Published in final edited form as:
J Neurosci Methods. 2013 September 30; 219(1): 131–141. doi:10.1016/j.jneumeth.2013.07.003.

Real-time segmentation of burst suppression patterns in critical care EEG monitoring

M. Brandon Westover[a,*,1], **Mouhsin M. Shafi**[d,1], **ShiNung Ching**[b,1], **Jessica J. Chemali**[b,c], **Patrick L. Purdon**[b,c], **Sydney S. Cash**[a], and **Emery N. Brown**[b,c]

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

[b]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Boston, MA

[c]Department of Anaesthesia and Critical Care, Massachusetts General Hospital, Boston, MA

[d]Department of Neurology, Beth Israel Deaconess Medical Center, United States

Abstract

Objective—Develop a real-time algorithm to automatically discriminate suppressions from non-suppressions (bursts) in electroencephalograms of critically ill adult patients.

Methods—A real-time method for segmenting adult ICU EEG data into bursts and suppressions is presented based on thresholding local voltage variance. Results are validated against manual segmentations by two experienced human electroencephalographers. We compare inter-rater agreement between manual EEG segmentations by experts with inter-rater agreement between human vs automatic segmentations, and investigate the robustness of segmentation quality to variations in algorithm parameter settings. We further compare the results of using these segmentations as input for calculating the burst suppression probability (BSP), a continuous measure of depth-of-suppression.

Results—Automated segmentation was comparable to manual segmentation, i.e. algorithm-vs-human agreement was comparable to human-vs-human agreement, as judged by comparing raw EEG segmentations or the derived BSP signals. Results were robust to modest variations in algorithm parameter settings.

Conclusions—Our automated method satisfactorily segments burst suppression data across a wide range adult ICU EEG patterns. Performance is comparable to or exceeds that of manual segmentation by human electroencephalographers.

Significance—Automated segmentation of burst suppression EEG patterns is an essential component of quantitative brain activity monitoring in critically ill and anesthetized adults. The

[*]Corresponding author at: Wang 7 Neurology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, United States. Tel.: +1 617 726 3311; fax: +1 617 724 6513.
[1]These authors contributed equally to the work.

Authorship details

MBW, SC, MMS, SSC, PLP and ENB conceptualized and designed the study. MBW, MMS, SC and JJC analyzed the data, conducted the statistical analysis, and drafted the original manuscript. All authors reviewed and revised the manuscript. MBW, MMS and SSC contributed equally to the overall work.

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.jneumeth.2013.07.003.

segmentations produced by our algorithm provide a basis for accurate tracking of suppression depth.

Keywords

Burst suppression; Medically-induced coma; Quantitative EEG; ICU EEG monitoring

1. Introduction

Burst suppression is a brain state of profound brain inactivation and unconsciousness, accompanied by an EEG pattern consisting of periods of depressed or flat electrographic background activity alternating with periods of higher voltage activity (Swank and Watson, 1949; Amzica, 2009; Brenner, 2005), formally defined by The International Federation of Societies for Electroencephalography and Clinical Neurophysiology (IF-SECN) as a "pattern characterized by theta and/or delta waves, at times intermixed with faster waves, and intervening periods of relative quiescence." Originally described by Derbyshire in anesthetized cats (Derbyshire et al., 1936), and further described by Niedermeyer (Niedermeyer et al., 1999; Chatrian, 1974), burst suppression occurs in a range of clinical settings during the course of continuous clinical EEG monitoring. Burst suppression patterns occur spontaneously in the course of neonatal development (e.g. trace alternant or trace discontinua) as well as in a variety of neonatal brain injuries and in patients of all ages with diffuse anoxic brain injury (Niedermeyer, 2009; Cloostermans et al., 2012; Nunes et al., 2005). Burst suppression may also be induced by anesthetic drug administration, as during surgical anesthesia or to treat refractory status epilepticus (Doyle and Matta, 1999; Shorvon, 2011), or by controlled hypothermia, as is commonly done during complete circulatory in cardiac surgery (Woodcock et al., 1987).

In pharmacologically induced burst suppression the duration of spontaneous bursts and suppressions varies systematically with brain anesthetic concentrations, with higher concentrations leading to progressively longer suppressions, eventually causing continuous EEG suppression (Vijn and Sneyd, 1998; Young, 2000; Brenner, 2005; Ching et al., 2012). In pathological states such as anoxic brain injury, longer spontaneous suppressions correlate with worse prognosis, and trends in the evolution of spontaneous burst suppression patterns hold important prognostic information (Young et al., 2004; Hallberg et al., 2010; Dandan et al., 2010). The *burst suppression ratio* (BSR), a measure of the percentage of time within an interval spent in the suppressed state, is a widely used indicator of the intensity of suppression which medical personnel actively control for therapeutic purposes. For example, BSR is used to guide treatment of refractory status epilepticus, with treatment protocols typically targeting a burst suppression pattern of 5–15 s suppressions alternating with 1–5 s bursts or equivalently BSR values of 50–95% (Cottenceau et al., 2008; Riker et al., 2003; Rossetti et al., 2011; Musialowicz et al., 2010).

Despite the importance of this measure and recommendations for its use in a wide variety of clinical settings, current practice consists of visual examination of the EEG and depth of burst suppression is most often assessed as a 'guesstimate' rather than using any truly quantitative, reproducible approach. Furthermore, clinical monitoring of continuous EEG data is highly labor intensive. Therefore, in contexts where the level of burst suppression must be closely monitored and/or tightly controlled, an objective, reproducible, clinically validated method for continuously quantifying the depth of suppression would be beneficial. The first step in quantifying the depth of burst suppression is *segmentation*, i.e. performing a binary division of the EEG into segments of various lengths classified as either suppressions or bursts. While several previous publications have described methods of varying complexity for automated segmentation of EEG data into burst and suppression epochs (see

Section 4), these have mostly dealt with study populations other than adult ICU patients (primarily neonates and neurologically healthy patients undergoing surgical anesthesia, see Supplemental Material, Table 1). Consequently, the results may not generalize to burst suppression EEG patterns of critically ill adults. Furthermore, validation of these methods against human expert clinical EEG segmentation has been limited (Thomsen et al., 1991; Lipping et al., 1995; Arnold et al., 1996; Griessbach et al., 1997; Sherman et al., 1997; Leistritz et al., 1999; Atit et al., 1999; Särkelä et al., 2002).

Here we present a straightforward algorithm that automatically and in real-time segments burst suppression EEG data across a wide spectrum of patterns recorded in the adult ICU setting, and validate the algorithm results against manual EEG segmentations from two experienced electroencephalographers. Performance of automated segmentation, measured in terms of agreement with expert manual segmentation, is demonstrated to be comparable in all cases with human expert performance.

2. Methods

2.1. Clinical data

A representative sample of burst–suppression EEG recordings from critically ill neurological patients was identified by retrospective review of clinical EEG reports from all ICU patients who underwent continuous EEG monitoring at the Massachusetts General Hospital between August 2010 and March 2012. From these, we selected the first 20 consecutive EEGs reported to show burst suppression patterns. Recordings were only included for analysis if patients were being treated within an intensive care unit (ICU) at the time of the recording; EEG recordings from surgical procedures were excluded. We restricted selection to recordings no longer that 90 min. Given the retrospective nature of the study, we did not control the length of these EEG recordings. Recording durations ranged from 21 to 76 min (see Table 1). Each record was manually segmented in its entirety (see below).

All EEGs were recorded using 19 silver/silver chloride electrodes, affixed to the scalp according to the international 10–20 system. Data were recorded at 512 or 256 Hz, using XLTEK clinical EEG equipment (Natus Medical Inc., Oakville, Canada), subsequently downsampled to 200 Hz.

Clinical information was gathered from review of written inpatient medical notes, imaging studies and reports, EEG reports, and discharge summaries. Baseline demographic data (age, gender), primary admission diagnosis, and the identity of anesthetic agents administered at the time of EEG recording were noted (Table 1). Review of clinical and EEG data was carried out with the approval of the local institutional review board. The study was approved by the Massachusetts General Hospital Human Research Committee.

2.2. Clinical expert classification of bursts and suppressions

A common approach to quantitatively distinguishing bursts and suppressions is to define suppressions as EEG segments of less than a threshold voltage value, typically set at between 0.5 and 20 μV, which last at least 0.5 s (Rampil and Laster, 1992; Niedermeyer et al., 1999). However, prescriptive definitions such as these are problematic, because burst suppression in different pathological conditions or induced by different anesthetics vary widely in amplitude and spectral characteristics (Särkelä et al., 2002; Niedermeyer et al., 1999; Niedermeyer, 2009); see Fig. 1. Moreover, such definitions are not used in clinical practice which relies instead on clinician visual pattern recognition. We therefore adopted an empirical approach to define bursts and suppressions by having two experienced clinical electroencephalographers independently segment all 20 EEG records manually, and defined

"definite" burst or suppression epochs as those portions of the EEG so classified by both reviewers. The algorithms developed below are optimized with respect to these consensus EEG epochs.

