%), spike wave (91.74%, SW), rhythmic delta activity (85.53%, RDA) and burst suppression (80.39%, BS) showed perfect agreements, while electrographic seizure patterns (69.91%, ESP), periodic discharge (66.82%, PD) and seizure suspicion (60.9%) showed substantial agreements. All independent items also were compared to the "expert opinion" (EXO) by averaging the interrater agreement (IRA) of the respondents. As expected, slightly higher IRAs could be achieved with a perfect agreement for interruption of recording (93.97%), SW (92.03%), RDA (90.25%) and BS (85.63%). Again, a substantial agreement was observed for ESP (77.96%), PD (74.93) and seizure suspicion (67.45%). Fig. 2 illustrates the differences between the multirater agreement (MRA) and the averaged IRA of the respondents compared to the EXO. We observed that independent items with a high MRA like SW or interruption of recordings do not differ as much from the EXO as items with a lower MRA like seizure suspicion. Similar effects were obtained in the receiver operating characteristic of all independent items (Fig. 3A). Seizure suspicion (79.10%), ESP (88.69%) and PD (87.98%) achieved lower specificities under 90% in comparison with SW (95.33%), BS (92.44%); RDA (93.88%) and interruption of seizures (99.04). Except PD (79.25%) all items obtain sensitivities over 80% with RDA (93.18%) and BS (92.82%) exceeding sensitivities of 90%.

In general, dependent items showed lower agreements as independent items with a perfect agreement for the localization of ESP (92.26%) and substantial agreements for the localization of PD (75.68%), the localization of SW (71.81%), the localization of RDA (70.83%) and seizure tendency (61.40%) (Table 1). Assessment of frequency consistency was highly dependent on the evaluated pattern type and showed agreements between 47.47% and 79.15%. The question whether the level of sedation was adequate in the presence of BS patterns achieved only a barely moderate agreement (41.10%). Dependent items with more than two choices are displayed as confusion matrices in Fig. 4. We obtained an acceptable result with a diagonal line of agreement for seizure tendency (Fig. 4B). All patterns were analyzed for pattern localization (Fig. 4A). We observed that generalized patterns were distinguished clearly from lateralized patterns but in many cases respondents could not assign the correct side for lateralized patterns.

Fig. 5 gives a detailed overview of all assessed items. Agreements of every single respondent compared to the EXO are displayed as a heat map. For every respondent, the individual performance score is displayed next to the respondent rank. Accordingly, Fig. 3B depicts how sensitive and specific the respondents' ratings were.

The respondent characteristics including sex, age, profession, gaming experience, EEG experience, work experience or computer

Table 1
Interrater and multirater agreement for tested items.

Item	Dependent item	Choices	Multirater agreement, AC1 (%)	Averaged IRA to EXO, AC1 (%)
Periodic discharge		Yes, no	66.82	74.93
	Localization	L,G,R		75.68
	Consistency of frequency	Yes, no		47.47
Rhythmic delta activity		Yes, no	85.53	90.25
	Localization	L,G,R		70.38
	Consistency of frequency	Yes, no		79.15
Spike wave		Yes, no	91.74	92.03
	Localization	L,G,R		71.81
	Consistency of frequency	Yes, no		64.53
Electrographic seizure pattern		Yes, no	69.91	77.96
	Localization	L,G,R		92.26
	Consistency of frequency	Yes, no		79.63
Burst suppression		Yes, no	80.39	85.63
	Level of sedation	Adequate, inadequate		41.10
Seizures suspicion		Yes, no	60.90	67.45
	Seizure tendency	T+, T=, T−		61.40
Interruption of recording		Yes, no	92.72	93.97

AC1, agreement coefficient of first order; EXO, expert opinion; IRA, interrater agreement; L, left; R, right; G, generalized; NA, not available; T+, improvement; T=, unchanged; T−, deterioration.

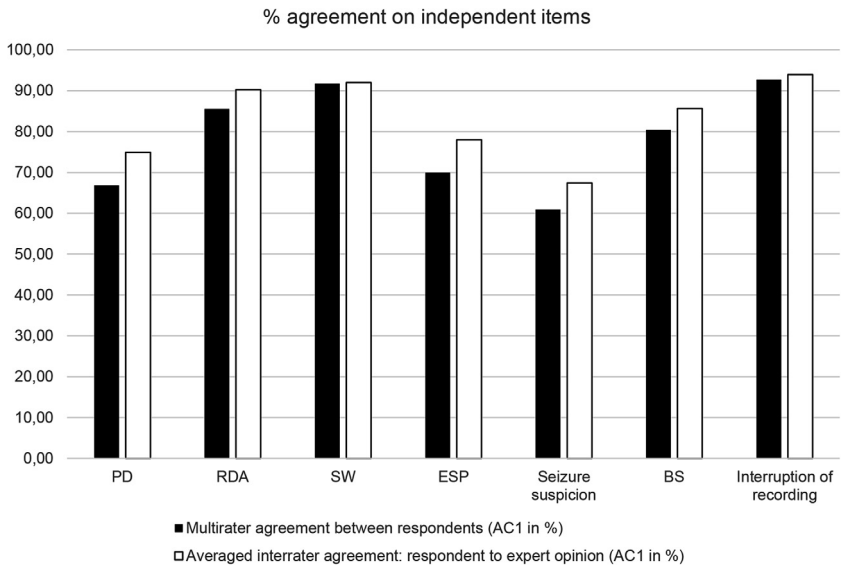

Fig. 2. Percentage of agreement (% agreement) on independent items. The multirater agreement coefficient of first order (MRA AC1; black bars) gives the agreement between the respondents. The interrater agreement coefficient of first order (IRA AC1; white bars) presents the averaged agreement between the respondents and the expert opinion. BS, burst suppression; PD, periodic discharge; RDA, rhythmic delta activity; ESP, electrographic seizure pattern; SW, spike wave.

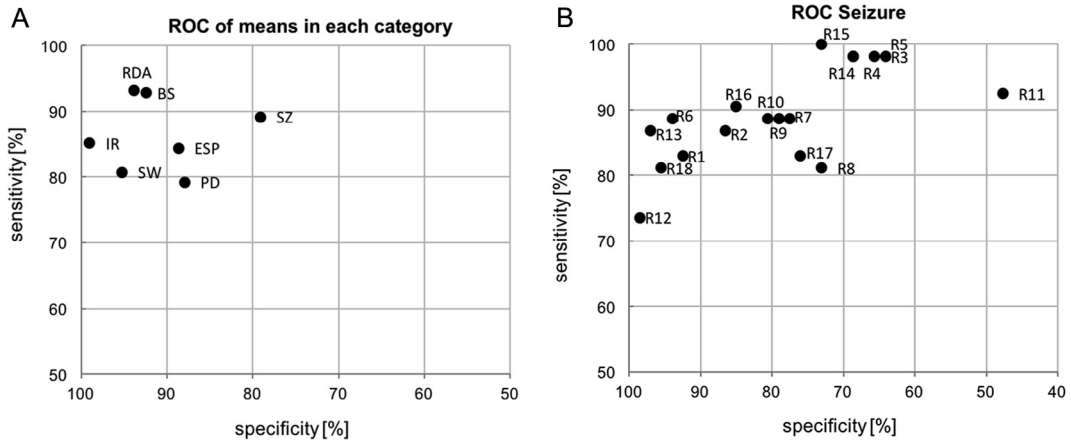

Fig. 3. Receiver operating characteristic. (A) shows the mean sensitivity and mean specificity for seven independent items calculated from the answers of 18 respondents. (B) illustrates the sensitivity and specificity for seizures of every respondent (R). To calculate sensitivity and specificity answers of the respondents were compared to the expert opinion. BS, burst suppression; PD, periodic discharge; RDA, rhythmic delta activity; ESP, electrographic seizure pattern; SW, spike wave; IR, interruption of recording; SZ, seizure suspicion; R, respondent.

skills had no significant influence on the individual performance score of each respondent. A detailed overview of characteristics, statistical tests used and results is given in Table 2. No respondent suffered from color blindness.

4. Discussion

The applicability of NeuroTrend (NT) as a continuous EEG (cEEG) bedside monitor was evaluated by multirater agreement (MRA) and interrater agreement (IRA). Perfect agreement was found for spike waves (SW), rhythmic delta activity (RDA) and burst suppression (BS) while electrographic seizure patterns (ESP), periodic discharges (PD) and seizure suspicion achieved substantial agreement among the respondents. Similar agreements were found when we compared the choices of the respondents with the expert opinion (EXO). Furthermore, sensitivity and specificity for all independent items showed results over 80% with two exceptions: sensitivity for PD with 79.25% and the specificity for seizure suspicion with 79.10%. These high agreements, high sensitivities and high specificities achieved for independent items indicate that with NT, briefly trained ICU personal can identify the occurrence of rhythmic and periodic EEG patterns that may indicate seizures. Also, over or under dosage of sedative drugs may be identified adequately.

NT displays trend data and not every short-lasting detection should be interpreted as an event. This provides an opportunity for interpretation and raises the question why some patterns showed higher agreements than others. SW and RDA had the highest agreements because they were mostly present over longer time periods and were therefore easy to detect, if displayed by the algorithm. The same is true for BS patterns where in some cases interpretation was facilitated if large amounts of anesthetic drugs were administered at the same time the pattern occurred. One possible source of error could have been that anesthetic drugs given at short notice for nursing care caused short periods of BS. Respondents knew only about the continuous application of intravenous anesthetics and were blinded to these short applications. This might have caused disagreement during the rating process. PD were the most difficult to evaluate pattern. PD occurred frequently with lots

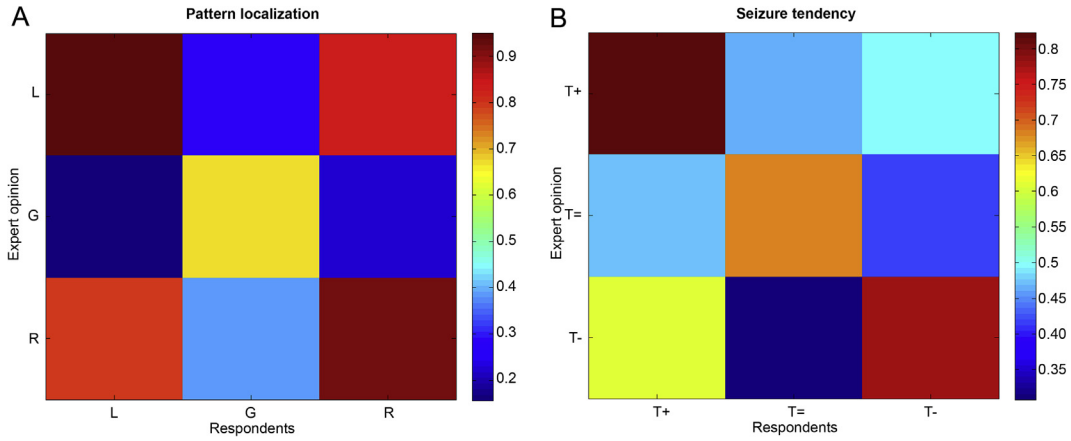

Fig. 4. Color coded confusion matrices for dependent items with more than two choices. Choices are shown on the vertical and horizontal axes while heat map intensities indicate the percentage of respondents choosing an available option. Only annotations were used if the required dependent item was previously chosen correctly. A dark red diagonal line from upper left to lower right would indicate a perfect result. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

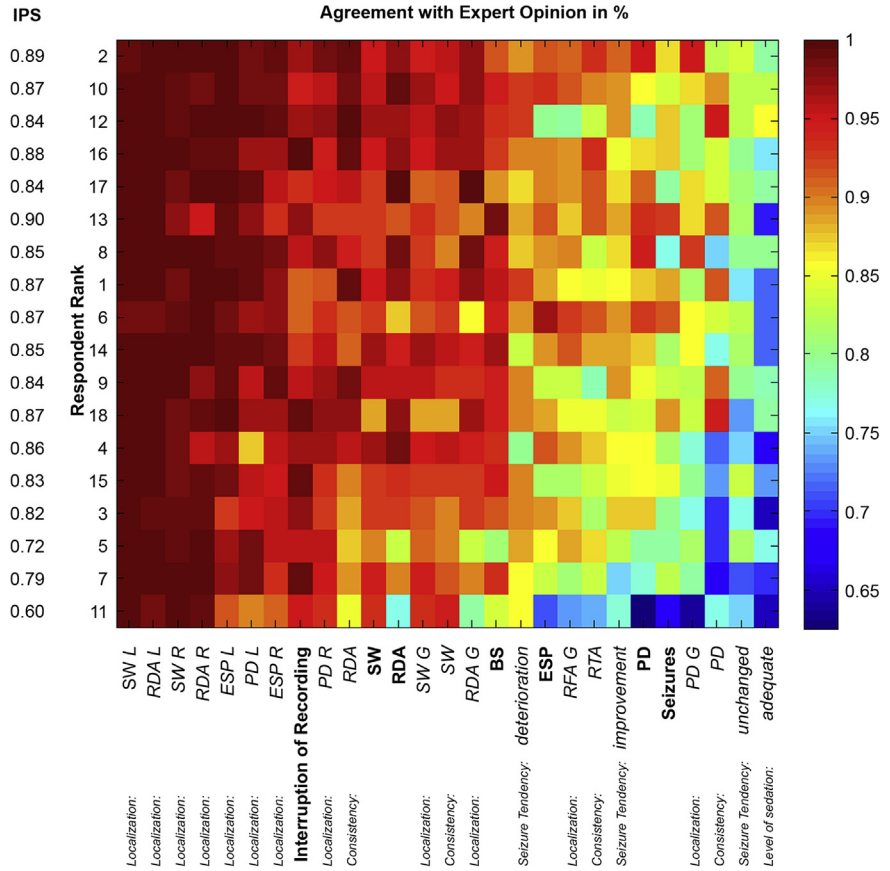

Fig. 5. The heatmap shows detailed individual scores of agreement coded by color according to the color bar on the right. For each respondent seven independent items (bold) and twenty dependent items (italic) have been tested. Scores for each item are calculated as percentage of answers in agreement between the respondents and the expert opinion. Additionally, on the left the individual performance score (IPS) of each respondent is displayed which takes all independent items except interruption of seizures into account. BS, burst suppression; PD, periodic discharge; RDA, rhythmic delta activity; ESP, electrographic seizure pattern; SW, spike wave; G, generalized; L, lateralized to the left; R, lateralized to the right. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of short-lasting detections. Only PD with a consistency in frequency over a period of 30 min should selected. This created a great scope for interpretation for example f single patterns appeared successively over longer periods of time. ESP on the contrary are patterns that are very rarely displayed incorrectly by NT. The difficulty is that these patterns, if present, mostly occurred in short periods of time and were therefore difficult to distinguish from misdetections resulting in an agreement of 77.96%.

