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.
© 2016 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

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/© 2016 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

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

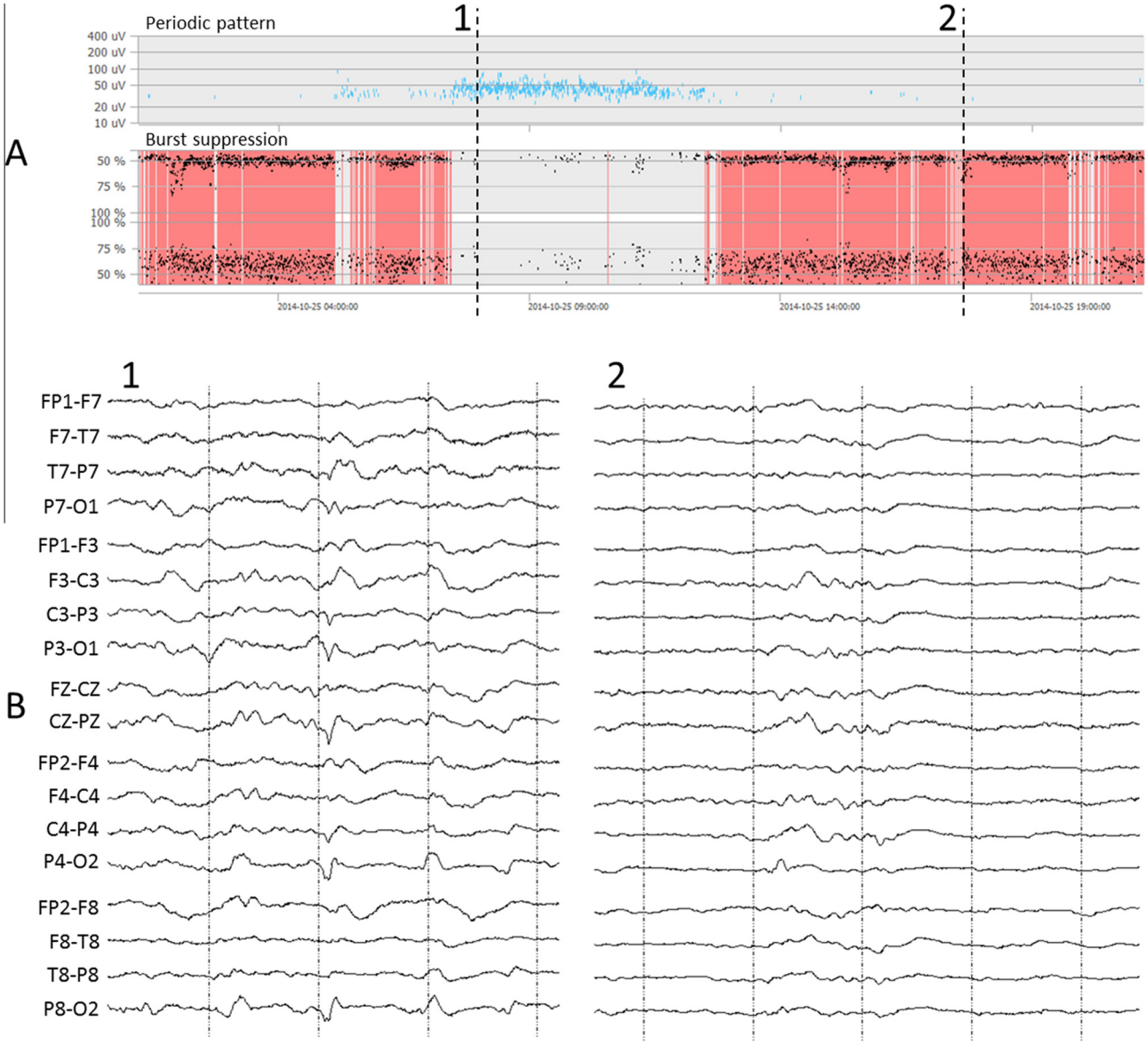

Fig. 1. Quantitative EEG interface (NeuroTrend, www.encevis.com) showing the detection results of a 20 h EEG recording registered in the intensive care unit. (A) The upper plot show periodic pattern detections (Fürbass et al., 2015a) which are continuously present for approx. 5 h in this patient. The 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.