2.3. Expert data segmentation

Two experienced clinical electroencephalographers (MBW, MMS) independently reviewed and each manually performed an exhaustive segmentation of all 20 burst–suppression EEGs using custom viewing and annotation software written in-house, implemented in Matlab (Natus, MA). The reviewers were instructed to mark the beginning and end of all instances of "suppressions"; all remaining EEG segments were classified as non-suppressions, referred to hereafter as "bursts". As explained above, to simulate actual clinical practice, reviewers were instructed not to apply explicit technical definitions of "bursts" and "suppressions"; rather, reviewers were to segment the each record according to clinical judgment based on experience. Data review and segmentation was conducted while viewing data from the complete array of recording channels displayed in average referential montage.

2.4. Assessment inter-rater agreement

Inter-rater agreement between manual segmentations of the two expert reviewers was examined in three ways. First, for each of the 20 records, percent agreement was calculated as the fraction of each EEG to which the two reviewers assigned the binary burst vs suppression classification; a global percent agreement was also calculated for all 20 records considered as a whole.

Second, we assessed inter-rater agreement using Cohen's Kappa (κ) values separately for each record and for all records taken as a whole (Carletta, 1996). κ values were calculated as: $\kappa = (a - c)/(1 - c)$, where 'a' is the fraction of EEG data points classified identically by both reviewers (observed inter-rater agreement), and 'c' is the hypothetical probability of chance agreement obtained by assuming that each rater randomly assigned categories to EEG data points using the observed relative frequencies of each category. (κ is a measure of the degree of inter-rater agreement beyond that theoretically achievable by chance; see Supplemental Material.)

Finally, a histogram was computed for the lengths of EEG segments, measured in seconds, on which the two raters disagreed. Summary statistics were computed for this "discrepancy histogram" including the median, mode, and standard deviation.

2.5. Automatic segmentation of burst suppression data

2.5.1. Recursive variance estimation—Our objective was to develop a causal burst suppression classifier, i.e. a classifier that declares each new incoming data point to be part of a burst or suppression based *only* on *past* data. We evaluated a simple classification scheme employing a thresholding operation applied to a recursive estimate of the local signal variance (Haykin, 1996), as expressed in the following three equations:

$$\mu_t = \beta \mu_{t-1} + (1 - \beta) x_t \quad (1)$$

$$\sigma_t^2 = \beta \sigma_{t-1}^2 + (1 - \beta)(x_t - \mu_t)^2 \quad (2)$$

$$z_t = \delta[\sigma_t^2 < \theta] \quad (3)$$

Here x_t is a single-channel EEG signal at time t, in our analysis taken to be the average of signals derived from the frontal EEG channels Fp1 and Fp2 (each derived from an average montage); μ_t is the current value of the recursively estimated local signal mean; σ_t^2 is the current value of the recursively estimated local signal variance; and $\delta[\cdot]$ is the indicator function, i.e. $\delta[\sigma_t^2<\theta]=1$ when $\sigma_t^2<\theta$, and otherwise equal to zero, thus z_t is a binary sequence of zeros (indicating data points classified as bursts) and ones (indicating data points classified as belonging to suppressions). The two free parameters in these equations, β (the "forgetting factor") and θ (the classification threshold), were chosen to optimize agreement on consensus EEG regions, as described below.

The first two equations (adaptive mean and variance) in the above adaptive variance estimation scheme effectively implement a low-pass filtering operation whose dynamics are tuned (by adjustment of the forgetting factor, β, see below) to optimize performance on the burst suppression segmentation problem.

2.5.2. Determination of optimal forgetting factor—The forgetting factor, β, in the recursive estimation equations is a value between 0 and 1 which determines balance between the influence of recent and past data on current estimates of the signal mean and amplitude, μ_t and σ_t, respectively. The forgetting factor is inversely related to the *forgetting time*, the time required for the influence of a unit impulse (i.e. a single 1 within a sequence of zeros) to decay by 63% to the value $1/e \approx 0.37$, by the expression $\tau = -1/(F_s \ln \beta)$, where F_s=200 Hz is the sampling rate for our data.

An optimal value for β for segmentation of adult ICU burst suppression data was determined by a fine grid search over possible values of the forgetting time τ within an broad interval judged plausible on clinical grounds, namely 0.01–3 s. The optimal value for τ (equivalently, β) was determined by first filtering all 20 EEGs using each candidate value of τ, solving for the corresponding optimal threshold value θ (see Section 2.5.3), and then selecting the value of τ which yielded the minimum classification error on the entire data set taken as a whole. Optimality was determined with respect *only* to data on which *both* expert reviewers agreed (consensus segments). Data segments classified differently by the two reviewers were ignored for the purposes of parameter optimization.

Robustness of the estimated optimal forgetting factor to differences in burst suppression characteristics between the 20 EEG recordings was assessed by evaluating the gain in classification accuracy when the forgetting factor is optimized (while simultaneously optimizing the classification threshold—see below) in a patient specific-manner, i.e. on a record-by-record basis.

2.5.3. Determination of optimal classification threshold—For a given value of the forgetting factor β, the optimal classifier threshold θ was efficiently determined via a modification of the method described by Viola and Jones (2004). For data points assigned to the same classes by the two expert reviewers (other data were ignored in parameter optimization) values of the corresponding local variance σ_t^2 were sorted by descending magnitude, and the optimal threshold given these local variance values was computed in a single pass over the sorted list of values as follows. Each value on the sorted list was considered as a potential threshold value, i.e. a point at which to "split" the data, classifying all data points with larger variance as belonging to bursts, and all points with smaller variance as belonging to suppressions. For each candidate threshold value in the sorted list four sums were computed: the total number of data points that would be classified as bursts, T^b, the total number of data points that would be classified as suppressions, T^s, the total number of actual burst data points below the current sample, B^b, and the total number of

actual suppression data points below the current sample, B^s (where "actual" is defined relative to human expert consensus). The error for a threshold value θ which splits the range between the current and previous data point in the sorted list is:

$$e = \min(B^b + (T^s - B^s), B^s + (T^b - B^b))$$

that is, the minimum of the error of labeling all data points below the current point as bursts and labeling the points above as suppressions vs the error of doing the opposite. These sums are easily updated using a single sequential search through the list. Optimal thresholds were determined with respect to all 20 burst suppression EEG records considered as a whole, and separately for each individual record.

2.6. Validation of automated segmentation: inter-rater agreement analysis

Statistical performance of automatic burst suppression segmentations achieved via the adaptive variance thresholding equations (1)–(3), with optimized parameters β, θ, was evaluated by computing global and record-by-record inter-rater agreement and Cohen's κ statistics for the agreement between the automated method and each of the two expert reviewers.

2.7. Quantification of burst suppression depth

The burst suppression probability (BSP) is a recently proposed measure of burst suppression depth which, compared with the more 'traditional burst suppression ratio' (BSR), provides a superior tradeoff between signal smoothness and responsiveness to changes in the EEG (see Supplementary Material). The BSP represents an estimate of the instantaneous probability that the EEG is in the suppressed state, and can be computed recursively on a sample-by-sample basis in real-time (Chemali et al., 2011, 2013). Examples of the result of passing the binary data output from the segmentation algorithm through the BSP algorithm with parameters optimized for adult ICU EEG data are presented in Fig. 5.

2.8. Validation of automated segmentation: root mean squared error of BSP

We further validated our segmentation algorithm by comparing the BSP traces computed from the output of our automated segmentation algorithm with those derived from manual segmentations. Signal comparisons were carried out using a root mean-squared error criterion:

$$\mathrm{RMSE}_{i,j} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\mathrm{BSP}_{i,t} - \mathrm{BSP}_{j,t})}$$

where the indices *i*, *j* have values 1 or 2 (human reviewer 1 or 2) or *A* (algorithm). Performance of the algorithm was assessed by comparing the discrepancy between the two expert's BSP signals with respect to each other, $\mathrm{RMSE}_{1,2}$, with the discrepancies between the automated algorithm's BSP signal and those of each human experts, i.e. $\mathrm{RMSE}_{1,A}$ and $\mathrm{RMSE}_{2,A}$.