Obviously, items that needed interpretation like seizure suspicion, seizure tendency, consistency of frequency and level of sedation showed lower MRAs/IRAs than items that had to be selected by mere presence, no matter if items were dependent or indepen-

Table 2
Influences of respondent characteristics on the individual performance score.

Respondent characteristics (n = 18)	Number of /mean	Statistical test	P value (α = 0.05)
Sex	Male: 5 Female: 13	Independent t-test of unequal variance	ns (0.141)
Age (in years)	$\bar{x}$ = 35.44	Pearson's correlation coefficient (0.033)	ns (0.909)
Profession	Nurse:15 BMA: 3	Independent t-test of unequal variance	ns (0.514)
Gaming experience	Available: 10 Not available: 8	Independent t-test of unequal variance	ns (0.828)
EEG experience	Available: 3 Not available: 15	Independent t-test of unequal variance	ns (0.514)
Work experience (in years)	$\bar{x}$ = 8.83	Pearson's correlation coefficient (−0.006)	ns (0.958)
Computer skills	Very good: 2 Good: 10 Satisfactory: 6 Sufficient: 0 Not sufficient: 0	One way ANOVA	ns (0.623)

ANOVA, analysis of variance; BMA, biomedical analyst; ns, not significant.

dent. Here we want to emphasize that only a brief presentation of the study protocol was chosen to train the respondents. This approach was not only chosen to recruit as many respondents as possible but also in order to test the feasibility and usability of our newly designed NT review scheme for nurses. We are convinced that better agreements are achievable with longer and repetitive trainings, especially for items that demand interpretation. Moreover, it has to be emphasized that the less well recognized dependent item "consistency of frequency" might have had a negative influence on the ratings of seizure suspicion and seizure tendency for patterns like PDs and RDA as they needed a consistent frequency to classify for seizures.

The strengths of our study are (1) the large amount of prospectively collected continuous EEG data, (2) a large number of respondents who rated the NT results and (3) a study design that is implementable in everyday clinical life. Implementation is not only possible from a technical point of view but also because we presented an applicable NT review scheme for ICU staff members. On the one hand, most of the nurses and biomedical analysts (BMA) at our hospital work in 12-h shifts, making it easy and feasible to check the NT monitor every 4 h. On the other hand, a 4-h period is easily readable on the NT screen and small EEG changes are still detectable. The mean time it took the respondents to evaluate the 120 slides was 1 h and 22 min. Therefore, a brief screening of a 4-h segment is not time consuming and can be completed in less than one minute.

The short time period in which respondents had to learn and understand the assessment process may be an advantage in terms of simulating 'real life' conditions but also limits our conclusions concerning items that needed more interpretation and instructions. For example, the independent item BS showed a perfect MRA and IRAs. In contrast the corresponding and therefore dependent item 'level of sedation' had an IRA of only 41.10%. In our study protocol an adequate level of sedation was defined by the occurrence of BS patterns at the time a patient suffered from ongoing seizures or showed an elevated ICP. The low level of agreement in this case reflects the insecurities concerning certain definitions that had not been sufficiently internalized by the respondents.

Automatic seizure detection is still rarely used in the ICU because of the many false alarms. NeuroTrend therefore tries to present the complex EEG in simplified form. When it comes to abnormalities, trained staff will alert the specialist. Ideally, a learning process should be initiated and the rate of false alarms should be reduced. This has yet to be verified in future studies. Our experience has shown that increased awareness increases the number of detected seizures in the ICU. Whether this results in over treatment is still to be evaluated and currently discussed by experts (Jordan and Hirsch, 2006; Ferguson et al., 2013).

Artefacts were neglected in the study because they were (1) either detected by the algorithm and removed, or (2) of such a short duration that they did not meet the seizure criteria of the study. Likewise, patterns such as frontal intermittent rhythmic delta activity (FIRDA) and stimulus-induced rhythmic, periodic, or ictal discharges (SIRPIDs), which are frequently encountered in the ICU, have not been dealt with in detail, since it is always necessary to carry out a careful medical examination in case of seizure suspicion. Thus, an initial alert by the nursing staff, even if it is a false alarm, is welcome if these patterns occur.

Because we wanted to simplify the rating process, next to the pattern types SW, RDA, PD and BS we implemented the term ESP where we summarized fast, rhythmic and unequivocal seizure patterns in the alpha and theta range (labeled "rhythmic theta activity; RTA" and "rhythmic alpha activity; RAA" in NT) according to CCET and non-convulsive seizure criteria (Chong and Hirsch, 2005; Hirsch et al., 2013). These patterns are highly suspicious for the presence of subclinical or non-convulsive status epilepticus in neurological critical care patients, because normal alpha activities nearly never occur in this highly selective patient cohort. Furthermore, pathological alpha and theta patterns like alpha and theta coma would be misinterpreted by the algorithm but at least recognized by the responsible physician (Westmoreland et al., 1975; Synek and Synek, 1984). Therefore, and to simplify matters for the respondents involved we defined the NT labels "rhythmic theta activity" (RTA) and "rhythmic alpha activity" (RAA) as clear seizure patterns aware of a possible error. Another weakness of the study is that only one neurophysiologist formed the EXO. However, this shortcoming must be put into relation, since the study is about the recognition and interpretation of displayed patterns and not about the assessment of the computer algorithm itself.

Overall, we believe patients may benefit from the use of computer algorithms like NT as a bedside monitor. We showed that most ICU staff members can easily read and interpret the NT results after a brief training with no restrictions concerning their age, their computer skills or their work experience. Ultimately, this may lead to an increased use of continuous EEG at ICUs as well as to an increased awareness of frequently occurring subclinical seizures and non-convulsive status epilepticus. In many ICUs cEEG is not applied because of the enormous amount of data and the resulting effort of tedious interpretation. NT can facilitate and fasten the evaluation process and with the help of trained ICU personnel cEEG may lose its arduous character and develop into an easy

applicable, non-invasive tool to detect seizures and monitor sedation depth.

5. Conclusion

In the present study, the applicability of a computer algorithm called NeuroTrend (NT) as a bedside monitor for ICU patients who undergo long-term EEG monitoring was assessed. In this specific scenario, NT results were assessed by briefly trained nurses ($n = 15$) and biomedical analysts ($n = 3$). Occurrence of patterns that could indicate seizures as well as evaluation of sedation depth were of particular interest. Detection of seizure patterns showed perfect to substantial agreement in the multirater (MRA) as well as the interrater agreement (IRA). While burst suppression (BS) patterns were clearly identified among the respondents the interpretation of an adequate sedation could only reach moderate agreement. We therefore assume that NT is perfectly suited as a bedside neuro-monitor used by various ICU staff members if an adequate amount of time is invested in staff trainings.

Funding

This study was funded by the Austrian Research Promotion Agency (grant number 826816).

Acknowledgements

We want to thank all nurses and biomedical analysts who participated in the study. We also wish to acknowledge the persons in charge of the intensive care unit who enabled the study.

Conflict of interest: Johannes Herta and Johannes Koren were both partially supported by the Austrian Research Promotion Agency grant as part of their PhD project (personal fees).

Algorithm development was conducted by the "Austrian Institute of Technology" including the authors Franz Fürbass and Manfred Hartmann. The Austrian Institute of Technology is the manufacturer of the EEG software package "encevis", which will include the NeuroTrend algorithms.

Christoph Baumgartner, Andreas Gruber, Angelika Zöchmeister and Arthur Hosmann declare that they have no conflict of interest related to the present work.

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.clinph.2017.04.002.

References

Blood E, Spratt KF. Disagreement on agreement: Two alternative agreement coefficients. Paper 186–2007. SAS Global Forum 2007:1–12.
Chong DJ, Hirsch LJ. Which EEG patterns warrant treatment in the critically ill? Reviewing the evidence for treatment of periodic epileptiform discharges and related patterns. J. Clin. Neurophysiol. 2005;22:79–91.
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.
Eisenberg HM, Frankowski RF, Contant CF, Marshall LF, Walker MD. High-dose barbiturate control of elevated intracranial pressure in patients with severe head injury. J. Neurosurg. 1988;69:15–23.
Ferguson M, Bianchi MT, Sutter R, Rosenthal ES, Cash SS, Kaplan PW, et al. Calculating the risk benefit equation for aggressive treatment of non-convulsive status epilepticus. Neurocrit. Care 2013;18:216–27.
Friedman D, Claassen J, Hirsch LJ. Continuous electroencephalogram monitoring in the intensive care unit. Anesth. Analg. 2009;109:506–23.
Fürbass F, Herta J, Koren J, Westover MB, Hartmann MM, Gruber A, et al. Monitoring burst suppression in critically ill patients: Multi-centric evaluation of a novel method. Clin. Neurophysiol. 2016;127:2038–46.
Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. Clin. Neurophysiol. 2015;126:1124–31.
Gwet KL. Handbook of inter-rater reliability. In: Adv. Anal.. fourth ed. LLC; 2014.
Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: Automatic EEG artifact removal for epilepsy monitoring. Neurophysiol. Clin. 2014;44:479–90.
Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. Epilepsy Behav. 2015;49:273–9.
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.
Husain AM, Horn GJ, Jacobson MP. Non-convulsive status epilepticus: usefulness of clinical features in selecting patients for urgent EEG. J. Neurol. Neurosurg. Psychiatr. 2003;74:189–91.
Jordan KG, Hirsch LJ. In nonconvulsive status epilepticus (NCSE), treat to burst-suppression: pro and con. Epilepsia 2006;47:41–5.
Kaplan PW. Assessing the outcomes in patients with nonconvulsive status epilepticus: nonconvulsive status epilepticus is underdiagnosed, potentially overtreated, and confounded by comorbidity. J. Clin. Neurophysiol. 1999;16:341–52.
Koren J, Herta J, Draschtak S, Pötzl G, Pirker S, Fürbass F, et al. Prediction of rhythmic and periodic EEG patterns and seizures on continuous EEG with early epileptiform discharges. Epilepsy Behav. 2015;49:286–9.
Ney JP, van der Goes DN, Nuwer MR, Nelson L, Eccher MA. Continuous and routine EEG in intensive care: utilization and outcomes, United States 2005–2009. Neurology 2013;81:2002–8.
Sutter R, Stevens RD, Kaplan PW. Continuous electroencephalographic monitoring in critically ill patients: indications, limitations, and strategies. Crit. Care Med. 2013;41:1124–32.
Synek VM, Synek BJ. Theta pattern coma, a variant of alpha pattern coma. Clin. Electroencephalogr. 1984;15:116–21.
Vespa PM, Nuwer MR, Nenov V, Ronne-Engstrom E, Hovda DA, Bergsneider M, et al. Increased incidence and impact of nonconvulsive and convulsive seizures after traumatic brain injury as detected by continuous electroencephalographic monitoring. J. Neurosurg. 1999;91:750–60.
Westmoreland BF, Klass DW, Sharbrough FW, Reagan TJ. Alpha-coma. Electroencephalographic, clinical, pathologic, and etiologic correlations. Arch. Neurol. 1975;32:713–8.
Young GB, Jordan KG, Doig GS. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. Neurology 1996;47:83–9.

6.7 Paper A7: Assessment of rhythmic and periodic detections

Title: Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend

Authors: Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, Gruber A

Published in: Epilepsy & Behavior

Year: 2015

Authors' contribution: Fürbass Franz developed the computer algorithms for detection of rhythmic and periodic patterns that were used in this study. Franz Fürbass worked on statistical analysis of computer detections and visualizations of results for the manuscript. Further, manuscript editing of parts of the method section concerning the algorithms was done by Franz Fürbass. Initial manuscript writing, data annotations, EEG recording and all clinical work was done by first and coauthors.

Epilepsy & Behavior 49 (2015) 273–279

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: A new approach for screening continuous EEG in the intensive care unit

J. Herta [a,*], J. Koren [b], F. Fürbass [c], M. Hartmann [c], T. Kluge [c], C. Baumgartner [b], A. Gruber [a]

[a] Department of Neurosurgery, Medical University of Vienna, Vienna, Austria
[b] Karl Landsteiner Institute for Clinical Epilepsy Research and Cognitive Neurology, 2nd Neurological Department, General Hospital Hietzing with Neurological Center Rosenhuegel, Vienna, Austria
[c] AIT Austrian Institute of Technology GmbH, Digital Safety & Security Department, Vienna, Austria

ARTICLE INFO

Article history:
Revised 24 April 2015
Accepted 28 April 2015
Available online 23 May 2015

Keywords:
Epileptic seizure detection
Automatic
Prospective multicenter study
Continuous EEG
Intensive care unit
Screening device

ABSTRACT

Background: NeuroTrend is a computational method that analyzes long-term scalp EEGs in the ICU according to ACNS standardized critical care EEG terminology (CCET) including electrographic seizures. At present, it attempts to become a screening aid for continuous EEG (cEEG) recordings in the ICU to facilitate the review process and optimize resources.
Methods: A prospective multicenter study was performed in two neurological ICUs including 68 patients who were subjected to video-cEEG. Two reviewers independently annotated the first minute of each hour in the cEEG according to CCET. These segments were also screened for faster patterns with frequencies higher than 4 Hz. The matching annotations (2911 segments) were then used as gold standard condition to test sensitivity and specificity of the rhythmic and periodic pattern detection of NeuroTrend.
Results: Interrater agreement showed substantial agreement for localization (main term 1) and pattern type (main term 2) of the CCET. The overall detection sensitivity of NeuroTrend was 94% with high detection rates for periodic discharges (PD = 80%) and rhythmic delta activity (RDA = 82%). Overall specificity was moderate (67%) mainly because of false positive detections of RDA in cases of general slowing. In contrast, a detection specificity of 88% for PDs was reached. Localization revealed only a slight agreement between reviewers and NeuroTrend.
Conclusions: NeuroTrend might be a suitable screening tool for cEEG in the ICU and has the potential to raise efficiency of long-term EEG monitoring in the ICU. At this stage, pattern localization and differentiation between RDA and general slowing need improvement.