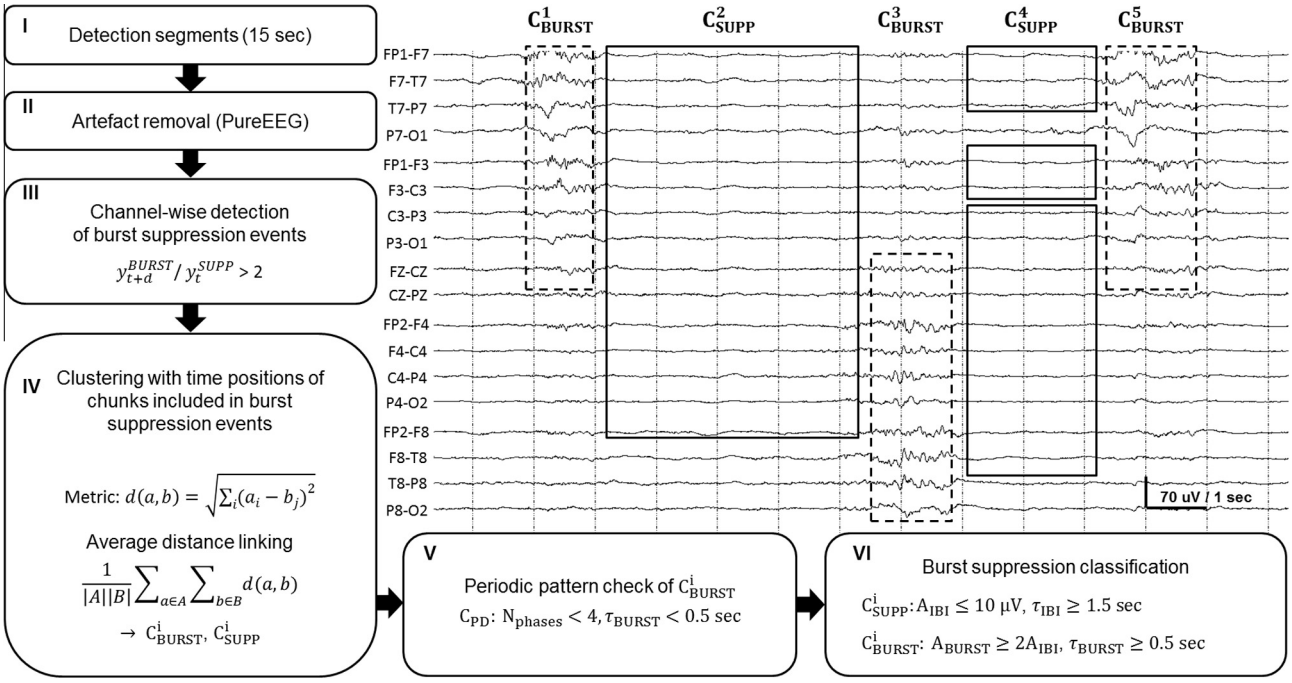

Fig. 2. Block diagram of the automatic burst suppression detection method with the resulting clusters visualized in an EEG segment. The burst and suppression clusters C_{BURST} (dashed line) and C_{SUPP} (solid line) do not overlap in time and may include several channels. The EEG sample shows a bihemispheric 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.