3. Results

3.1. Clinical data

The study population consisted of 13 male and 7 female patients ranging in age from 22 to 83 years; see Table 1. The most common admission neurological diagnoses were, taken

separately, anoxic brain injury (11/20, 55%), refractory status epilepticus (13/20, 65%), coma due to traumatic brain injury (2/20, 10%), and metabolic encephalopathy (2/20, 10%). The remaining causes, of which there was one case (1/20, 5%) for each, included subarachnoid hemorrhage, bacterial meningitis, status epilepticus in an epileptic patient due to subtherapeutic anticonvulsant levels, recent neurosurgical removal of an arteriovenous malformation, and brain intraparenchymal hemorrhage. Most patients had more than one of the above admission neurological diagnoses. Anoxic brain injury was the underlying cause of seizures in greater than half of patients with refractory status epilepticus (7/13, 54). Burst suppression was at least partly attributable to anesthetic administration in 19/20 (95%) of cases. The most commonly used anesthetic agent was propofol, administered during the EEG recording in 16/20 (80%) of cases, often in combination with midazolam (7/16, 44%). Midazolam was used alone in 2 (10%) of cases. Phenobarbital and lorazepam were each used in a single case.

Qualitatively the 20 burst suppression EEG records were divisible into 3 classes as determined by consensus of the two clinical reviewers: (1) s*pikes*: records in which bursts contained high amplitude epileptiform discharges, singly or in runs, obtained from patients with status epilepticus or anoxic brain injuries; (2) *distinct:* records for which bursts and suppressions were easily distinguished; and (3) *indistinct:* records for which bursts and suppressions were difficult to distinguish. Two representative examples from the *spikes*, *distinct*, and *indistinct* burst suppression classes are shown in Fig. 1. These three classes varied in the subjective effort required to segment bursts from suppressions, as explored further below. While we did not assume a priori that bursts were homogenous within any given recording, nevertheless, as a qualitative empirical observation, within each EEG bursts did tend to be quite homogenous. Based on our clinical experience, this within-record homogeneity is typical of most cases. Exceptions do occur, for example, in cases of burst suppression as a result of severe cerebral anoxia, bursts occasionally evolve over time such that the overall amplitude decreases and the spectral content shifts toward lower frequencies (unpublished observations).

3.2. Inter-rater agreement

Inter-rater agreement between the two expert manual segmentations is summarized in Table 2. Agreement on the binary classification of data points as belonging to bursts vs suppressions ranged from 70% to 99%, with an average over all 20 records of 90%. Average agreement was highest for records qualitatively classified as *distinct*, and lowest for EEGs in the *indistinct* burst suppression class. Inter-rater κ values followed a similar pattern, ranging from 0.05 up to 0.89, with group average κ values of 0.64 and 0.69 for records in the *distinct* and *spikes* groups, roughly twice the average κ value for records classed as *indistinct*. These results indicate that burst suppression EEGs encountered in the ICU setting encompass a spectrum of patterns of varying difficulty, which human expert reviewers can generally nevertheless parse with a relatively high degree of overall consensus.

To investigate the temporal precision of agreement between manual segmentations of burst suppression EEG patterns, we computed a histogram of the lengths of discrepancies (contiguous data segments which were classified differently in the two expert manual segmentations); see Fig. 2. Discrepant regions tended to be brief relative to the length of consensus regions. In particular, median length, mean length, and the cutoff below which 95% of discrepant segments lay were 0.37, 0.78, and 2.5 s, respectively, as compared with 1.50, 5.2, and 18.65 s for consensus burst regions, and 2.14, 5.76, and 19.09 s for consensus suppression regions. On manual visual review longer discrepancies were nearly always seen to be due to actual disagreements regarding classification of difficult background regions (e.g. over whether a segment was “sufficiently flat” to be classified as a suppression)

whereas brief discrepancies, which constituted the majority, were attributable to inherent limits in the temporal precision of manual segmentation. The median discrepancy values can be taken as an estimate of the level of temporal precision which may be reasonably demanded from any automated segmentation method if it is to be at least as good as expert manual segmentation of burst suppression EEG patterns, a level of precision indeed achieved by our method (see below).

3.3. Parameter estimation

To explore the dependence of performance on the value of the forgetting time (equivalently, the forgetting factor), we performed a series of global optimization and leave-one-out cross validation analyses, in each case setting the value of the other parameter, the segmentation threshold θ, equal to its patient-specific optimal value. The globally optimal forgetting time, calculated with respect to consensus regions from all 20 manual expert EEG segmentations, was $\tau = 104.7$ ms (equivalently, the globally optimal forgetting factor was $\beta = 0.9534$) and yielded a mean (std) accuracy over all records, of 95.5 (3.8)%. To investigate the performance that can be expected on out-of-sample/previously unseen EEG recordings, we performed a leave-one-out cross validation investigation. In this approach, for each EEG recording the performance of the segmentation algorithm is evaluated using the value of τ that is the optimal for the set of all records *except* the record under study. The sample mean (std) value for τ calculated by leave-one-out analysis was $\tau = 100.4$ (20.4) ms, quite close to the globally optimal value given above, and this value produced comparable segmentation accuracy, with mean (std) of 95.4 (4.0)%. As a more stringent validation measure, we also compared the performance achievable by specifically optimizing τ for each individual record alone (without reference to any other records) with the τ value calculated using the leave-one-out method. Performance was again similar, with a mean accuracy achieved with patient-specific τ values of 95.9 (3.4)%, and a mean and maximum performance improvement of 0.5% and 4.6%, respectively (see Fig. 3A). We conclude that patient-specific optimization of the forgetting factor generally does not significantly improve performance achievable using the globally optimized forgetting factor value, hence the algorithm may be regarded as containing essentially only one tunable parameter.

Patient-specific optimization of classification thresholds, θ, resulted in more significant performance gains. Bar graphs in Fig. 3B compare the classification accuracy for each EEG using (a) the population-optimal values for β and θ, (b) the population value for β while individually optimizing θ, or (c) optimizing both β and θ. Relative to classification performance using the population values, individualized optimization of θ yielded significant improvements for several EEG records, with median, mean, and maximal performance improvements of 1.3%, 4.3%, and 30.4%. Joint patient-specific optimization of β and θ yielded only minor further improvements, with median, mean, and maximal performance improvements of 0.1%, 0.4%, and 2.8%.

Given the above figures, a reasonable approach to implement the automated segmentation algorithm presented herein is to use the globally-optimal forgetting factor value estimated for β, but to tune the classification threshold θ applied to the recursive variance estimates on a patient-by-patient basis. This can be done in practice by initializing the classification threshold θ to the globally optimized value, then further tuning it to optimize discrimination between an individual patient's bursts and suppressions. To estimate the duration of data on which patient-specific tuning of θ should be based, we divided each of our 20 EEGs into initial 'training' segments of duration 5, 10, or 15 min, and a second segment of 'testing' data, consisting of all of the data except the initial 15 min. We then estimated the classification threshold θ using only the training segments, and calculated the percentage of the corresponding testing segment which was correctly as classified as bursts and

suppressions. Mean (std) classification accuracy over all 20 records when using training data of length 5, 10, 15 or 20 min was 88.6 (21.0)%, 89.0 (21.0)%, 93.6 (9.3)%, and 93.8 (9.3)%, respectively. Results for individual cases are shown in the bar graphs of Fig. 3C. Thus, in most cases 10 min provided sufficient data for highly accurate segmentation of subsequent burst suppression data, and 15 min of burst suppression data provided sufficient data to ensure good subsequent segmentation performance in all 20 cases.

In all of the following analysis the globally estimated forgetting factor β is used while classification thresholds are optimized on a patient-by-patient basis using the first 15 min of each recording.

3.4. Automatic segmentation performance

Agreement between automated vs human expert segmentation is quantified in Table 2. For the *spikes* burst–suppression EEGs, the mean overall agreement and κ statistic were, respectively, 89%, 69% with the first human rater and 84%, 63% with the second human rater, as compared with human-vs-human agreement of 83%, 61%. The same performance statistics for the *distinct* burst–suppression class were 93%, 75% (algorithm vs human 1) and 94%, 77% (algorithm vs human 2), comparable to the human-vs-human statistics for this class which were 92%, 73%. Finally, for the *indistinct* burst–suppression class, performance statistics were 85%, 36% and 90%, 36% for the algorithm-vs-human comparisons, compared with 82%, 33% for the human vs human comparison. Thus, on average, and for each of the 20 individual records, agreement between the algorithm's and each human's segmentation results was comparable to human–human agreement.