This article is part of a Special Issue entitled "Status Epilepticus".

1. Introduction

The increased use of continuous EEG (cEEG) in the intensive care unit (ICU) for patients with critical illness has been propagated lately by many authors [1–7]. This is due to the fact that nonconvulsive seizures (NCSs) and nonconvulsive status epilepticus (NCSE) occur more often than previously anticipated [8]. Sutter et al. revealed that, after implementing cEEG into clinical practice, the rate of NCS diagnosis increased significantly compared with previous diagnostics. This might be not only due to higher observer awareness and greater availability of EEG but also due to longer observation periods [1]. Incident rates diverge a lot, as the studied patient populations are seldom homogeneous and inclusion criteria for cEEG vary between studies. i.e., 19% of the patients had NCSs in a study from Claassen [5] compared with 34% found in a study of Jordan [9]. Patients who suffered from convulsive status epilepticus often convert to NCSE after their convulsions have stopped [10]. Also, patients with altered state of consciousness and clinical features like subtle motor activity and abnormal eye movements may suffer from NCE or NCSE [11]. Privitera could demonstrate that in 198 patients with altered state of consciousness, 37% had NCSs [12]. In comatose patients, there is nearly no evidence of

Abbreviations: ACNS, American Clinical Neurophysiology Society; AIT, Austrian Institute of Technology; BI, bilateral independent; cEEG, continuous electroencephalography; CCET, critical care EEG terminology; EEG, electroencephalography; FN, false negative; FP, false positive; GHV, General Hospital Vienna; G, generalized; ICU, intensive care unit; κ, kappa; L, lateralized; MF, multifocal; MT1, main term 1; MT2, main term 2; NCR, Neurological Center Rosenhuegel; NCSs, nonconvulsive seizures; NCSE, nonconvulsive status epilepticus; NOPAT, no pattern; NT, NeuroTrend; QEEG, quantitative EEG; PD, periodic discharge; RAA, rhythmic alpha activity; RDA, rhythmic delta activity; RDA + S, rhythmic delta activity plus frequent intermixed sharp waves or spikes; RTA, rhythmic theta activity; SE, sensitivity; SP, specificity; SW, rhythmic spike-and-wave activity; TN, true negative; TP, true positive.

* Corresponding author at: Medical University of Vienna, Department of Neurosurgery, Währinger Gürtel 18-20, 1090 Vienna, Austria. Tel.: +43 1 40400 25650; fax: +43 1 40400 45660.
E-mail address: johannes.herta@meduniwien.ac.at (J. Herta).

http://dx.doi.org/10.1016/j.yebeh.2015.04.064
1525-5050/

seizure activity without EEG. Towne showed that in 236 coma patients with unclear genesis, 8% had NCSE [13]. Therefore, cEEG still remains the gold standard for reliable diagnosis of NCSs and NCSE. Whether NCSE is a predictor for bad outcome in patients with critical illness is difficult to assess because treatment effects, causative medical disorder, and complications are difficult to separate. Until now, seizure duration and delayed diagnosis of NCSs and NCSE are the only two independent parameters known to increase morbidity and mortality [14].

Recently, the use of cEEG in patients with critical illness has been reported to be associated with a favorable outcome [15]. Continuous analysis of cEEG by a trained expert reviewing segments of 10 s each is virtually impossible but would enable early and adapted treatment for the patient. Quantitative EEG (QEEG) addressed this important problem by evaluating the EEG in real time and by showing amplitude, power, frequency, and rhythmicity in compressed time scales [16]. The downside of QEEG techniques is the oversimplified approach to extract EEG information. This leads to a predisposition to false positive errors, and seizure activity can be missed in the shadow of high-amplitude artifacts [17].

For a long time period, many authors tried to define and classify NCSs and NCSE including or excluding EEG patterns frequently seen in patients with critical illness such as periodic discharges and fluctuating rhythmic patterns [5,11,14,18–21]. In 2013, the American Clinical Neurophysiology Society (ACNS) developed a standardized critical care EEG terminology (CCET) to facilitate communication between researchers [19].

Based on the CCET, the computational encephalography research group of the Austrian Institute of Technology (AIT) developed an automated detection and trending method called NeuroTrend (NT) with the aim to assist and facilitate the review process of cEEG [22]. In this work, we evaluate the performance of NT in terms of sensitivity, specificity, and interrater agreement.

2. Methods

NeuroTrend (NT) is a computational method that automatically detects rhythmic and periodic patterns in surface EEG and visualizes the results graphically. The definition of rhythmic and periodic patterns follows the guidelines of the American Clinical Neurophysiology Society Terminology [19]. Additionally, rhythmic patterns of more than 4 Hz are detected to cover the whole spectrum of electrographic seizure patterns. The aim of this work is to evaluate the sensitivity and specificity of detected patterns compared with manual-annotated EEG segments. The technical methodology used in the rhythmic and periodic pattern detection was described recently by Fürbass [22]. In this work, NeuroTrend version 1.1 was used for the calculation of all detections (NeuroTrend V1.1, www.eeg-vienna.com).

2.1. Data acquisition and patient selection

We prospectively recorded long-term video-EEGs (n = 68) using the international 10–20 electrode system with a sampling rate of 256 Hz. The recording was done at the neurological ICU of the Neurological Center Rosenhuegel (NCR) and the neurosurgical ICU of the General Hospital Vienna (GHV) using a Micromed EEG recording system (SystemPLUS Evolution 1.04.95) between March 1, 2013 and September 1, 2014. Only cEEGs with a recording period longer than 20 h were included. At least nineteen of twenty-one cup electrodes (including reference and ground electrode) had to be functional over the whole recording period. Gold cup electrodes (Genuine Grass Gold Disc electrodes) as well as conductive plastic cup electrodes (Ives EEG Solutions) were used for recordings. Gold cup electrodes were used preferentially. Plastic cup electrodes were used in cases where CT scans had to be carried out regularly.

The treating physician conducted patient selection according to the following criteria:

a) Remote eye movement abnormalities or subtle myoclonus
b) Short time period since patient's admission and neurologic injury
c) Low Glasgow Coma Scale (GCS).

The criteria applied were expected to filter out as many cases of NCSs/NCSE as possible according to Husain et al. [11] and Claassen et al. [5]. Patients younger than 18 years and patients with a high risk of infection (e.g., because of expanded wounds) were excluded from the study.

2.2. Validation strategy

In a first step, two clinical neurophysiologists from the recording centers NCR and GHV were asked to annotate the first minute of each hour in the video-EEG recording of their own center. The reviewers who were naïve to these video-EEGs had to screen for mechanical ventilation artifacts, electrocardiogram artifacts, and rhythmic movements. Electroencephalography pieces including these artifacts were labeled accordingly. Video and sound data were then separated from the EEG, and the EEGs were anonymized. The anonymized EEGs from both sites were then merged, resulting in a dataset of 68 long-term EEG recordings.

In a second step, both evaluators were asked to annotate rhythmic and periodic patterns in the one-minute annotation segments of all 68 EEGs from both centers. The definition of these patterns followed the main term 2 definition (MT2) in the CCET guidelines [19]. The MT2 definition was extended to include rhythmic pattern of more than 4 Hz. Both reviewers were firm with the recent version of CCET and had used ACNS training slides several times. The reviewers could use the EEG viewer without any restriction in relation to montage or filters. Several nonoverlapping annotations were allowed. Each annotation may have an arbitrary start and an end position but has to be fully included in the annotation minute. For each annotation, the reviewer was allowed to choose between one of the following pattern types: periodic discharges (PDs), rhythmic delta activity (RDA), rhythmic theta activity (RTA), rhythmic alpha activity (RAA), and rhythmic spike-and-wave activity (SW). If the reviewer did not insert any annotation in the one-minute interval, it was counted as no pattern (NOPAT).

In addition to the pattern type, a localization property had to be defined. This property was defined according to the CCET [19] as main term 1 (MT1): generalized (G), lateralized (L), multifocal (MF), and bilateral independent (BI).

The annotations from the two reviewers were then used as gold standard condition to test the sensitivity and specificity of the rhythmic and periodic pattern detection of NT. Evaluation scripts were used to automatically read the reviewer annotations and to calculate the detection performance numbers. Artifact annotations from the first annotation step were only assessed if no other markers were placed in the annotation segment.

2.3. Statistical methods

The detection performance was defined by assigning one of four possible test conditions to each annotation minute: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). A pattern was counted as TP if one of the detected patterns in the annotation minute matched the gold standard annotation. A gold standard annotation was defined as an agreement between both reviewers. If no agreement between the two reviewers was met, the annotation interval was excluded from the calculation. A gold standard annotation without a matching detection in the annotation minute was counted as FN. An annotation segment with one or several detections that do not match the

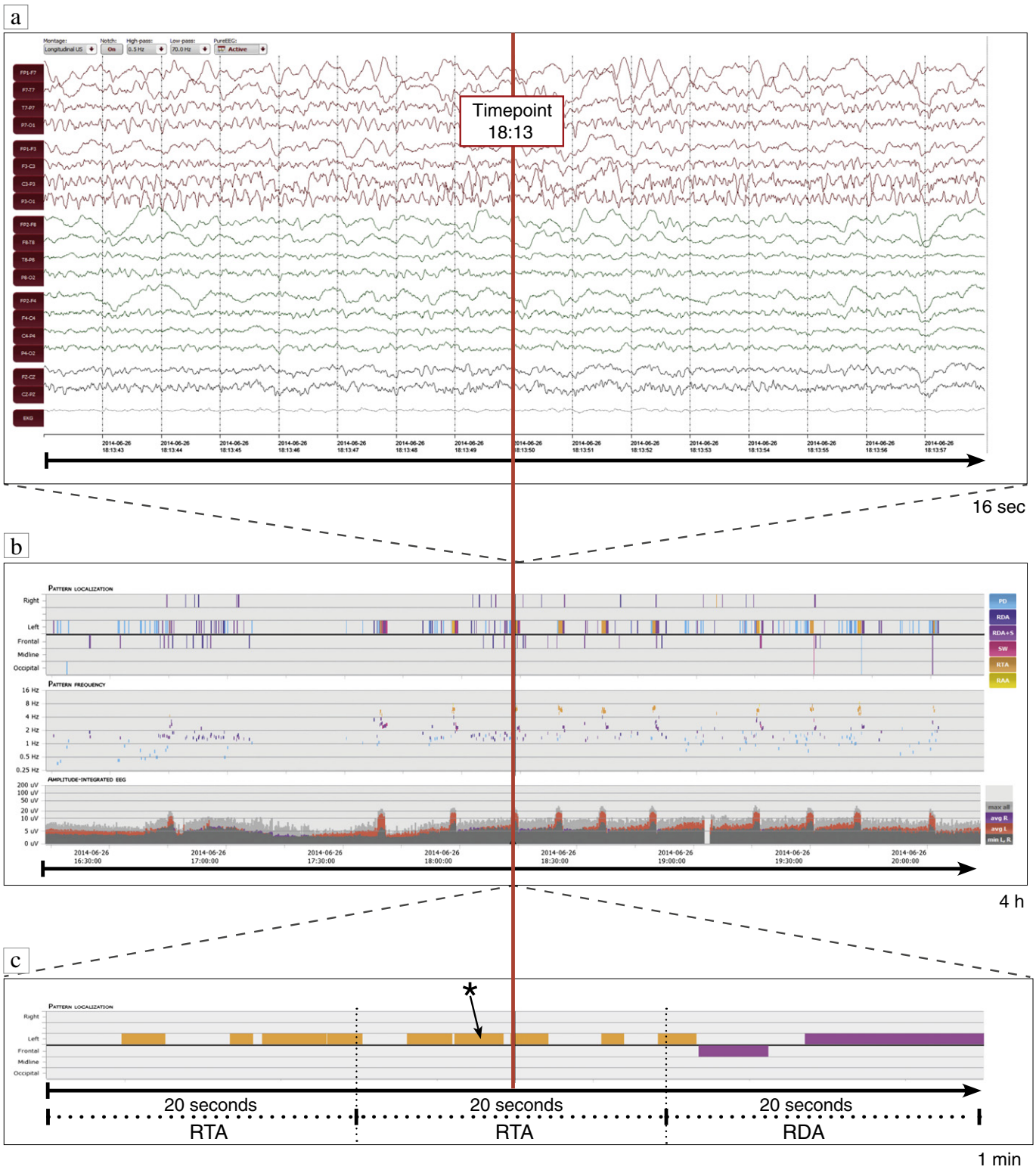

Fig. 1. Explanation of validation strategy: a) shows 16 s of raw EEG with left-sided rhythmic theta activity (RTA). The same EEG at the same time point is represented as a vertical red line in the NeuroTrend data illustration underneath. b) Demonstrates the usage of NeuroTrend as it displays 4 h (variable adjustment of time) of cEEG color-coded on one page. A clear repetition of RTA (orange bars, color code is displayed at the right side) occurring nearly every 15 min can be seen in the left hemisphere. AEEG also shows the 15-minute intervals but not the type of the pattern. c) One minute was extended out to illustrate the assessment process. The whole 1-minute interval shows detections of RTA (orange) and RDA + S (violet) when used for sensitivity and specificity calculation respectively. When divided into 20-second segments, segments 1, 2, and 3 show RTA, but only segment 3 includes RDA detection. For calculation of Cohen's kappa (κ) values, the pattern type and localization with the highest percentage of time coverage are used. The 1-minute segment, therefore, counts as lateralized RTA (*). The 20-second segments are counted twice as lateralized RTA and once as frontal RDA.

gold standard annotation was counted as FP. An annotation segment without gold standard annotation (NOPAT) and without any detection was counted as TN. Sensitivity (SE) and specificity (SP) were calculated according to the following formulas:

$$SE[\%] = \frac{TP}{TP + FN} * 100$$
$$SP[\%] = \frac{TN}{TN + FP} * 100.$$

To verify the interrater agreement between both reviewers, kappa (κ) values were calculated for each annotation interval. The same approach was used for the comparison between reviewers and NT for the categorical parameters MT1 and MT2. Both κ-statistics were calculated in two passes. First, the standard annotation interval of 60 s was used. Second, the same segment was divided into three shorter segments of 20 s, each offering a more detailed analysis. In each annotated segment, the annotation with the longest duration time was used to calculate κ-statistics. The rationale behind having two different evaluation intervals lies in NT's intentional usage as trending software and is demonstrated in Fig. 1. While we hypothesize that the 20-second time interval gives us a statement about the actual hit rate, the 60-second time interval should reflect the progression of EEG patterns and their trend.