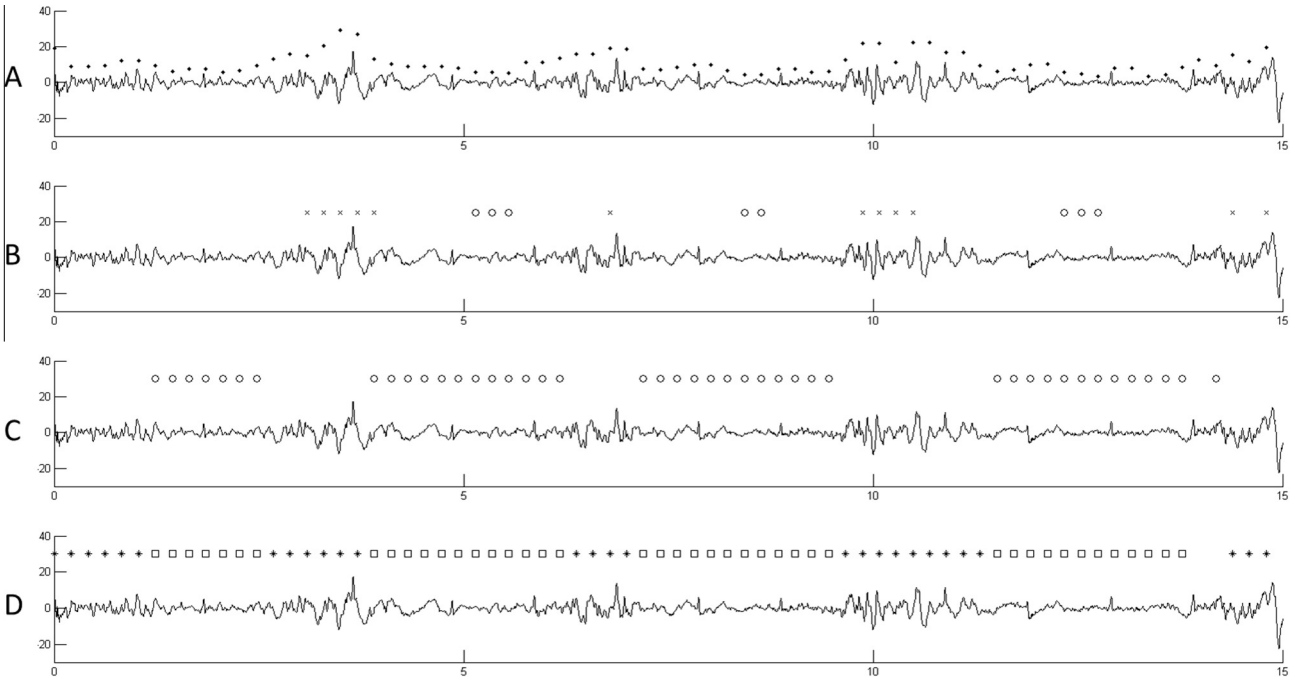

Fig. 3. Processing steps of the channel-wise burst suppression event detection. (A) The peak-to-peak amplitudes measured for each chunk of 0.4 s length are shown as dots. (B) Based on a smoothed time series of these chunk amplitudes each chunk with a preceding double amplitude chunk is marked as part 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)	

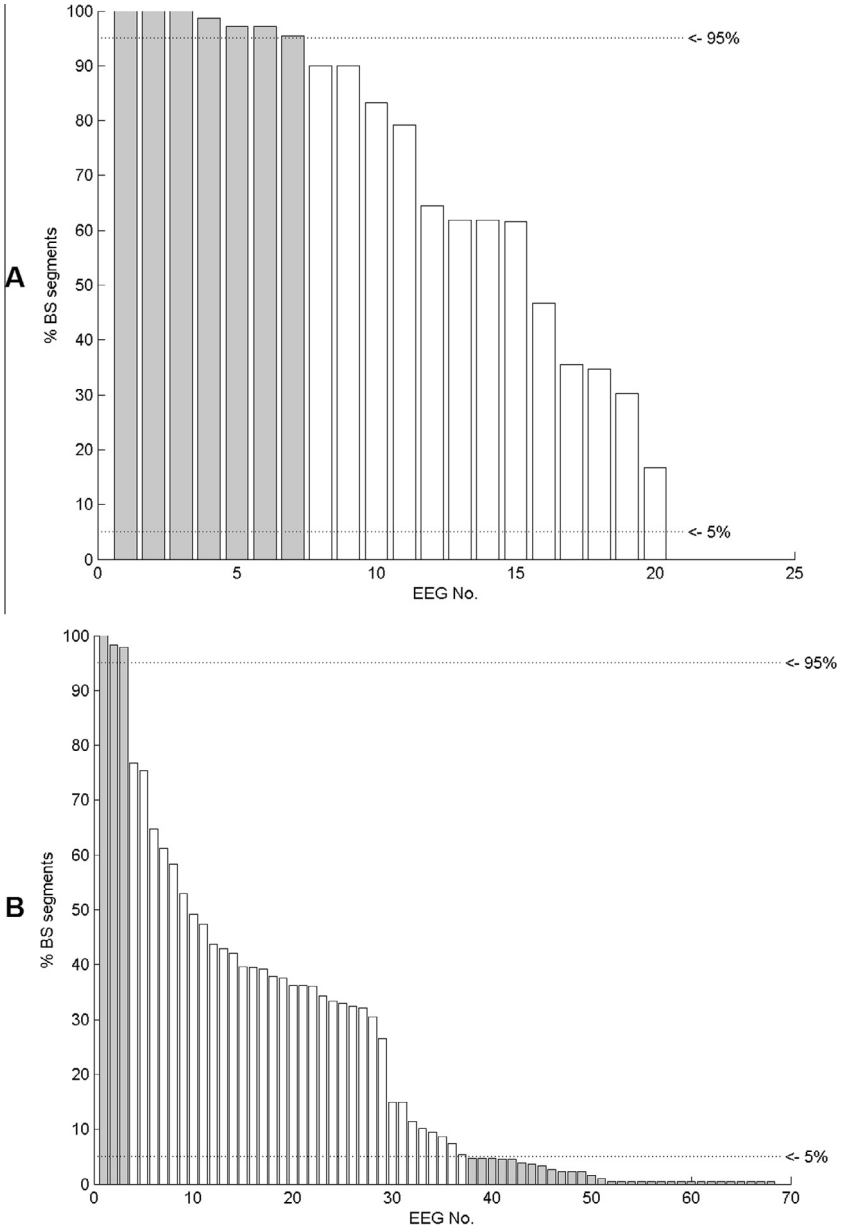

Fig. 5. Prevalence of segments with burst suppression patterns (BS) for each patient in dataset MGH (A) and dataset VIEN (B). Patients with less than 5% or more than 95% of burst suppression segments are marked in grey color.

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.