Performance of the algorithm was also excellent when measured in terms of sensitivity and specificity for bursts and suppressions, and overall accuracy (% of record correctly classified), calculated with respect to EEG regions given identical classifications by both human readers. Summary statistics for bursts and suppressions for the entire set of all 20 EEG recordings, given as median [min,max]% are as follows: for suppressions, sensitivity was 93 [64,99]%, and specificity was 97 [64,99]%, for bursts, sensitivity was 97 [73,99]%, and specificity was 93 [73,99]%. Overall classification accuracy on consensus EEG regions was accuracy was 96 [87,99]%. Performance was best on records qualitatively classed in Table 1 as "Spikes" or "Distinct".

3.5. Burst suppression probability

The binary sequence of ones and zeros produced by manual and automated segmentation of the EEG were next processed to produce a continuous time-varying measure of the depth of burst suppression, the burst suppression probability (BSP). Excellent overall agreement was achieved between the BSP computed from the automatic EEG segmentation ('automated BSP') vs BSP signals computed from each human expert's manual segmentation ('manual BSP'); see Fig. 4. For all 20 burst suppression EEG records, agreement between automated vs manual BSP signals was comparable to between-expert agreement; in 70% of cases automated vs human agreement was higher, i.e. the automated BSP signal effectively interpolated between the two manual BSP signals. Thus, automated computation of time-varying BSP signals using the methods described herein are comparable to BSP signals computed from EEG segmentations made by human experts. Samples of the raw-EEG data, automated segmentation results, and BSP signals are shown in Fig. 5.

4. Discussion

Burst suppression EEG patterns are commonly seen in patients undergoing brain activity monitoring during surgical anesthesia, medically induced coma for status epilepticus, or

after hypoxic ischemic brain injury. In these data-intensive contexts automated quantitative methods for extracting and compressing essential information from the raw EEG signal are critical complements to manual interpretation by clinical electroencephalographers. In this study we have presented a method for automatically segmenting burst suppression EEG data from adult ICU patients in real time, and shown that these segmentations can be used for quantifying the depth of burst suppression by computing a continuous time-varying estimate of the burst suppression probability (BSP). Furthermore, we have validated our automated segmentation method by comparison with manual EEG segmentations by two experienced clinical electroencephalographers.

Our results show excellent agreement between the automated and expert segmentations of burst suppression adult ICU EEG recordings. Indeed machine vs human agreement were as good or superior to human vs human agreement across a wide spectrum of clinical cases encompassing a variety of neurological etiologies and anesthetic agents. In none of the 20 cases studied was human vs human agreement substantially superior to human vs algorithm agreement.

Our results indicate that for EEG burst suppression patterns encountered in the adult ICU setting, a very simple algorithm suffices to match human performance. Our segmentation algorithm requires minimal memory and computation and requires optimization of only two parameters: the forgetting factor, β, which determines the effective duration of the past data over which to compute the local mean and variance; and the threshold θ, below which data are classified as suppressions. Of these parameters, our analysis shows that patient-specific optimization significantly improves performance only for the threshold, θ. By contrast, patient specific tuning of the forgetting factor β confers negligible performance improvement, thus the population estimate presented herein may generally be substituted.

At least eight previous publications have presented methods for automated segmentation of burst suppression EEG patterns, based on a variety of EEG features and classification methods, including the nonlinear energy operator, spectral feature clustering, and neural-network classification of adaptive Hilbert-transformed EEG features (Thomsen et al., 1991; Lipping et al., 1995; Arnold et al., 1996; Griessbach et al., 1997; Sherman et al., 1997; Leistritz et al., 1999; Atit et al., 1999). These prior efforts are summarized and improved on in Särkelä et al. (2002), the method most closely related to our method. To place the contributions of our work in context, we therefore briefly comment on the ways in which we have attempted to add value to this literature by improving on the work of Särkelä et al. (2002). Our work adds new value in the following respects. First, we employed a larger and more diverse amount of training data, totaling 12.7 h over 20 subjects, compared with 1.3 h from 17 subjects (5 min from each subject). Second, our EEG data were from ICU patients, many of whom were severely neurologically ill and had burst suppression patterns different in character from those seen in healthy patients, whereas Särkelä et al. report results from patients undergoing anesthesia for elective surgery. This is potentially important given the wide variation in burst suppression patterns between individuals, brain pathology, and anesthetic agents (see Fig. 1); hence EEG segmentation methods validated under different conditions (e.g. inhalational anesthesia, the neonatal setting, or animal studies) cannot a priori be assumed to generalize to the adult ICU setting. Third, we have employed more rigorous methods of validation. By requiring all EEG records to be manually segmented by two experts, we have been able to estimate the degree of precision (e.g. in marking onsets and offsets of bursts) required to achieve human expert-level segmentation accuracy (see Fig. 2). The importance of this point highlighted by the existence of EEG recordings classified in this study as “indistinct”. In these cases, agreement between human expert segmentation is modest, reflecting the intrinsic difficulty of objectively segmenting certain examples of EEG burst suppression. In such cases, treating a single human expert’s

segmentation as the “ground truth” would produce either an unrealistically pessimistic picture of algorithmic performance or overfitting to uninformative features of the EEG. In addition, herein we have performed a thorough cross validation, allowing an estimate of the degree to which algorithm performance is sensitive to precise adjustment of parameter values. By contrast, Särkelä et al. refer to a single ground truth manual segmentation, and no measure of inter-rater agreement is provided. Finally, our method has effectively only one tunable parameter. (Though it in fact has two parameters, we have shown that performance is not sensitive to fine-tuning of the forgetting factor, β, which may be set to a generic population-appropriate value.) By contrast, Särkelä’s et al.’s method involves 13 threshold values (3 thresholds applied to signal amplitudes, and 10 thresholds applied to 3 internal counters used in IF-THEN conditionals), in addition to parameters required in the design and tuning of one high-pass and one low-pass linear filter. Finally, in the present work we have explored the performance improvements achievable by, and the amount of data required, for patient-specific tuning of our algorithm. Särkelä et al. did not explore patient-specific parameter optimization.

These above points notwithstanding, we hasten to add that our work on burst suppression segmentation convinces us that this is a “high signal-to-noise” problem, and that a variety of approaches can be made to yield excellent performance. In particular, we believe that the method of Särkelä et al. can probably be made to perform on ICU burst suppression EEG data as well as our new method. Indeed, it would be surprising if it could not given its larger number of free parameters. On the other hand, it is not obvious a priori on which parameters performance most depends, or whether patient-by-patient tuning would be practical. Nevertheless, our algorithm, despite its relatively simplicity, achieves acceptable performance event without patient-by-patient tuning (using population-derived parameter values), and can be improved in a straightforward and practical way by tuning a single threshold parameter on a small amount of training data, a feature which our experience suggests is valuable in patient populations with widely varying burst suppression EEG characteristics (Fig. 1).

An important issue not directly addressed in the present work is the automated detection or rejection of artifacts. In particular, in this study we chose to use the Fp1 and Fp2 (forehead) EEG electrode positions to monitor burst suppression, as is common practice in the operating room setting during general anesthesia. In awake patients, forehead EEG channels are vulnerable to myogenic and blinking artifacts (Astolfi et al., 2006; Delorme et al., 2007; Tatum et al., 2011), raising potential concerns that non-cortical (i.e. artifactual) voltage fluctuations might “look like” bursts to our algorithm. In our dataset the choice of Fp1/Fp2 made no significant difference in performance, as we saw essentially identical performance results with other choices of adjacent electrode pairs (we explored Fz/Cz and F3/C3 as well; data not shown). On reflection, this is not surprising, for two reasons. First, myogenic artifact tends to be minimal during ICU EEG recordings of burst suppression, because burst suppression is, neurologically, a state of profound coma in which the patient tends to lie motionless (Ching et al., 2012; Niedermeyer, 2009; Brown et al., 2011). Secondly, the cortical voltage signals in burst suppression tend to look very similar at all electrode locations, because burst suppression as a rule is a diffusely synchronous brain state, with burst onsets and offsets appearing essentially simultaneously at all scalp electrodes positions (Ching et al., 2012; Niedermeyer, 2009). Nevertheless, for applications more demanding than passive monitoring (e.g. closed-loop control of anesthesia), rigorous automated distinction between artifactual activity and cortically generated neural activity is essential to ensure safety. Artifact detection in burst suppression has been explored in at least one previous publication on automated segmentation of burst suppression in patients undergoing elective surgery, in which large artifacts were present during the induction of anesthesia were detectable by straightforward methods (Särkelä et al., 2002), some of which may

generalized to EEG recordings in the ICU environment. Artifact rejection in ICU EEG recordings is a focus of ongoing research in our group.