The following qualitative classifications are used to categorize κ values into different ranges: poor agreement: ≤0; slight agreement: 0.01–0.20; fair agreement: 0.20–0.40; moderate agreement: 0.40–0.60; substantial agreement: 0.60–0.80; and almost perfect agreement: 0.80–1.00 [23].

3. Results

3.1. Patient characteristics

In the study period, 80 patients were monitored with continuous video-EEG. Five patients were excluded from the study because of insufficient data quality, long time periods with detached electrodes, or less than 17 electrodes at the beginning of recording. Another 7 patients were excluded because of a recording duration of less than 20 h. The mean age was 58 (±16.5) years with a female to male ratio of 35:33. Plastic cup electrodes were used in 27 cases, while the majority of patients (n = 41) were monitored with gold cup electrodes. Continuous Electroencephalography (cEEG) of 4813 h were recorded in total with a median length of 48 h. This led to 2911 segments of 1 min each available for evaluation purposes.

3.2. Interrater agreement

Interrater agreement of main term 1 (MT1) as well as main term 2 (MT2) according to standardized critical care EEG terminology (CCET) as well as electrographic seizures was performed between the two reviewers [19]. Main term 1 showed a substantial agreement in both short (20-second) and long (60-second) annotation intervals. The same Cohen's kappa (κ) values could be found for MT2 and are presented in Table 1. A good agreement between reviewers was crucial to enable further validation between reviewers and automated pattern detection. Looking at each MT2 pattern separately, rhythmic delta activity (RDA) was the pattern with the highest disagreement between reviewers and deteriorated further in the more detailed 20-second analysis (Fig. 2). Rhythmic alpha activity (RAA), spike-and-wave activity (SW), and rhythmic theta activity (RTA) showed a good agreement but occurred in very low numbers.

3.3. Validation of main term 2 (MT2)

Sensitivity and specificity of NeuroTrend (NT) for MT2 patterns are shown in Fig. 3. While sensitivity for the detection of any MT2 pattern is high (0.94), specificity is low (0.67) for 60-second annotations with a positive predictive value of 0.2 and a negative predictive value of 0.99. In the shorter 20-second time interval, sensitivity declines to 0.84, while specificity rises to 0.78. The same can be seen for periodic discharges (PDs) and RDA. Sensitivity declines from 0.8 to 0.59 for PD and 0.82 to 0.71 for RDA if compared between the 60-second annotation interval and the shorter 20-second annotation interval. Specificity inversely rises from 0.88 to 0.93 for PD and from 0.72 to 0.83 for RDA. Rhythmic theta activity and rhythmic spike-and-wave activity showed high specificity and sensitivity, while RTA was detected with a high specificity solely. Because of the low number of RTA, SW, and RAA in our study, no serious conclusions can be drawn for these patterns.

κ-Statistic showed similar results in regard to the agreement between NT and the reviewer gold standard (Table 2). Included segments in which no patterns were found (NOPAT) κ-statistic revealed a fair agreement between NT and the gold standard with a decline of agreement from 0.38 for 60-second annotations to 0.24 for 20-second annotations. This decline cannot be reproduced if no patterns (NOPATs) are

Table 1
Interrater agreement of main term 2 according to standardized critical care EEG terminology including electrographic seizures between two independent reviewers. The results of the longer sixty-second evaluation intervals are shown as opposed to the shorter twenty-second intervals. Overall, there is substantial agreement between the two reviewers regardless of the chosen evaluation interval. It is evident that rhythmic slowing of the EEG is often difficult to differentiate from RDA. Numbers of RAA and RTA are too low to make a reasonable statement. NOPAT = no pattern, PD = periodic discharge, RAA = rhythmic alpha activity, RDA = rhythmic delta activity, RTA = rhythmic theta activity, SW = rhythmic spike-and-wave activity.

		Reviewer 2											
		NOPA		PD		RAA		RDA		RTA		SW	
		60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.
Reviewer 1	NOPA	**1489**	**7149**	38	168	0	3	24	57	1	3	0	2
	PD	21	105	**242**	**393**	0	0	4	9	1	2	0	0
	RAA	0	0	0	0	**2**	**3**	0	0	1	2	0	0
	RDA	43	103	6	5	0	0	**74**	**109**	0	0	1	2
	RTA	0	1	2	2	0	0	1	2	**5**	**10**	0	0
	SW	2	4	0	0	1	2	0	0	1	0	**8**	**13**
Cohens Kappa:		**0.79**	**0.67**										

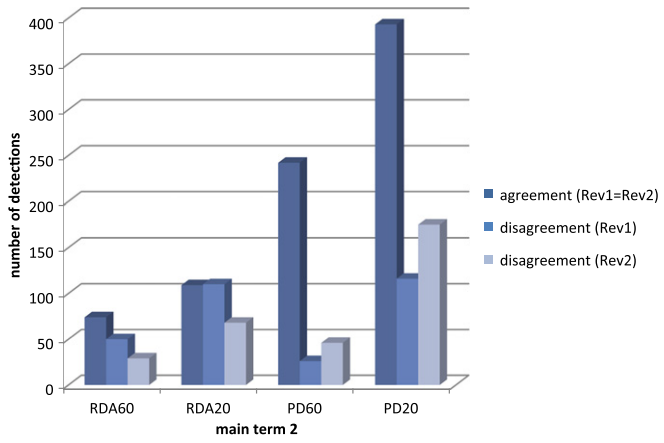

Fig. 2. The number of interrater agreements (Rev1 = Rev2) as well as disagreements (Rev1, Rev2) is shown for rhythmic delta activity (RDA) and periodic discharges (PDs) separately for 20- and 60-second annotation intervals. PDs show a substantial agreement with an even rise in agreements and disagreements in both annotation intervals. Agreement for RDA on the contrary deteriorates with more detailed analysis.

excluded from the calculation. This is mainly due to the fact that NT often detected RDA falsely when NOPAT was assigned. This observation was getting worse if a higher time resolution was used for calculation.

3.4. Validation of main term 1 (MT1)

While MT1 showed a substantial agreement (0.79) between reviewers, κ between reviewers and NT is poor (0.16) if NOPATs are not included in the statistic.

3.5. Artifacts

Artifacts in cEEG play a major role in the ICU and can disturb automatic pattern detection heavily. NeuroTrend, therefore, uses an artifact rejection module called "PureEEG" which has been described recently [24]. During the review process, cEEGs were reviewed for artifacts in the annotation segments with the help of video and sound recordings. Artifacts were marked but not excluded in NT analysis. During evaluation of these segments, it could be seen that NT was relatively stable for artifacts with 664 (60.3%) detections as NOPAT out of 1102 artifact markers resulting in 438 (39.7%) falsely labeled segments. In terms of pattern, mainly NT's RDA detection was triggered by artifacts and reached 295 (67.3%) false detections. False detection rate of an artifact as a rhythmic or periodic pattern was 108 (24.7%) for PD, 32 (7.3%) for RTA, and 3 (0.7%) for SW. Artifacts were never detected falsely as RAA.

4. Discussion

In this article, we assessed and validated the rhythmic and periodic pattern detection performance of an automated computer algorithm called NeuroTrend (NT) [22]. The aim of NT is the quick visualization of several hours of cEEG recordings based on ACNS standardized critical care EEG terminology (CCET) including rhythmic patterns with frequencies higher than 4 Hz [19].

While conventional QEEG displays compressed raw EEG in terms of technical measurements, NT transcribes automatic detections into neurophysiological established wording [24]. Another big difference between QEEG and NT consists in the prior usage of artifact rejection. Therefore, QEEG may facilitate the review process of larger cEEG files but comprises the risk of false interpretation due to processed artifacts. Both methods have got a strong data compression property. NeuroTrend allows the graphical representation of large cEEG files, giving the reviewer the possibility to screen several hours to days of cEEG on a few pages. It is important to stress that NT's focus lies in displaying trend data. It should depict changes in EEG over longer time periods and not exact values at a certain time point. That is why an unconventional approach to assess correct pattern analysis has been chosen. While most studies assess monitoring devices by interrater agreement at preselected time points, we tried to approach real cEEG testing conditions by using unselected time intervals [25,26]. Furthermore, final calculations of agreement were automated to minimize confounders. A time interval of 1 min every hour was randomly chosen regardless of recording quality, presence or absence of artifact, and EEG pattern. To enable the evaluation of correct EEG pattern detection during this minute, we separately analyzed a segment of 60-second as well as three 20-second fragments. It might seem that the detection of shorter segments

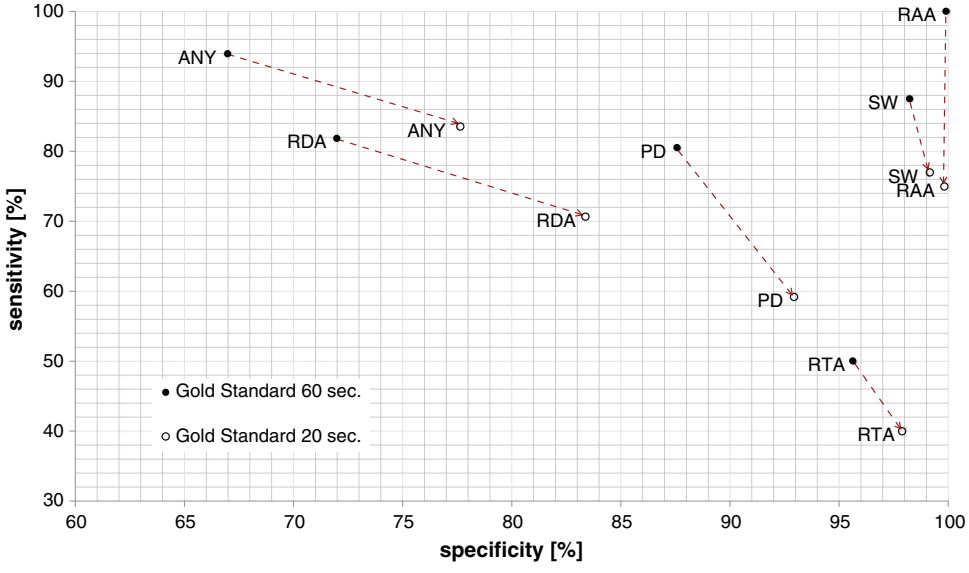

Fig. 3. Detection performance of NeuroTrend. NeuroTrend has a high sensitivity compared with the gold standard of two reviewers (60 s) for PD, RDA, and ANY patterns. Therefore, it is suitable as a screening tool. Specificity for PD is high, while specificity for RDA and, therefore, also for ANY is moderate. The shorter 20-second annotation interval showed us a shift towards a higher specificity at the cost of a lower sensitivity. For the patterns SW, RTA, and RAA, not enough data were collected to be significant. PD = periodic discharge, RAA = rhythmic alpha activity, RDA = rhythmic delta activity, RTA = rhythmic theta activity, SW = rhythmic spike-and-wave activity, ANY = all patterns previously mentioned together.

Table 2
Kappa statistic of main term 2 according to standardized critical care EEG terminology including electrographic seizures between reviewers (gold standard) and NeuroTrend. The results of the longer sixty-second evaluation intervals are shown as opposed to the shorter twenty-second intervals. Overall, there is a fair agreement with a considerable drawback in the shorter evaluation interval. This is due to the increased number of false positive detections, especially for RDA. Excluding NOPAT and considering only intervals where reviewers and NeuroTrend found a pattern highlight this finding. No difference can be found in Cohen's kappa between the sixty-second evaluation and the twenty-second evaluation anymore. NOPAT = no pattern, PD = periodic discharge, RAA = rhythmic alpha activity, RDA = rhythmic delta activity, RTA = rhythmic theta activity, SW = rhythmic spike-and-wave activity.

		NeuroTrend											
		NOPA		PD		RAA		RDA		RTA		SW	
		60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.	60 sec.	20 sec.
Gold standard	NOPA	**1052**	**5607**	130	438	0	10	300	999	7	77	0	18
	PD	18	73	**155**	**216**	0	0	58	85	0	0	12	21
	RAA	0	0	0	0	**0**	**1**	0	0	3	3	0	0
	RDA	3	15	17	19	0	0	**58**	**74**	2	1	0	0
	RTA	0	0	5	6	0	0	2	6	**1**	**2**	0	0
	SW	0	0	2	4	0	0	2	4	0	0	**4**	**5**
	Cohens Kappa (NOPA included):			**0.38**	**0.24**								
	Cohens Kappa (NOPA excluded):			**0.38**	**0.36**								

is more precise to predict the hit rate and, therefore, sensitivity and specificity. However, in contrast to spike detection and other alarm devices, the detection of a single measurement by the algorithm is not its primary purpose. Similar to a human EEG reviewer who has the ability to focus on relevant EEG changes, NT should display the progression of predominant ongoing patterns. That is why our primary study outcome was measured by the longest pattern available during 60 s as presented in Fig. 1.