One final minor limitation of the present study is the relatively small number of anesthetic agents contributing to burst suppression in the 20 EEGs that we analyzed. The majority of our cases involved propofol or midazolam, alone or in combination; we had only single cases involving phenobarbital (case 10), lorazepam (case 14), or no anestheic agent (case 13, burst suppression due to brain injury alone). While these are the agents used most commonly in U.S. intensive care units, the character of burst suppression does vary somewhat depending on the choice of anesthetic (e.g. burst tend to be briefer under sevoflurane inhaled anesthesia; unplublished observations), and it is possible that our algorithm may need to specifically tuned when monitoring burst suppression under anesthetics not included or not well represented in our case set. Nevertheless, as we have shown, the algorithm presented here can readily be tuned for specific patients using a small sample (<15 min) of that patient's burst suppression EEG data.

The validated, real-time EEG segmentation method presented herein is suitable for several clinical and basic scientific applications which are the focus of ongoing work in our group. These applications include tracking of burst suppression in the operating room and ICU, and studies of the underlying physiological mechanisms and evolution of burst suppression patterns in hypoxic ischemic brain injury. In the interest of 'reproducible computational research' (Gentleman, 2005; Schwab et al., 2000), the data (de-identified EEG data and manual segmentations) and code used to generate all figures included herein are available by writing to the corresponding author.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Disclosures

This study was supported by the following grants: American Brain Foundation (MBW), NIH/NINDS NS062092 (MBW, SSC); NIH Director's Pioneer Award DP1OD003646 and NIH R01-MH071847 (ENB, SC), and NIH New Innovator Award DP2-OD006454 and K-Award K25-NS057580 (PLP).

References

Amzica F. Basic physiology of burst–suppression. Epilepsia. 2009 Dec; 50(Suppl. 12):38–39. http://dx.doi.org/10.1111/j.1528-1167.2009.02345.x. [PubMed: 19941521]

Arnold M, Doering A, Witte H, Dörschel J, Eisel M. Use of adaptive Hilbert transformation for EEG segmentation and calculation of instantaneous respiration rate in neonates. Journal of Clinical Monitoring. 1996 Jan; 12(1):43–60. [PubMed: 8732816]

Astolfi L, Cincotti F, Mattia D, Babiloni F, Marciani MG, De Vico Fallani F, et al. Removal of ocular artifacts for high resolution EEG studies: a simulation study. Conference proceedings: …Annual international conference of the IEEE engineering in medicine and biology society. IEEE engineering in medicine and biology society. Conference. 2006; vol. 1:976–979. [PubMed: 17946431]

Atit M, Hagan J, Bansal S, Ichord R, Geocadin R, Hansen C, et al. EEG burst detection: performance evaluation. BMES/EMBS conference, 1999. Proceedings of the first joint. 1999; vol. 1:441.

Brenner RP. The interpretation of the EEG in stupor and coma. The Neurologist. 2005 Sep; 11(5):271–284. [PubMed: 16148734]

Brown EN, Purdon PL, Van Dort CJ. General anesthesia and altered states of arousal: a systems neuroscience analysis. Annual Review of Neuroscience. 2011; 34(1):601–628. http://dx.doi.org/10.1146/annurev-neuro-060909-153200.

Carletta J. Assessing agreement on classification tasks: the kappa statistic. Computational Linguistics. 1996; 22(2):249–254.

Chatrian. A glossary of terms most commonly used by clinical electroencephalographers. Electroencephalography and Clinical Neurophysiology. 1974 Nov; 37(5):538–548. [PubMed: 4138729]

Chemali J, Solt K, Purdon PL, Ching S, Brown EN. Burst suppression probability algorithms: state-space methods for tracking EEG burst suppression. Journal of Neural Engineering. 2013 in Press.

Chemali, J.; Kevin Wong, KF.; Solt, K.; Brown, EN. A state-space model of the burst suppression ratio; Conference proceedings: …Annual international conference of the IEEE engineering in medicine and biology society. IEEE engineering in medicine and biology society. Conference; 2011. p. 1431-1434.

Ching S, Purdon PL, Vijayan S, Kopell NJ, Brown EN. A neurophysiological-metabolic model for burst suppression. Proceedings of the National Academy of Sciences of the United States of America. 2012 Feb; 109(8):3095–3100. http://dx.doi.org/10.1073/pnas.1121461109. [PubMed: 22323592]

Cloostermans MC, van Meulen FB, Eertman CJ, Hom HW, van Putten MJAM. Continuous electroencephalography monitoring for early prediction of neurological outcome in postanoxic patients after cardiac arrest: a prospective cohort study. Critical Care Medicine. 2012 Oct; 40(10): 2867–2875. http://dx.doi.org/10.1097/CCM.0b013e31825b94f0. [PubMed: 22824933]

Cottenceau V, Petit L, Masson F, Guehl D, Asselineau J, Cochard J-F, et al. The use of bispectral index to monitor barbiturate coma in severely brain-injured patients with refractory intracranial hypertension. Anesthesia and Analgesia. 2008 Nov; 107(5):1676–1682. http://dx.doi.org/10.1213/ane.0b013e318184e9ab. [PubMed: 18931232]

Dandan Z, Jia X, Ding H, Ye D, Thakor NV. Application of Tsallis entropy to EEG: quantifying the presence of burst suppression after asphyxial cardiac arrest in rats. IEEE Transactions on Bio-Medical Engineering. 2010 Apr; 57(4):867–874. http://dx.doi.org/10.1109/TBME.2009.2029082. [PubMed: 19695982]

Delorme A, Sejnowski T, Makeig S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage. 2007 Feb; 34(4):1443–1449. http://dx.doi.org/10.1016/j.neuroimage.2006.11.004. [PubMed: 17188898]

Derbyshire AJ, Rempel B, Forbes A, Lambert EF. The effects of anesthetics on action potentials in the cerebral cortex of the cat. American Journal of Physiology–Legacy Content. 1936; 116(3):577–596.

Doyle PW, Matta BF. Burst suppression or isoelectric encephalogram for cerebral protection: evidence from metabolic suppression studies. British Journal of Anaesthesia. 1999 Oct; 83(4):580–584. [PubMed: 10673873]

Gentleman R. Reproducible research: a bioinformatics case study. Statistical Applications in Genetics and Molecular Biology. 2005; 4 http://dx.doi.org/10.2202/1544-6115.1034, Article2.

Griessbach G, Eiselt M, Dörschel J, Witte H, Galicki M. Common optimization of adaptive preprocessing units and a neural network during the learning period. Application in EEG pattern recognition. Neural Networks: The Official Journal of the International Neural Network Society. 1997 Aug; 10(6):1153–1163. [PubMed: 12662508]

Hallberg B, Grossmann K, Bartocci M, Blennow M. The prognostic value of early aEEG in asphyxiated infants undergoing systemic hypothermia treatment. Acta Paediatrica (Oslo, Norway: 1992). 2010 Apr; 99(4):531–536. http://dx.doi.org/10.1111/j.1651-2227.2009.01653.x.

Haykin, SS. Adaptive filter theory. Prentice Hall; 1996.

Leistritz L, Jäger H, Schelenz C, Witte H, Putsche P, Specht M, et al. New approaches for the detection and analysis of electroencephalographic burst–suppression patterns in patients under sedation. Journal of Clinical Monitoring and Computing. 1999 Aug; 15(6):357–367. [PubMed: 12578031]

Lipping T, Jäntti V, Yli-Hankala A, Hartikainen K. Adaptive segmentation of burst–suppression pattern in isoflurane and enflurane anesthesia. International Journal of Clinical Monitoring and Computing. 1995; 12(3):161–167. [PubMed: 8583169]

Musialowicz T, Mervaala E, Kälviäinen R, Uusaro A, Ruokonen E, Parviainen I. Can BIS monitoring be used to assess the depth of propofol anesthesia in the treatment of refractory status epilepticus? Epilepsia. 2010 Aug; 51(8):1580–1586. http://dx.doi.org/10.1111/j.1528-1167.2009.02514.x. [PubMed: 20132290]

Niedermeyer E, Sherman DL, Geocadin RJ, Hansen HC, Hanley DF. The burst–suppression electroencephalogram. Clinical EEG (Electroencephalography). 1999 Jul; 30(3):99–105. [PubMed: 10578472]

Niedermeyer E. The burst–suppression electroencephalogram. American Journal of Electroneurodiagnostic Technology. 2009 Dec; 49(4):333–341. [PubMed: 20073416]

Nunes ML, Giraldes MM, Pinho AP, Costa da Costa J. Prognostic value of non-reactive burst suppression EEG pattern associated to early neonatal seizures. Arquivos de Neuro-psiquiatria. 2005 Mar; 63(1):14–19. S0004-282X2005000100003. [PubMed: 15830058]

Rampil IJ, Laster MJ. No correlation between quantitative electroencephalographic measurements and movement response to noxious stimuli during isoflurane anesthesia in rats. Anesthesiology. 1992 Nov; 77(5):920–925. [PubMed: 1443747]

Riker RR, Fraser GL, Wilkins ML. Comparing the bispectral index and suppression ratio with burst suppression of the electroencephalogram during pentobarbital infusions in adult intensive care patients. Pharmacotherapy. 2003 Sep; 23(9):1087–1093. [PubMed: 14524640]

Rossetti AO, Milligan TA, Vulliémoz S, Michaelides C, Bertschi M, Lee JW. A randomized trial for the treatment of refractory status epilepticus. Neurocritical Care. 2011 Feb; 14(1):4–10. http://dx.doi.org/10.1007/s12028-010-9445-z. [PubMed: 20878265]

Särkelä M, Mustola S, Seppänen T, Koskinen M, Lepola P, Suominen K, et al. Automatic analysis and monitoring of burst suppression in anesthesia. Journal of Clinical Monitoring and Computing. 2002 Feb; 17(2):125–134. [PubMed: 12212991]

Schwab M, Karrenbach M, Claerbout J. Making scientific computations reproducible. Computing in Science and Engineering. 2000

Sherman, DL.; Brambrink, AM.; Walterspacher, D.; Dasika, VK.; Ichord, R.; Thakor, NV. Detecting EEG bursts after hypoxic-ischemic injury using energy operators. Engineering in medicine and biology society, 1997; Proceedings of the 19th annual international conference of the IEEE; 1997. p. 1188-1190.

Shorvon S. The treatment of status epilepticus. Current Opinion in Neurology. 2011 Apr; 24(2):165–170. http://dx.doi.org/10.1097/WCO.0b013e3283446f31. [PubMed: 21378567]

Swank RL, Watson CW. Effects of barbiturates and ether on spontaneous electrical activity of dog brain. Journal of Neurophysiology. 1949 Mar; 12(2):137–160. [PubMed: 18114367]

Tatum WO, Dworetzky BA, Schomer DL. Artifact and recording concepts in EEG. Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society. 2011 Jun; 28(3):252–263. http://dx.doi.org/10.1097/WNP.0b013e31821c3c93. [PubMed: 21633251]

Thomsen CE, Rosenfalck A, Nørregaard Christensen K. Assessment of anaesthetic depth by clustering analysis and autoregressive modelling of electroencephalograms. Computer Methods and Programs in Biomedicine. 1991 Mar; 34(2–3):125–138. [PubMed: 2060286]

Vijn PC, Sneyd JR. I.v. anaesthesia and EEG burst suppression in rats: bolus injections and closed-loop infusions. British Journal of Anaesthesia. 1998 Sep; 81(3):415–421. [PubMed: 9861133]

Viola P, Jones MJ. Robust real-time face detection. International Journal of Computer Vision. 2004; 57(2):137–154.

Woodcock TE, Murkin JM, Farrar JK, Tweed WA, Guiraudon GM, McKenzie FN. Pharmacologic EEG suppression during cardiopulmonary bypass: cerebral hemodynamic and metabolic effects of thiopental or isoflurane during hypothermia and normothermia. Anesthesiology. 1987 Aug; 67(2): 218–224. [PubMed: 3605748]

Young GB. The EEG in coma. Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society. 2000 Sep; 17(5):473–485. [PubMed: 11085551]

Young GB, Wang JT, Connolly JF. Prognostic determination in anoxic-ischemic and traumatic encephalopathies. Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society. 2004 Oct; 21(5):379–390. [PubMed: 15592010]

HIGHLIGHTS

- Monitoring the depth of EEG burst suppression is critical in caring for patients with certain neurological critical illnesses.
- We present an automated method for real-time segmentation of adult burst suppression ICU EEGs.
- Our method achieves expert-level automated real-time segmentation of burst suppression EEG data.

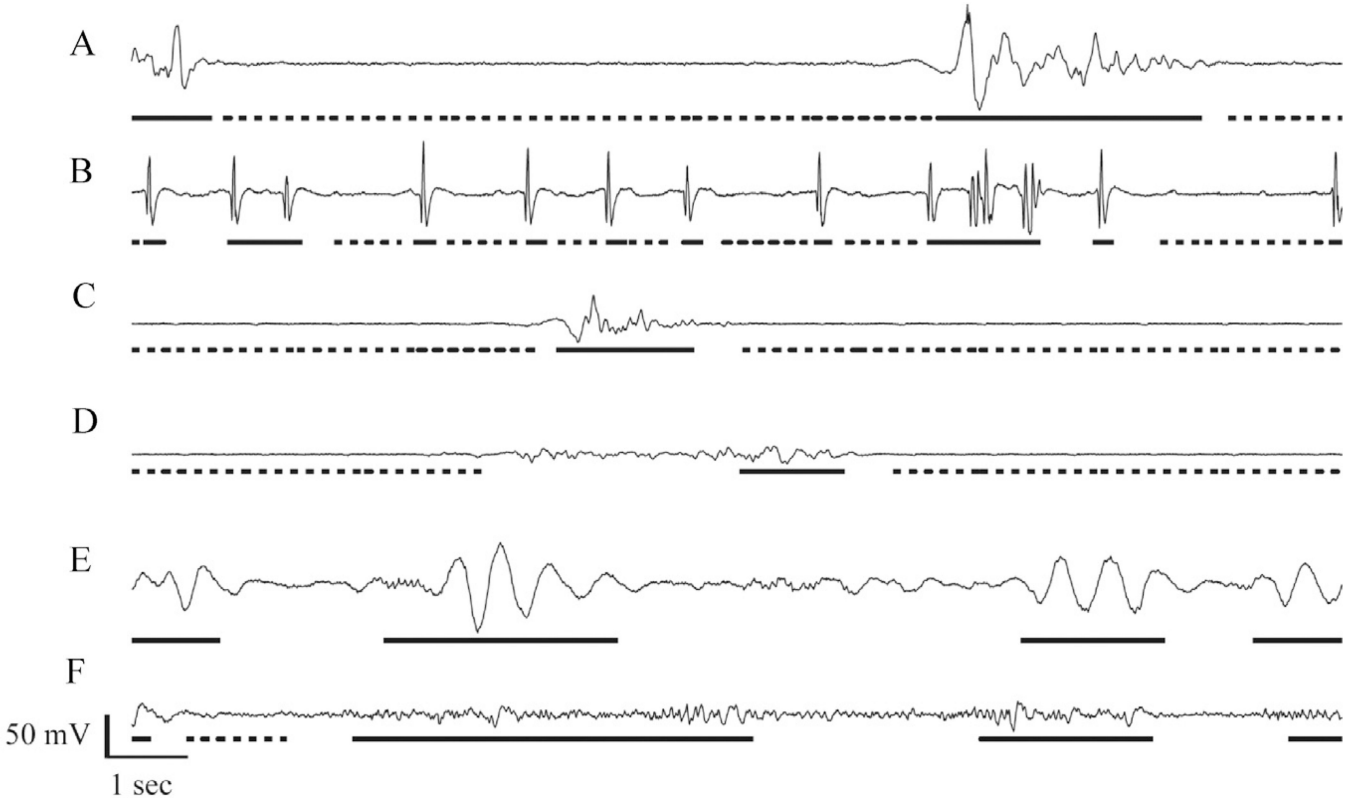

Fig. 1.

Examples of ICU burst suppression EEG patterns. From top to bottom are shown two examples from each of three qualitative classes: (A and B) patterns containing epileptiform spikes (patients 1 & 4); (C and D) patterns with distinct burst and suppressions (patients 8 & 13); and (E and F) patterns with indistinct bursts (patients 18 & 20). Underlining indicates consensus regions: segments with dashed underlining were marked by both reviewers as suppressions, whereas segments with solid underlining were marked by both reviewers as non-suppressions (bursts). Segments not underlined represent regions of disagreement. For example, in D there was apparently disagreement over the precise location of the border between the burst and surrounding suppression, whereas in E and F there is apparent disagreement over whether the background is “flat enough” between consensus bursts to constitute suppression.