Because the validation process should be as close to real screening conditions as possible, all cEEG data were recorded prospectively, no cEEG file was used for preceding algorithm development, and artifacts were not removed.

Interrater agreement between both reviewers showed a substantial agreement (0.6–0.8) in MT1 and MT2. A good agreement was essentially required to establish a gold standard condition against which NT could be tested. Our findings of a high interrater agreement correspond to previous interobserver studies that tested the 2012 version of the ACNS nomenclature [27,28].

Because EEG segments were not preselected, the marker NOPAT for EEG segments without rhythmic or periodic pattern was introduced.

The assessment of NT showed that it might be a useful screening tool for cEEG. On the one hand, NT revealed a high overall sensitivity (0.94) for MT2 patterns and a low rate of false negative detections. On the other hand, overall specificity (0.67) was low with only one true hit out of five detections. It should not be overlooked that sensitivity declines and specificity rises when the shorter 20-second interval is used for evaluation. Specificity for PD is good (0.88) while specificity for any pattern and RDA is moderate. Table 2 illustrates nicely the large number of false positive RDA detections in segments where the reviewers assigned NOPAT. Likewise to moderate results in specificity for RDA in our study and poor raw percentage of positive agreement for RDA (57%), a recent study showed that RDA is often difficult to distinguish from general slowing [29].

Evaluation of MT1 revealed that NT has a tendency to detect patterns as lateralized. Patterns with anterior - posterior lag and hemispheric differences are the cause of this behavior.

Electroencephalography is prone to artifacts, and many forms of artifacts in cEEG at the ICU have been described [30]. Like already outlined above NT can distinguish itself from other screening tools by an artifact rejection property called "PureEEG" [24]. Artifacts can falsely trigger pattern detections of NT in 39.7% of all prior labeled artifacts. Especially RDA was triggered by artifacts, which lead to a high false positive detection of RDA.

Limitations of the study can be seen in a small number of patients and the unequal distribution of patterns. These issues had to be condoned to enable a prospective study. Furthermore, it could be argued that including only segments where solely two reviewers gave an agreement may exclude potential difficult patterns from evaluation. Once again, it has to be stressed that interrater agreement for CCET has been proven high in our as well as in previous studies [27,28].

5. Conclusion

NeuroTrend might become a suitable screening tool for cEEG and has the potential to raise the efficiency of long-term EEG monitoring in the ICU. As it still offers the possibility to switch between trend data and raw EEG, it does not interfere the review process and can be used complementary to raw EEG, which remains gold standard for EEG interpretation. At this stage, pattern localization and differentiation between RDA and general slowing need further improvement.

Disclosure

Research and development of NeuroTrend were supported by The Austrian Research Promotion Agency grant 826816 (EpiMon).

Johannes Herta and Johannes Koren were both partially supported by the FFG grant.

Algorithm development was conducted by the "Austrian Institute of Technology" including the authors Franz Fürbass, Manfred Hartmann, and Tilman Kluge. The Austrian Institute of Technology is the manufacturer of the EEG software package "Encevis", which will include the NeuroTrend algorithms.

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

[1] Sutter R, Fuhr P, Grize L, Marsch S, Rüegg S. Continuous video-EEG monitoring increases detection rate of nonconvulsive status epilepticus in the ICU. Epilepsia 2011;52:453–7.
[2] 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.
[3] Vespa PM, Nuwer MR, Nenov V, Ronne-Engstrom E, Hovda DA, Bergsneider M, et al. Increased incidence and impact of nonconvulsive and convulsive seizures after traumatic brain injury as detected by continuous electroencephalographic monitoring. J Neurosurg 1999;91:750–60.
[4] Friedman D, Claassen J, Hirsch LJ. Continuous electroencephalogram monitoring in the intensive care unit. Anesth Analg 2009;109:506–23.
[5] 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.
[6] Oddo M, Carrera E, Claassen J, Mayer SA, Hirsch LJ. Continuous electroencephalography in the medical intensive care unit. Crit Care Med 2009;37:2051–6.
[7] Jette N, Claassen J, Emerson RG, Hirsch LJ. Frequency and predictors of nonconvulsive seizures during continuous electroencephalographic monitoring in critically ill children. Arch Neurol 2006;63:1750–5.
[8] Kaplan PW. Assessing the outcomes in patients with nonconvulsive status epilepticus: nonconvulsive status epilepticus is underdiagnosed, potentially overtreated, and confounded by comorbidity. J Clin Neurophysiol 1999;16:341–52 [discussion353].
[9] Jordan KG. Continuous EEG, monitoring in the neuroscience intensive care unit and emergency department. J Clin Neurophysiol 1999;16:14–39.
[10] DeLorenzo RJ, Hauser WA, Towne AR, Boggs JG, Pellock JM, Penberthy L, et al. A prospective, population-based epidemiologic study of status epilepticus in Richmond, Virginia. Neurology 1996;46:1029–35.
[11] Husain AM, Horn GJ, Jacobson MP. Non-convulsive status epilepticus: usefulness of clinical features in selecting patients for urgent EEG. J Neurol Neurosurg Psychiatry 2003;74:189–91.
[12] 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.
[13] Towne AR, Waterhouse EJ, Boggs JG, Garnett LK, Brown AJ, Smith Jr J, et al. Prevalence of nonconvulsive status epilepticus in comatose patients. Neurology 2000;54:340.
[14] Young GB, Jordan KG, Doig GS. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. Neurology 1996;47:83–9.
[15] Ney JP, van der Goes DN, Nuwer MR, Nelson L, Eccher MA. Continuous and routine EEG in intensive care: utilization and outcomes, United States 2005–2009. Neurology 2013;81:2002–8.
[16] Foreman B, Claassen J. Quantitative EEG for the detection of brain ischemia. Crit Care 2012;16:216.
[17] Nuwer M. Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society. Neurology 1997;49:277–92.
[18] Chong DJ, Hirsch LJ. Which EEG patterns warrant treatment in the critically ill? Reviewing the evidence for treatment of periodic epileptiform discharges and related patterns. J Clin Neurophysiol 2005;22:79–91.
[19] 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.
[20] Holtkamp M, Meierkord H. Nonconvulsive status epilepticus: a diagnostic and therapeutic challenge in the intensive care setting. Ther Adv Neurol Disord 2011;4:169–81.
[21] Sutter R, Kaplan PW. Electroencephalographic criteria for nonconvulsive status epilepticus: synopsis and comprehensive survey. Epilepsia 2012;53(Suppl. 3):1–51.
[22] Fürbass F. Automatic detection of rhythmic and periodic patterns in critical care EEG based on ACNS standardized terminology. Neurophysiol Clin 2015 [submitted for publication].
[23] Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977;33:363–74.
[24] Hartmann MM, Schindler K, Gebbink TA, Gritsch G, Kluge T. PureEEG: automatic EEG artifact removal for epilepsy monitoring. Neurophysiol Clin 2014;44:479–90.
[25] Wilson SB, Turner CA, Emerson RG, Scheuer ML. Spike detection II: automatic, perception-based detection and clustering. Clin Neurophysiol 1999;110:404–11.
[26] Sierra-Marcos A, Scheuer ML, Rossetti AO. Seizure detection with automated EEG analysis: a validation study focusing on periodic patterns. Clin Neurophysiol 2015;126:456–62.
[27] 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.
[28] Mani R, Arif H, Hirsch LJ, Gerard EE, LaRoche SM. Interrater reliability of ICU EEG research terminology. J Clin Neurophysiol 2012;29:203–12.
[29] 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.
[30] Young GB, Campbell VC. EEG monitoring in the intensive care unit: pitfalls and caveats. J Clin Neurophysiol 1999;16:40–5.

6.8 Paper A8: Effect of electrode reduction

Title: Reduced electrode arrays for the automated detection of rhythmic and periodic patterns in the intensive care unit: frequently tried, frequently failed?

Authors: Herta J, Koren J, **Fürbass F**, Hartmann M, Gruber A, Baumgartner C

Published in: Clinical neurophysiology

Year: 2017

Authors' contribution: Fürbass Franz developed the computer algorithms for detection of rhythmic and periodic patterns that were used in this study. Franz Fürbass also worked on modifications of algorithms to enable electrode reduction as well as preparation of reduced EEG files for computer analysis. Statistical analysis of computer detections and graphical evaluation for the manuscript were done by Franz Fürbass. Initial manuscript editing, data annotation, EEG recording and all clinical work were done by first and coauthors.

Clinical Neurophysiology 128 (2017) 1524–1531

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Reduced electrode arrays for the automated detection of rhythmic and periodic patterns in the intensive care unit: Frequently tried, frequently failed?

J. Herta [a,*], J. Koren [b], F. Fürbass [c], M. Hartmann [c], A. Gruber [a], C. Baumgartner [b,d]

[a] Department of Neurosurgery, Medical University of Vienna, Vienna, Austria
[b] Karl Landsteiner Institute for Clinical Epilepsy Research and Cognitive Neurology, 2nd Neurological Department, General Hospital Hietzing with Neurological Center Rosenhuegel, Vienna, Austria
[c] AIT Austrian Institute of Technology GmbH, Digital Safety & Security Department, Vienna, Austria
[d] Department of Epileptology and Clinical Neurophysiology, Sigmund Freud University, Vienna, Austria

See Editorial, pages 1519–1521

ARTICLE INFO

Article history:
Accepted 18 April 2017
Available online 26 April 2017

Keywords:
Epileptic seizure detection
Prospective multi-center study
Continuous EEG
Intensive care unit
Computer algorithm

HIGHLIGHTS

- First study that systematically evaluates the effect of automated electrode reduction on pattern detection.
- Effect of electrode reduction on pattern detection sensitivity was evaluated by a computer algorithm.
- Guidance which reduced EEG array may offer the highest detection results in specific situations.

ABSTRACT

Objective: To investigate the effect of systematic electrode reduction from a common 10-20 EEG system on pattern detection sensitivity (SEN).
Methods: Two reviewers rated 17130 one-minute segments of 83 prospectively recorded cEEGs according to the ACNS standardized critical care EEG terminology (CCET), including burst suppression patterns (BS) and unequivocal electrographic seizures. Consensus annotations between reviewers were used as a gold standard to determine pattern detection SEN and specificity (SPE) of a computational algorithm (baseline, 19 electrodes). Electrodes were than reduced one by one in four different variations. SENs and SPEs were calculated to determine the most beneficial assembly with respect to the number and location of electrodes.
Results: High automated baseline SENs (84.99–93.39%) and SPEs (90.05–95.6%) were achieved for all patterns. Best overall results in detecting BS and CCET patterns were found using the "hairline + vertex" montage. While the "forehead + behind ear" montage showed an advantage in detecting ictal patterns, reaching a 15% drop of SEN with 10 electrodes, all montages could detect BS sufficiently if at least nine electrodes were available.
Conclusion: For the first time an automated approach was used to systematically evaluate the effect of electrode reduction on pattern detection SEN in cEEG.
Significance: Prediction of the expected detection SEN of specific EEG patterns with reduced EEG montages in ICU patients.

Abbreviations: ACNS, American clinical neurophysiology society; BAM, "Banana" montage; BS, burst suppression patterns; CCET, American clinical neurophysiology society standardized critical care EEG terminology; cEEG, continuous electroencephalography; CRM, "Crown" montage; D15, drop of detection sensitivity of more than 15%; EEG, electroencephalography; FOM, "Forehead + behind ear" montage; HAM, "Hairline + vertex" montage; ICU,, intensive care unit; NCS,, nonconvulsive seizures; NOPAT, no pattern; PD, periodic discharge; RAA, rhythmic alpha activity; RDA, rhythmic delta activity; RTA, rhythmic theta activity; SEN, sensitivity; SPE, specificity.

* Corresponding author at: Medical University of Vienna, Department of Neurosurgery, Währinger Gürtel 18-20, 1090 Vienna, Austria. Fax: +43 01 40400 45660.
E-mail address: johannes.herta@meduniwien.ac.at (J. Herta).

http://dx.doi.org/10.1016/j.clinph.2017.04.012
1388-2457/

1. Introduction

Continuous EEG (cEEG) allows noninvasive monitoring of brain function with a high temporal resolution. Especially in the intensive care unit (ICU) it can add important information where conclusions from clinical examination may often be limited. For many applications, such as the detection of nonconvulsive seizures (NCS), the guidance of seizure treatment and the management of pharmacological induced coma, cEEG is considered the primary diagnostic tool (Jordan, 1999; Friedman et al., 2009). But even with an increased awareness of seizures in the ICU and huge advancements in computer technology, the use of EEG remains limited in everyday clinical practice. This is mainly due to the significant efforts associated with EEG. Besides the negligible costs of the recording device, personnel resources represent the major limiting factor. On the one hand, specially trained, 24-h available physicians are needed to review several hours of EEG. On the other hand, EEG technician must attach and maintain the electrode setup. In an ICU setting a trained EEG technician needs about 30–45 min to setup 19 cup electrodes. But collodion will dry out within the first six hours and needs accurate maintenance (Young et al., 2006). To increase availability and simplify the EEG setup, several studies assessed the possibility to work with a reduced number of electrodes (Bridgers and Ebersole, 1988; Foldvary et al., 2000; Tekgul et al., 2005; Kolls and Husain, 2007; Shellhaas and Clancy, 2007; Wusthoff et al., 2009; Young et al., 2009; Karakis et al., 2010; Nitzschke et al., 2011; Rubin et al., 2014; Tanner et al., 2014; Brenner et al., 2015; Lepola et al., 2015; Muraja-Murro et al., 2015).

A reduced electrode setup may have more potential benefits than just time saving. It can come in handy for patients where proper lead placement due to head wounds or drains is not possible. Furthermore, it may encourage physicians to use cEEG more frequently and consolidate acceptance among nursing staff. Previous studies reported frequent delays in the diagnosis of NCS (Dunne et al., 1987). Since mortality increases with seizure duration (Young et al., 1996) a reduced and easy applicable electrode setup should facilitate prompt diagnosis of NCS and benefit critical care patients.

Until now various approaches of electrode reduction have been published, that can be roughly summarized into three groups. Group-one tried to use a single-channel EEG (e.g. C3, C4). This was mainly used in neonates where most of seizures originate from the central midline (Schultz et al., 1992; Shellhaas and Clancy, 2007; Wusthoff et al., 2009). Group-two tried to cover as much of the scalp as possible, maintaining the 10–20 system based locations of electrodes (e.g. F3, F4, T7, Cz, T8, O1, O2) (Foldvary et al., 2000; Tekgul et al., 2005; Kolls and Husain, 2007; Karakis et al., 2010; Rubin et al., 2014; Lepola et al., 2015). Group-three's main interest was to develop an electrode setup which was easy to use and fast to apply in emergency cases (Bridgers and Ebersole, 1988; Young et al., 2009; Brenner et al., 2015; Muraja-Murro et al., 2015). In this setting it should be possible to place electrodes, without the help of an EEG technician, under the hairline on the forehead and behind the ear (e.g. Fp2, Fp1, F8, F7, Sp1, Sp2, T9, T10). Concerning seizure detection, nearly all major studies showed a tendency towards poor sensitivity (SEN). The common denominator of all these studies was to predefine a reduced electrode setup and compare its seizure detection rates with that of a standard 10–20 system.

In the present study, we reduced the electrodes of the International 10–20 EEG system systematically one by one, which to the best of our knowledge has never been done before. A computational algorithm assessed each reduction step. Four different variations of final electrode arrays, mainly derived from previously published reduced EEG montages were evaluated. Detection sensitivities (SEN) and specificities (SPE) for unequivocal electrographic seizures (spike-wave > 3 Hz, evolving discharges > 4 Hz), patterns defined by the ACNS Standardized Critical Care EEG Terminology (CCET) and burst suppression patterns (BS) were calculated (Hirsch et al., 2013). The aim of the study was to observe and illustrate the change in detection SEN and SPE for every reduced electrode and pattern of interest, to allow an individual assessment in cases where reduced setups are needed.

2. Methods

2.1. Dataset

A dataset of 92 prospectively recorded cEEGs in a neurological and a neurosurgical ICU (Neurological Center Rosenhuegel, General Hospital Vienna) was used. EEGs were recorded with a Micromed EEG recording system (SystemPLUS Evolution 1.04.95, Micromed S.p.A., Veneto, Italy) using the International 10–20 electrode system with a sampling rate of 256 Hz. Inclusion criteria for this study were 1) recordings longer than 24 h and 2) artefact-free recordings from a full set of 19 electrodes for more than 90% of the overall recording time. 7 EEGs were recorded with less than 19 electrodes. Another 2 patients had a recording time under 24 h. This left 83 patients for the study (6733 h, mean individual recording duration 73 h). Two types of electrodes were used for recordings: gold cup electrodes (Genuine Grass Gold Disc electrodes) and conductive plastic cup electrodes (Ives EEG Solutions). Research was prior approved by the institutional ethics committee.

2.2. NeuroTrend

NeuroTrend is a computational method that facilitates screening of long-term EEGs. It automatically detects rhythmic and periodic patterns in surface EEG and displays their localization and frequency in a graphical user interface. Results are visualized with a focus on data and time compression. Therefore, hours of cEEG can be compressed and displayed on a single screen. The definition of rhythmic and periodic EEG patterns follows the guidelines of CCET adding unequivocal electrographic seizures including generalized spike-wave discharges at 3 Hz or faster as well as evolving discharges that reach frequencies of more than 4 Hz and BS (Hirsch et al., 2013). Fürbass et al. (Fürbass et al., 2015) described the technical background of the algorithm, while Herta et al. (Herta et al., 2015) recently performed a validation of NeuroTrend. For this study a newer version of the algorithm was used. Especially RDA, which showed a high rate of false positive detections due to general slowing in the past, improved in terms of detection SEN and SPE as seen in Table 1. NeuroTrend is part of the encevis software package, in this work version V1.3 of encevis was used (http://www.encevis.com).

2.3. Data processing and statistical methods

The first minute of each hour of the raw cEEG recordings were identified and reviewed by two clinical neurophysiologists. In these segments the reviewers could assign one of four possible labels (1) periodic discharge (PD), (2) rhythmic delta activity (RDA), (3) ictal group (4) burst suppression patterns (BS). In each one-minute EEG segment multiple annotations could be made if they occurred consecutively. If no annotation was made the specific segment was labeled no pattern (NOPAT). Periodic and rhythmic delta patterns were rated according to the CCET guidelines. The ictal group included unequivocal electrographic seizures including generalized spike-and-wave discharges at 3 Hz or faster as well as evolving discharges that reach frequencies of more than 4 Hz.

Table 1
Detection performance of NeuroTrend. Two reviewers rated multiple segments of cEEG. Interrater agreement between the reviewers was calculated by Cohen's kappa (κ) statistics. Agreements were used as consensus annotations and compared to the detection results of the computer algorithm "NeuroTrend". Corresponding interrater agreement between the algorithm and the reviewers as well as detection sensitivity and specificity of the algorithm are shown. As a baseline calculation, a standardized 10-20 EEG system with 19 electrodes was used.

Category	n Annotation Segments	Segment length [sec]	n Agreements between both Reviewers	Sensitivity [%]	Specificity [%]	κ between Reviewers	κ between Reviewers & NeuroTrend
PD	17130	20	1305	87.36	90.05		
RDA	17130	20	121	93.39	91.87		
Ictal Group SW (n = 20) RTA* (n = 125) RAA* (n = 7)	17130	20	152	90.07	95.60	0.75	0.67
BS	5710	60	653	84.99	91.52	0.71	0.64

BS, burst suppression; κ, Cohens Kappa; PD, periodic discharges; RAA, rhythmic alpha activity; RDA, rhythmic delta activity; RTA, rhythmic delta activity; SW, spike-wave.
[a]Confirmed electrographic seizure activity in theta or alpha range.

Because BS typically lasted for longer periods, the whole one-minute segment was annotated either as a segment with or without BS. All other patterns could only be present for a few seconds. Therefore, annotations of these patterns were split into three non-overlapping 20-s segments.

Cohen's kappa statistic was used to calculate an interrater agreement. All segments that showed agreement between the two reviewers were considered as consensus annotations and used for further analysis. 10–20 system based cEEGs with 19 electrodes (excluding reference and ground electrode) were analyzed by the computer algorithm NeuroTrend. Consensus annotations were compared with the results of NeuroTrend. Detection performance of NeuroTrend was assessed by assigning one of four possible results to each annotation: True positive, false positive, true negative and false negative. A pattern was counted as true positive if one of the patterns detected by NeuroTrend in the annotation segment matched the consensus annotations of the reviewers. A consensus annotation without a matching NeuroTrend detection in the annotation minute was counted as false negative. An annotation segment with one or several NeuroTrend detections that did not match the consensus annotations was counted as false positive. An annotation segment without consensus annotation and without any detected pattern by NeuroTrend was counted as true negative. SEN and SPE were calculated according to the following formulas:

$$SEN\ [\%] = \frac{True\ Positive}{True\ Positive + False\ Negative} * 100$$

$$SPE\ [\%] = \frac{True\ Negative}{True\ Negative + False\ Positive} * 100$$

2.4. Electrode reduction

After having established an automated baseline for SEN and SPE using all 19 leads according to the International 10-20 EEG system (Fp1, F3, C3, P3, O1, Fp2, F4, C4, P4, O2, F7, T7, P7, F8, T8, P8, Fz, Cz, Pz), electrodes were reduced in a stepwise fashion. SEN, SPE and their confidence intervals were calculated separately for each electrode eliminated from the setup. Four variations of electrode reductions, depending on their local distribution, were used and labeled "forehead + behind ear montage" (FOM), "hairline + vertex montage" (HAM), "banana montage" (BAM) and "crown montage" (CRM). 12 to 13 reduction steps were calculated leaving six or seven electrodes for final calculations. Below, the steps of electrode reduction are shown as superscript numbers. The final EEG montages are shown in bold (Fig. 1):

Forehead + behind ear montage (FOM, Fig. 1A):

$O1^{1}$-$O2^{2}$-$P3^{3}$-$P4^{4}$-Pz^{5}-$T7^{6}$-$T8^{7}$-$C3^{8}$-$C4^{9}$-Cz^{10}-$F3^{11}$-$F4^{12}$ - **Fp1-Fp2-F7-Fz-F8-P7-P8[13]**

Hairline + vertex montage (HAM, Fig. 1B):

$P7^{1}$-$P8^{2}$-$F7^{3}$-$F8^{4}$-$P3^{5}$-$P4^{6}$-$F3^{7}$-$F4^{8}$-Fz^{9}-Pz^{10}-$C3^{11}$-$C4^{12}$ - **Fp1-Fp2-T7-Cz-T8-O1-O2[13]**

Banana (= Longitudinal) montage (BAM, Fig. 1C):

Cz^{1}-Pz^{2}-Fz^{3}-$P7^{4}$-$P8^{5}$-$F7^{6}$-$F8^{7}$-$T7^{8}$-$T8^{9}$-$P3^{10}$-$P4^{11}$-$F3^{12}$-$F4^{13}$ - **Fp1-Fp2-C3-C4-O1-O2[14]**

Crown (= Transversal) montage (CRM, Fig. 1D):

Cz^{1}-$O1^{2}$-$O2^{3}$-$Fp1^{4}$-$Fp2^{5}$-$C3^{6}$-$C4^{7}$-$T7^{8}$-$T8^{9}$-$P3^{10}$-$P4^{11}$-$F3^{12}$-$F4^{13}$ - **F7-Fz-F8-P7-Pz-P8[14]**

The detection results based on the EEGs of each reduction step were compared to the consensus annotations of the reviewers. SEN and SPE were quantified and the number of electrodes for which detection SEN dropped more than 15% (D15) was determined.

2.5. Validation of computational results

To validate the computational NeuroTrend results a single reduced EEG dataset was annotated a second time by the two reviewers. For reevaluation, we chose the montage that achieved a D15 with the least number of electrodes for every evaluated pattern. Furthermore, the number of electrodes was reduced to the half (reduction step 10, 9 electrodes). For this reduced montage 50 EEG segments from each of the four pattern groups and 50 EEG segments without patterns were randomly selected resulting in 250 EEG segments.

The same reviewers who established the primary consensus annotations annotated the 250 segments again. They were blinded to the distribution of patterns. The reduced montages were presented to the reviewers as short 20-s EEG segments. Switching between longitudinal, transversal and referential montages was allowed during the review process.

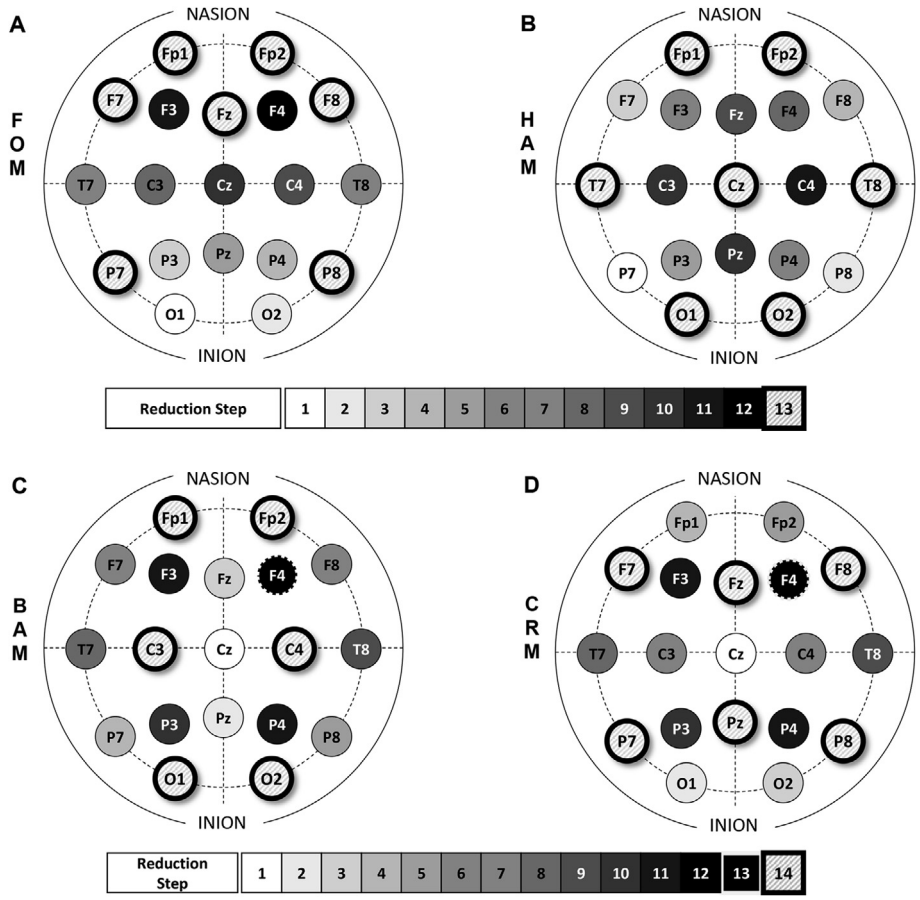

Fig. 1. Four different variations of electrode reductions from a common 10-20 EEG montage are shown. The different shades of gray (Reduction Step) numbers the succession of lead reduction starting with 'step 1' in white. The final electrode array is shown striped and boldly encircled ('step 13' or 'step 14'), labeling the different final montages. (A) A "forehead + behind ear" EEG montage (FOM), also known as "subhairline" montage, is shown. Because it is quickly installed and easy to use, it is commonly used in the emergency department. (B) The "hairline + vertex" montage (HAM) tries to cover the whole scalp but leads to double distances between electrodes in the final reduced montage. (C, D) A longitudinal "banana" montage (BAM) as well as a transversal "crown" montage (CRM) is shown. They were thought to be advantageous in detecting patterns with different local distributions.