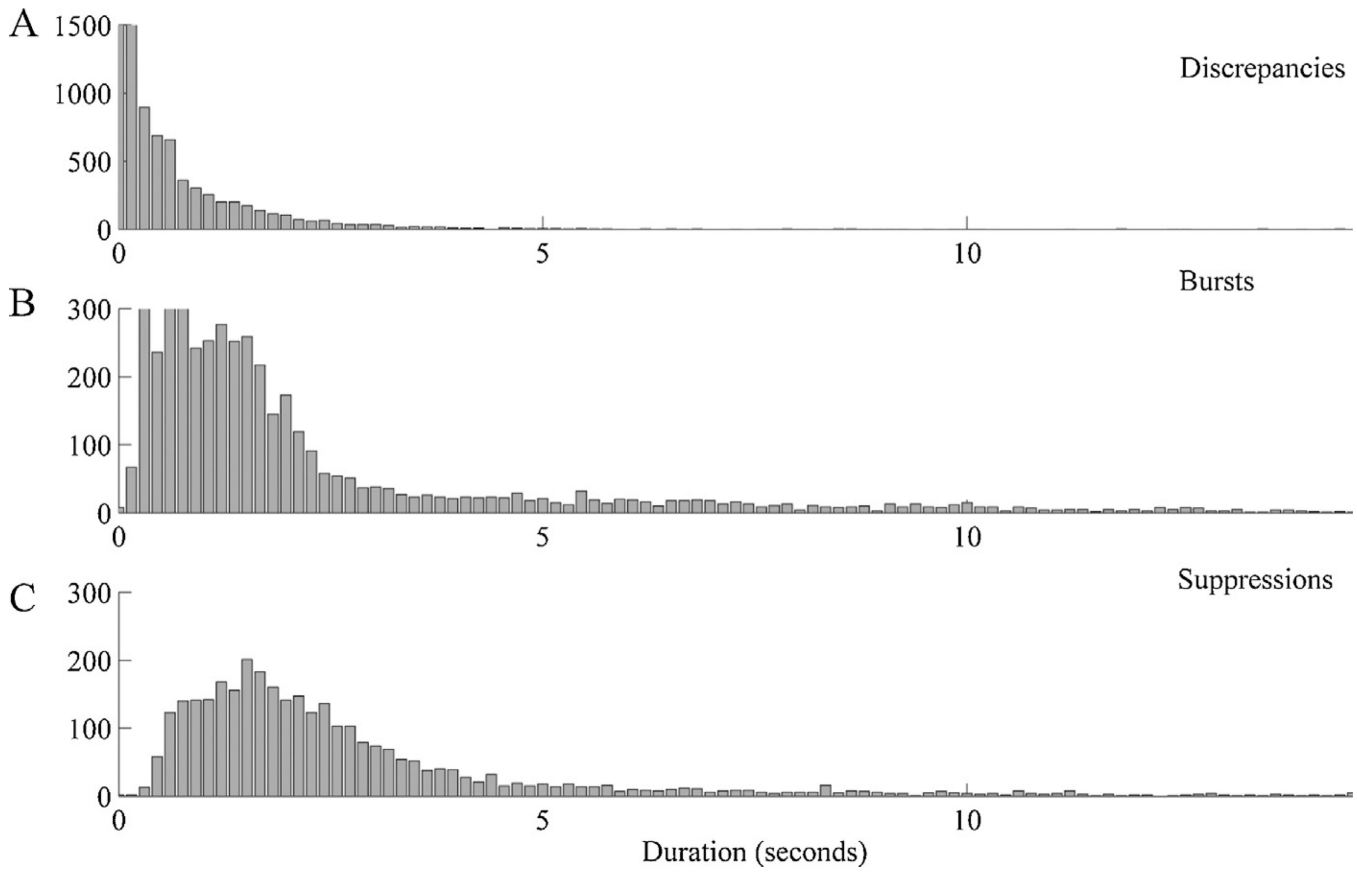

Fig. 2.

Length distributions for EEG segments on which the two manual segmentations disagreed (A, “discrepancies”), for segments agreed to belong to bursts (B), and for segments agreed to belong to suppressions (C).

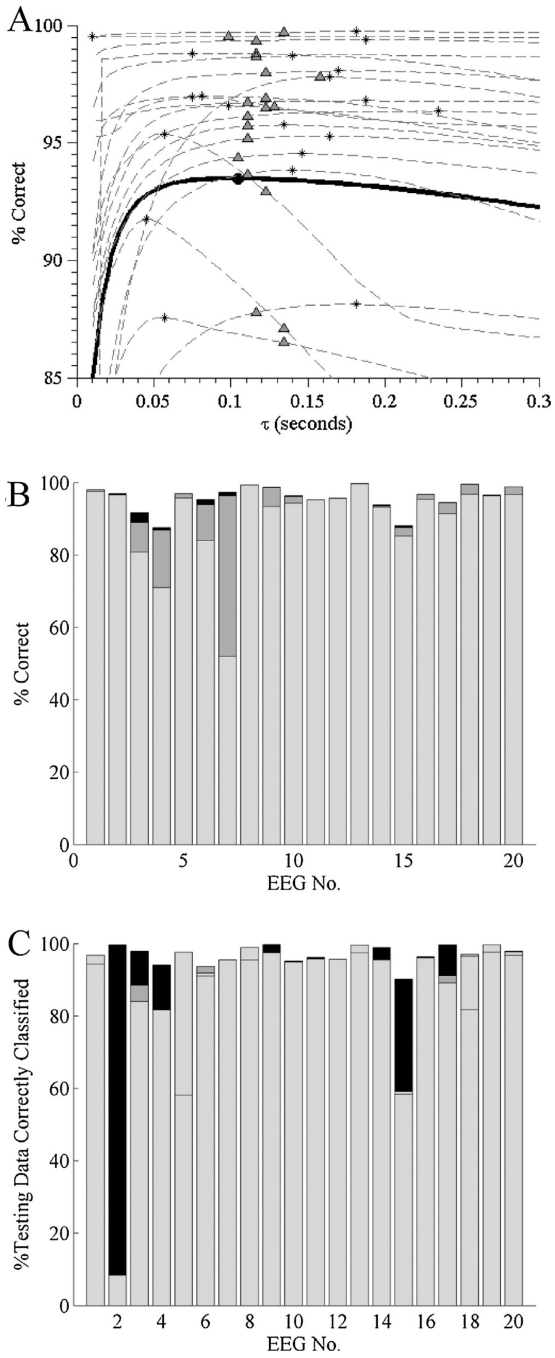

Fig. 3.

(A) Sensitivity analysis for the value of the forgetting factor, β, used in the adaptive variance thresholding segmentation algorithm. The solid curve denotes global performance, i.e. the percentage of data points correctly classified (relative to consensus regions of manually segmented data) over all 20 EEGs; the globally optimal forgetting factor value is marked with a solid filled circle. Dashed curves indicate classification performance vs forgetting factor value for each of the 20 individual patient records. Asterisks mark the patient-specific optimal value of β; open circles mark the performance on the same patient's EEG when the globally optimal forgetting factor is used instead. For every point on every curve we use the corresponding optimal value (for that value of β) of θ, the segmentation threshold (see Eq.

(3)). (B) Relative contributions to performance of patient-specific optimization of the forgetting factor, β, and the classification threshold, θ, for each of the 20 ICU burst suppression EEGs under study. Overall bar height reflects performance when both parameters are individualized. The height of the light gray bar shows performance using the globally optimized parameters; the height of the dark gray bar shows the performance gained by optimizing the threshold while using the population-level forgetting factor; and the black bar shows the (in many cases negligible) peformance gain upon patient-specific optimization of the forgetting factor.