The review results of the 250 EEG samples with the reduced 9 electrodes were compared to the primary consensus annotations with the full electrode setup to determine if samples were annotated equally. SEN of these annotations was calculated for each reviewer to quantify the loss of SEN and agreement.

Then consensus annotations between the two reviewers based on the reduced nine electrodes EEG samples were determined and classified as correct or incorrect by using the primary consensus annotations. The computational result was evaluated on the same samples to define correct or incorrect detections. The number of samples with correct annotations and incorrect computer result are defined as c. The number of samples with incorrect annotations and correct computer result are defined as b. To prove that no statistically significant difference between human and computational annotations exist the McNemar test with the test statistic $\chi^2 = \frac{(|b-c|-1)^2}{b+c}$ and a critical value of 3.841 for $\alpha = 0.95$ was used on this paired nominal data.

3. Results

3.1. Detection performance of NeuroTrend

17,130 20-s annotations showed agreement between the two reviewers and were considered as consensus annotations. In these 17,130 segments 1578 rhythmic and periodic EEG patterns were found and compared with the detection results of the computer algorithm. Baseline detection SENs and SPEs of NeuroTrend were calculated with a full set of 19 electrodes. Table 1 illustrates the consensus annotations found for different pattern groups as well as the detection performance of NeuroTrend. For BS 5710 60-s annotations showed agreement between the reviewers. 653 BS were found and compared to the computer algorithm (Table 1).

3.2. Electrode reduction

For most pattern categories and reduction montages, a reduction of electrodes caused a continuous decline in SEN, while SPE increased as illustrated in Fig. 2. Table 2 shows the number of electrodes used for which a D15 occurred in different pattern types and corresponding montages.

3.2.1. Periodic discharge (PD)

We detected a stable decrease in SEN for PD, no matter which electrode reduction montage was used (Fig. 2B). PDs occurred with 58.52% disproportionately often considering the distribution of all pattern groups. D15 was encountered in the HAM (SEN: 76.32%, SPE: 92.61%) with 13 electrodes, which was the best result compared to all other montages. A stable decline in SEN occurred until the 9th electrode was removed (10 electrodes remaining), a rapid decrease was observed thereafter. After the final reduction step, very poor SENs were observed, ranging from 42.76% to 53.26%.

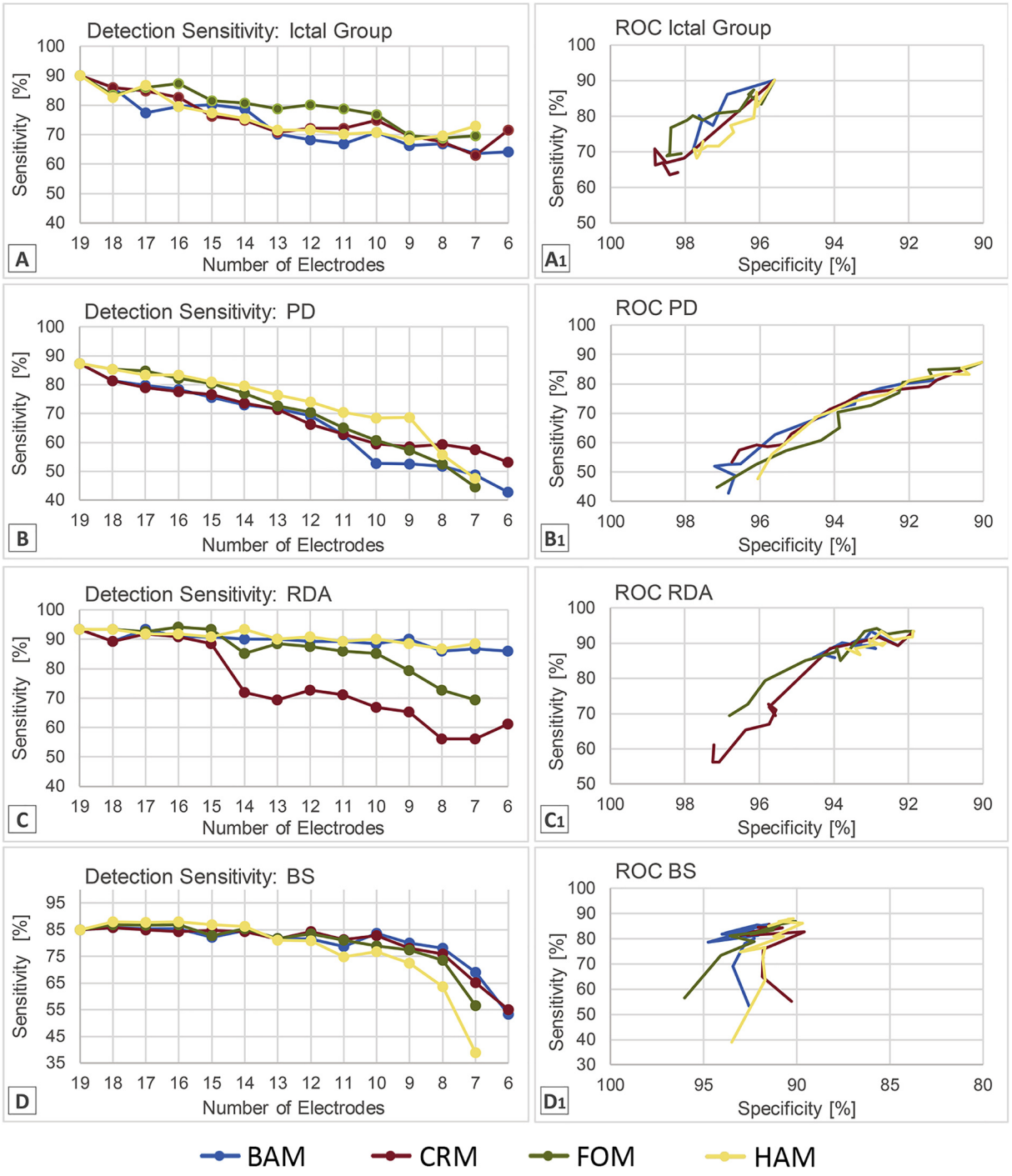

Fig. 2. Changes in detection sensitivity (left) and specificity (right) of NeuroTrend with a decreasing number of electrodes are illustrated for different pattern types and different reduction montages. BAM, banana montage; CRM, crown montage; FOM, forehead + behind ear montage; HAM, headband + vertex montage; PD, periodic discharge; RDA, rhythmic delta activity; Ictal group, spike-wave >3 Hz or evolving discharges >4 Hz; BS, burst suppression.

3.2.2. Ictal group (RTA, RAA, SW)

In the ictal group the overall best performing reduction montage was FOM with a D15 at 10 remaining electrodes (SEN: 76.82%, SPE: 98.38%). Further reduction of electrodes caused a slight but not explainable detection increase in some montages with final SENs between 64.24% and 72.85% (Fig. 2A).

3.2.3. Rhythmic delta activity (RDA)

RDA was the only group where different reduction montages diverged strongly. HAM and BAM nearly showed no decrease of SEN and even did not reach a D15 with the last reduction step. With the final array, HAM showed the best detection SEN of 88.43% (SPE: 93.67%). On the contrary FOM reached a D15 with 10 remaining electrodes and CRM with 15, respectively. Low SENs of 69.42% and 61.16% made these two montages unsuitable for the detection of RDA with only 6 to 7 leads (Fig. 2C).

3.2.4. Burst suppression patterns (BS)

In all reduction montages BS showed a uniform decline after a D15 with eight to nine electrodes remaining. The best montage

Table 2
Number of electrodes for which detection sensitivity dropped more than 15%.

Category	Item	"Banana" montage (BAM)	"Crown" montage (CRM)	"Forehead + behind ear" montage (FOM)	"Hairline + vertex" montage (HAM)
PD (periodic discharge)	Number of electrodes	15	15	14	13
	SEN (%)	75.56	76.70	77.01	76.32
	SPE (%)	93.40	93.26	92.28	92.61
Ictal Group (electrographic seizures)	Number of electrodes	15	16	10	15
	SEN (%)	80.13	82.78	76.82	77.48
	SPE (%)	97.63	97.45	98.38	96.78
RDA (rhythmic delta activity)	Number of electrodes	6	15	10	6
	SEN (%)	85.95	88.43	85.12	88.43
	SPE (%)	93.97	94.10	94.76	93.67
BS (burst suppression)	Number of electrodes	8	8	8	9
	SEN (%)	78.10	75.96	73.51	72.43
	SPE (%)	92.79	91.79	94.09	91.82

SEN, sensitivity; SPE, specificity.

at D15 was BAM with a SEN of 78.10% and a SPE of 92.27% with 8 leads after which an exponential decline occurred (Fig. 2D). The final array with BAM showed a SEN of 53.29% and a SPE of 92.55%, which was slightly inferior to the final array of FOM, which revealed a SEN of 56.51% and a SPE of 96.03%.

3.3. Validation of computational results

The overall most sensitive montage at reduction step 10 (nine remaining electrodes) was HAM with a SEN of 76.12% and a SPE of 87.58%. At this reduction step, the computer algorithm calculated a SEN of 68.21% for the ictal group, 88.43% for RDA, 68.74% for PD and 72.44% for BS. Corresponding SPE were high with 97.76 for the ictal group, 92.62 for RDA, 94.52 for PD and 91.69 for BS. Results with confidence intervals are shown in Fig. 3 (NeuroTrend; 9 electrodes). With the same reduced set of nine electrodes, 250 EEG segments were reevaluated by the two reviewers to validate the calculations of the computer algorithm.

In the ictal group the two reviewers reached detection SENs of 82% (SPE 94%) and 78% (SPE 98%), respectively. RDA could be detected with SENs of 84% (SPE 90% & 100%) each. Lower agreements were seen for PD with SENs of 82% (SPE 100%) and 64% (SPE 100%) as well as BS with SENs of 80% (SPE 94%) and 56% (SPE 98%). The detection SEN of NT calculated with nine electrodes compared to 19 electrodes declined between a range of 5–22% for different pattern groups. For the reviewers, comparing annotations with 9 electrodes with the consensus annotations, a decline between 16–44% was found.

The elimination of EEG segments without consensus annotations from the two reviewers resulted in overall $n = 156$ EEG segments that were used for the test statistic. The initial 50 samples of each pattern group reduced to $n_{PD} = 38$, $n_{RDA} = 42$, $n_{ictal\ group} = 38$, $n_{BS} = 38$ samples. The hypothesis that no statistically significant difference between the consensus annotations of two reviewers and the computational results existed could not be rejected for all four subgroups ($\chi^2_{PD} = 0.57$, $\chi^2_{RDA} = 0.5$, $\chi^2_{ictal\ group} = 3.2$, $\chi^2_{BS} = 1.2$) and for the combined 156 samples ($\chi^2 = 0.41$). This shows that the average detection SEN based on computational and human review has to be considered as equal.

4. Discussion

The present study aims to investigate the effect of electrode reductions from a standard 10-20 EEG system. Unlike previous studies this was done automatically by a computer algorithm, making it feasible (1) to determine detection SENs for every single electrode that was removed (2) to observe the effect of different sequences in which electrodes were removed. This automated and technical approach of analyzing the effect of decremental electrode reduction on pattern detection, distinguishes the study from others. The vast majority of previous studies used more clinically orientated approaches to determine if a certain number of predefined electrodes were sufficient to detect seizures (Bridgers and Ebersole, 1988; Foldvary et al., 2000; Tekgul et al., 2005; Kolls and Husain, 2007; Shellhaas and Clancy, 2007; Wusthoff et al., 2009; Young et al., 2009; Karakis et al., 2010; Nitzschke et al., 2011; Rubin et al., 2014; Tanner et al., 2014; Brenner et al., 2015; Lepola et al., 2015; Muraja-Murro et al., 2015). These studies only gave a brief insight into a small selection of existing possibilities because they missed the flexibility to change montages or add and remove electrodes. Furthermore, their main goal was to detect seizures, while the effect of electrode reduction on other patterns was not investigated. Our study on the other hand should be seen as a proof of concept. We tried to demonstrate that automation is a feasible and reasonable method to asses reduced electrode arrays, taking also patterns defined by CCET into account. The results may be expected but have never been accurately illustrated. Most of the previously published studies used between four to ten electrodes. According to our data a clear decline in pattern detection begins after the 10th electrode is removed. Therefore, our results form a foundation for further, more clinically oriented studies.

HAM outperformed all other reduction montages by reaching a D15 with the lowest number of electrodes. Best results could be achieved with six electrodes for ictal group patterns and RDA as well as nine electrodes for PD. HAM is easy and fast to apply because anatomic landmarks can be used to estimate correct electrode placement. Important drawback of the montage is the poor performance in detecting BS, which could be used to monitor treatment effects and estimate sedation depth in the ICU. Detection rates (SEN 72.85%, SPE 97.51%) of HAM match previous studies that used similar reduced arrays. Rubin et al. reviewed 50 ictal and 50 non-ictal EEG records for the presence or absence of seizures (Rubin et al., 2014). They used the electrodes F3, F4, T7, Cz, T8, O1, O2 and reviewed the EEG with transverse, longitudinal and referential to Cz montages. A detection SEN of 70% and SPE of 96% for seizures was found. They concluded that this was an unacceptable poor SEN for seizure detection. The same was suggested by Kolls and Hussain after the review of 120 preselected "clear" pattern clips by five epileptologists (Kolls and Husain, 2007). A six-channel montage including the electrodes Fp1, Fp2, F7, F8, T3, T4, T5, T6 (longitudinal bipolar, referential to ipsilateral ear, referential to contralateral ear) was used. Reviews were compared to medical records and showed SEN rates of 72% for seizures and 54% for PDs. Higher detection rates were shown by Karakis et al. (Karakis et al., 2010). They reviewed 38 preselected EEG samples, including only

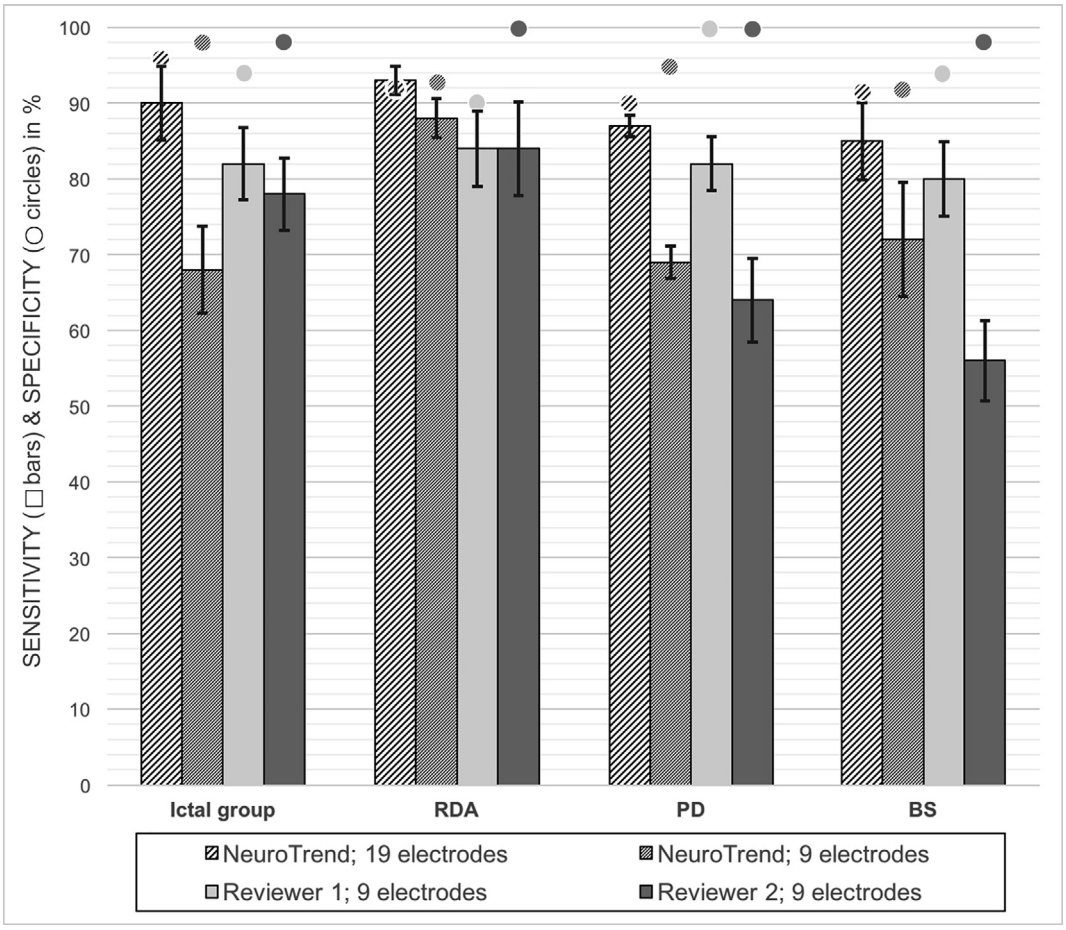

Fig. 3. Detection sensitivities (SEN, bars) and specificities (SPE, circles) for different pattern groups are shown. Bars & circles with oblique stripes illustrate SENs/SPEs of the computer algorithm (NeuroTrend) for 19 and 9 electrodes, respectively. Filled bars and circles illustrate SENs/SPEs of two different reviewers for 9 electrodes. PD, periodic discharge; RDA, rhythmic delta activity; Ictal group, spike-wave >3 Hz or evolving discharges >4 Hz; BS, burst suppression.

10 samples with seizures. A seizure detection rate of 85% was found with a six-channel EEG (Fp1, Fp2, T3, T4, O1, O2, Cz; double diamond, circumferential, referential to Cz) compared to 92.5% for the 10-20 EEG. The difference in the detection SEN of approximately 10–15% compared to our data could be explained by the general lower detection rate of the computer algorithm compared to consensus annotations in our study.

FOM has found application in the emergency department and predefined electrode bands that adhere to the skin of the forehead and behind the ear are already available (Myllymaa et al., 2013; Muraja-Murro et al., 2015). In our study P7 and P8 had to replace the electrodes behind the ear. The overall performance of FOM was mediocre with a D15 ranging between 8 to 14 electrodes depending on the pattern type observed. Nevertheless, there may be a potential use for the FOM in the emergency department as detection SEN was high in the ictal group before reducing the 10th electrode (SEN of 76.82%). Similar detection rates were shown by Young et al. (Young et al., 2009). Two epileptologists reviewed 70 cEEGs of 24 h with a standard 10–20 system as well as with a reduced array of four frontal channels. 31 patients suffered from seizures which were detected in 68% of all cases by the reduced montage. PDs showed lower rates of 39%, which can be confirmed by our observation (44.67%). A lower SEN was found by Tanner et al. (Tanner et al., 2014) who retrospectively reviewed 170 patients of which 8% had seizures. They found a seizure detection rate of 54% with a reduced setup of seven to eight leads. Contradictory findings were presented by Bridgers and Ebersole (Bridgers and Ebersole, 1988). They performed an interrater agreement assessing 25 patients with epileptiform abnormalities. One epileptologist reviewed 16 channel EEG data while the other epileptologist had only seven channels available. 91% of all epileptiform complexes were detected by reviewing seven channels with a false positive rate of 10% and a false negative rate of 8%. Other studies that investigated the influence of FOM on detection rates were affected by small numbers of evaluated patients or a low incidence of seizures (Brenner et al., 2015; Lepola et al., 2015; Muraja-Murro et al., 2015).

CRM and BAM were not previously described as reduced montages in literature. They were thought to offer advantages in detecting strictly localized patterns such as PD. This hypothesis was not met as both arrays showed a poor performance for PDs and were mediocre in detecting patters of the ictal group. Interestingly both montages scored highest in detecting BS while diverging in the detection of RDA. CRM showed a steep decline in SEN after frontopolar and occipital electrodes were removed. This may be explained by over-interpretation of RDA in the consensus annotations if generalized frontal slowing occurred. When interpreting the raw EEG, double distances between electrodes must be considered as they strongly influence the EEG curve. The used computer algorithm could handle these double distances by calculating results with a common average of remaining electrodes even though it is based on visual detection of EEG data.

Validation of NeuroTrend results showed that the computer algorithm scored a little worse than the reviewers except for RDA but no significant discrepancies could be observed. This time, no consensus annotations between the reviewers were established,

as this would have biased the validation by leaving only clear and easy to recognize patterns. On the one hand, persistent low detection rates of the algorithm in comparison to the reviewers would have implied an unusable poor algorithm. On the other hand, persistent high detection rates compared to the reviewers might have indicated an implausible result since the algorithm is based on visual analyzes. Results between the two reviewers varied a lot for PDs and BS indicating difficulties in annotating these patterns with a reduced number of electrodes.

The strength of the study, to keep the focus on clinical relevant, non-selected data comprises some limitations. Because of the huge amount of work in annotating hours of cEEG to establish consensus annotations, a limited number of patients (n = 83) was enrolled. "Real-world" conditions immanent are numerous EEG segments with no specific patterns and an unequal distribution of patterns (Table 1).

The study lacks information about pattern localization because the algorithm showed a low performance in distinguishing lateralized from generalized patterns in a previous study (Herta et al., 2015). This limits the statement about the advantages and disadvantages of the individual assemblies. Furthermore, it must be stressed that all observed patterns frequently occur together during cEEG in the ICU. It would be misguided to assume that in a given patient a certain montage may be superior to another montage for clinical monitoring purposes based upon these results. For example, if seizure detection is the primary goal, not only ictal group patterns but also PD and RDA may classify as seizures and in the further course of treatment the detection of BS may become of interest.

The ACNS does not recommend the use of less than 19 electrodes as well as deviations of the International 10-20 system placement but recognizes the need of a smaller number of electrodes in some situations (ACNS, 1994). Hence it is very difficult to give recommendations for the use of a reduced EEG array, as seizures may not be detected at all if standard EEG is not available or not applicable as argued by Young et al. (Young et al., 2009).

Our aim was to demonstrate a new approach of testing the usability of reduced EEG montages. Clear advantages of an automated assessment comprise the possibility of rapidly processing huge amounts of data, clear visualization, exact determination of frequencies and amplitudes as well as identification of pattern localization. Abilities that may not only find application in research and science but also in clinical practice.

5. Conclusion

For the first time a computer algorithm was successfully used to evaluate the effect of decremental electrode reduction from the international 10-20 EEG system. The findings roughly reflect which reduced assembly may be the most appropriate in specific situations where a full 10-20 EEG system cannot be applied. However, studies on how the reduced montages perform in individual patients still have to be carried out. In the future, we expect more detailed and specific analyzes by our algorithm taking new variables as for example pattern localization into account.

Disclosure

Research and development of NeuroTrend was supported by The Austrian Research Promotion Agency (FFG) grant 826816 (EpiMon). Johannes Herta and Johannes Koren were both partially supported by the FFG grant.

Algorithm development was conducted by the "Austrian Institute of Technology" including the authors Franz Fürbass, Manfred Hartmann and Tilmann Kluge. The Austrian Institute of Technology is the manufacturer of the EEG software package "encevis", which will include the NeuroTrend algorithms.

The remaining authors have no conflicts of interest.

Acknowledgements

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

References

ACNS. Guideline one: minimum technical requirements for performing clinical electroencephalography. J Clin Neurophysiol 1994;11:2–5.
Brenner JM, Kent P, Wojcik SM, Grant W. Rapid diagnosis of nonconvulsive status epilepticus using reduced-lead electroencephalography. West J Emerg Med 2015;16:442–6.
Bridgers SL, Ebersole JS. EEG outside the hairline: detection of epileptiform abnormalities. Neurology 1988;38:146–9.
Dunne JW, Summers QA, Stewart-Wynne EG. Non-convulsive status epilepticus: a prospective study in an adult general hospital. Q J Med 1987;62:117–26.
Foldvary N, Caruso AC, Mascha E, Perry M, Klem G, McCarthy V, et al. Identifying montages that best detect electrographic seizure activity during polysomnography. Sleep 2000;23:221–9.
Friedman D, Claassen J, Hirsch LJ. Continuous electroencephalogram monitoring in the intensive care unit. Anesth Analg 2009;109:506–23.
Fürbass F, Ossenblok P, Hartmann M, Perko H, Skupch AM, Lindinger G, et al. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. Clin Neurophysiol 2015;126:1124–31.
Herta J, Koren J, Fürbass F, Hartmann M, Kluge T, Baumgartner C, et al. Prospective assessment and validation of rhythmic and periodic pattern detection in NeuroTrend: a new approach for screening continuous EEG in the intensive care unit. Epilepsy Behav 2015;49:273–9.
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.
Jordan KG. Nonconvulsive status epilepticus in acute brain injury. J Clin Neurophysiol 1999;16. 332–40- discussion 53.
Karakis I, Montouris GD, Otis JAD, Douglass LM, Jonas R, Velez-Ruiz N, et al. A quick and reliable EEG montage for the detection of seizures in the critical care setting. J Clin Neurophysiol 2010;27:100–5.
Kolls BJ, Husain AM. Assessment of hairline EEG as a screening tool for nonconvulsive status epilepticus. Epilepsia 2007;48:959–65.
Lepola P, Myllymaa S, Toyras J, Hukkanen T, Mervaala E, Maatta S, et al. A Handy EEG Electrode Set for patients suffering from altered mental state. J Clin Monit Comput 2015;29:697–705.
Muraja-Murro A, Mervaala E, Westeren-Punnonen S, Lepola P, Toyras J, Myllymaa S, et al. Forehead EEG electrode set versus full-head scalp EEG in 100 patients with altered mental state. Epilepsy Behav 2015;49:245–9.
Myllymaa S, Lepola P, Toyras J, Hukkanen T, Mervaala E, Lappalainen R, et al. New disposable forehead electrode set with excellent signal quality and imaging compatibility. J Neurosci Methods 2013;215:103–9.
Nitzschke R, Muller J, Engelhardt R, Schmidt GN. Single-channel amplitude integrated EEG recording for the identification of epileptic seizures by nonexpert physicians in the adult acute care setting. J Clin Monit Comput 2011;25:329–37.
Rubin MN, Jeffery OJ, Fugate JE, Britton JW, Cascino GD, Worrell GA, et al. Efficacy of a reduced electroencephalography electrode array for detection of seizures. Neurohospitalist 2014;4:6–8.
Schultz B, Bender R, Schultz A, Pichlmayr I. Reduction of the number of recorded EEG channels for routine monitoring in the intensive care unit. Biomed Tech (Berl) 1992;37:194–9.
Shellhaas RA, Clancy RR. Characterization of neonatal seizures by conventional EEG and single-channel EEG. Clin Neurophysiol 2007;118:2156–61.
Tanner AEJ, Sarkela MOK, Virtanen J, Viertio-Oja HE, Sharpe MD, Norton L, et al. Application of subhairline EEG montage in intensive care unit: comparison with full montage. J Clin Neurophysiol 2014;31:181–6.
Tekgul H, Bourgeois BFD, Gauvreau K, Bergin AM. Electroencephalography in neonatal seizures: comparison of a reduced and a full 10/20 montage. Pediatr Neurol 2005;32:155–61.
Wusthoff CJ, Shellhaas RA, Clancy RR. Limitations of single-channel EEG on the forehead for neonatal seizure detection. J Perinatol 2009;29:237–42.
Young GB, Ives JR, Chapman MG, Mirsattari SM. A comparison of subdermal wire electrodes with collodion-applied disk electrodes in long-term EEG recordings in ICU. Clin Neurophysiol 2006;117:1376–9.
Young GB, Jordan KG, Doig GS. An assessment of nonconvulsive seizures in the intensive care unit using continuous EEG monitoring: an investigation of variables associated with mortality. Neurology 1996;47:83–9.
Young GB, Sharpe MD, Savard M, Al Thenayan E, Norton L, Davies-Schinkel C. Seizure detection with a commercially available bedside EEG monitor and the subhairline montage. Neurocrit Care 2009;11:411–6.