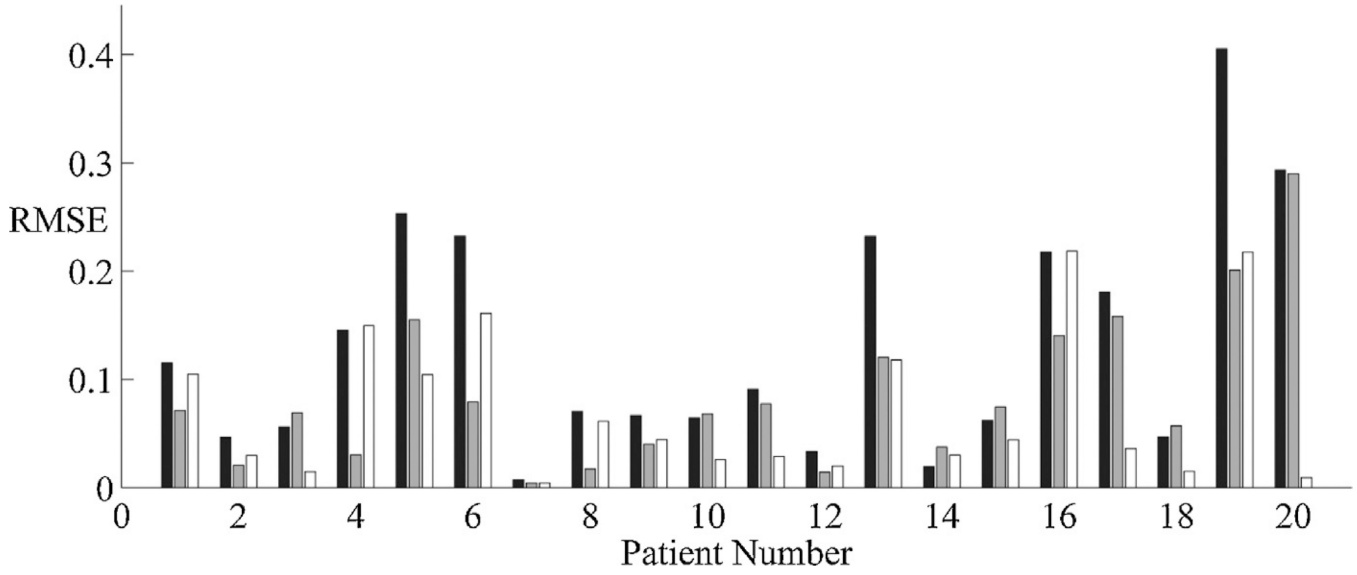

Fig. 4.

Comparison of time-varying burst suppression probability (BSP) signals resulting from processing the two human vs automated EEG segmentations for each of the 20 burst suppression EEGs (*x*-axis) in the study. Bars show the dissimilarity, measured as root mean squared error (RMSE), between the two manual clinical expert segmentations (black bars) and between the algorithm's segmentation and each manual segmentation (gray and white bars).

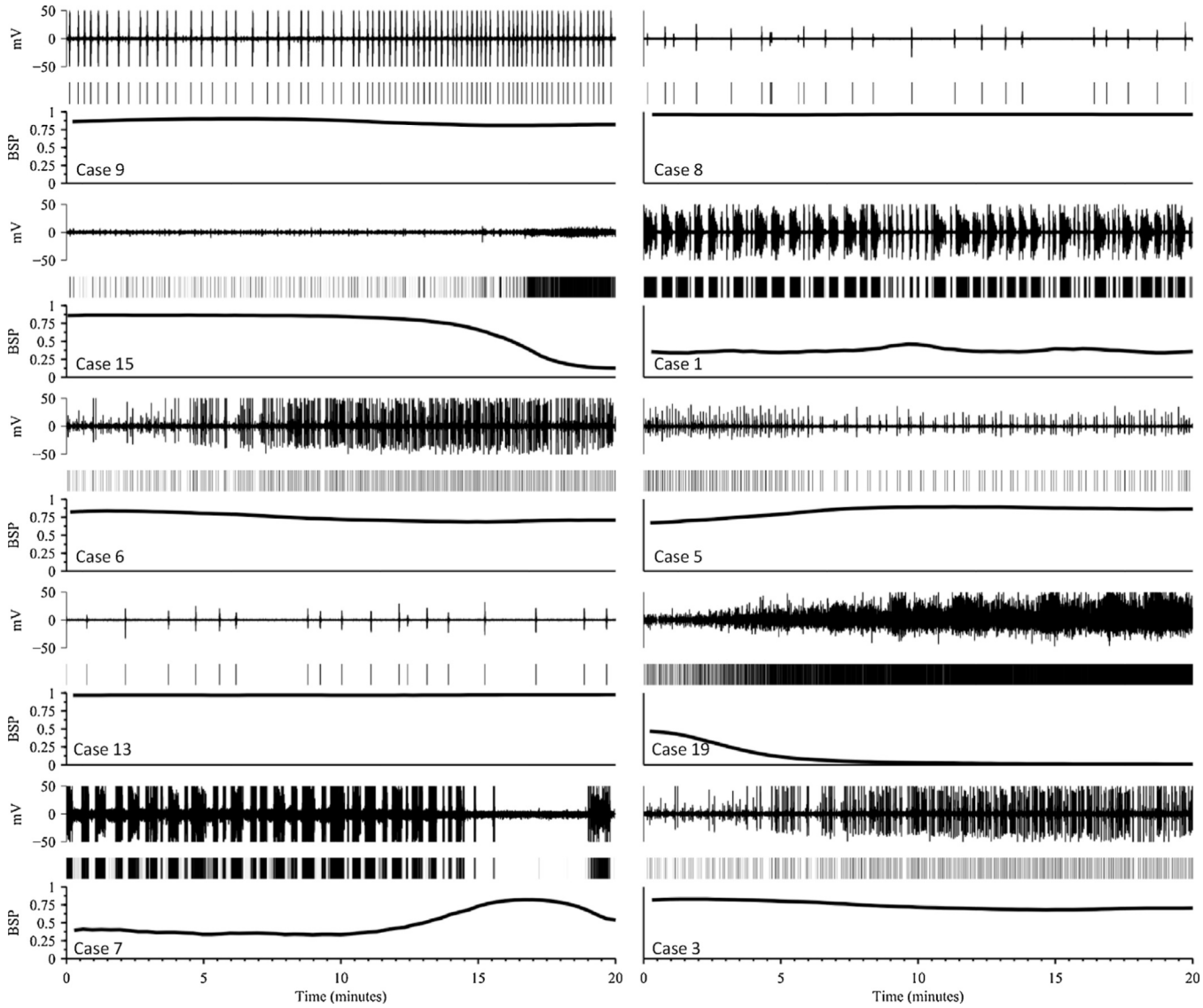

Fig. 5.

Ten examples of burst suppression probability (BSP) vs time as a quantitative measure of burst suppression depth. Twenty minute samples are shown for each of 10 EEG recordings. Each subfigure shows the EEG voltage (top), output of the segmentation algorithm (middle), and the estimated BSP. Case numbers are indicated, corresponding to the case numbering in Tables 1 and 2.

Table 1

Clinical data for 20 ICU patients with burst suppression EEG patterns.

	EEG no.	Age, sex	Recording duration (min)	Dx	Anesthetic
Spikes	1	59M	21	ABI, RSE	prop., midaz.
	2	80M	24	ABI, RSE	prop., midaz.
	3	67M	30	ABI, RSE	prop., midaz.
	4	83M	22	ABI, RSE	midaz.
	5	64M	76	Bacterial meningitis, RSE	prop.
	6	66F	21	ABI, RSE	prop.
	7	81F	64	ABI, RSE	prop. midaz.
Distinct	8	59M	30	ABI, RSE	prop.
	9	22F	39	Traumatic brain injury	prop.
	10	61M	35	Subdural hematoma, RSE	prop., midaz., phenobarb.
	11	60M	35	ABI	prop.
	12	55M	76	ICH, RSE	prop.
	13	47F	28	RSE, ABI	None
	14	59M	40	Hepatic encpephalopathy, RSE	lorazepam
Indistinct	15	53F	45	Hepatic encephalopathy	prop.
	16	56F	23	Post neurosurgery, RSE	prop., midaz.
	17	82M	21	ABI	midaz.
	18	45M	76	Epilepsy, RSE	prop.
	19	37F	30	Traumatic brain injury	prop.
	20	47M	30	ABI	prop., midaz.

Abbreviations: M, male; F, female; ABI, anoxic brain injury; RSE, refractory status epilepticus; ICH, intracerebral hemorrhage; prop., propofol; midaz., midazolam; phenobarb., phenobarbital.

Table 2

Inter-rater agreement statistics for electroencephalographers (1 vs 2) and between manual segmentations and automated segmentation method (1 vs *A* and 2 vs *A*).

Class	EEG#	Agreement (%)			κ (%)		
		1 vs 2	1 vs *A*	2 vs *A*	1 vs 2	1 vs *A*	2 vs *A*
Spikes	1	94	96	95	88	92	88
	2	96	94	96	70	54	63
	3	85	86	80	64	63	53
	4	81	81	79	63	62	58
	5	76	83	88	52	62	77
	6	75	87	79	44	62	54
	7	71	94	72	44	89	45
	Mean	83	89	84	61	69	63
Distinct	8	99	99	99	89	85	85
	9	95	97	96	84	88	85
	10	92	93	92	78	80	77
	11	92	90	93	76	68	76
	12	90	89	93	70	68	76
	13	97	99	98	57	71	72
	14	76	83	84	55	66	67
	Mean	92	93	94	73	75	77
Indistinct	15	77	84	73	49	57	42
	16	93	92	95	46	29	29
	17	82	82	90	44	48	56
	18	97	97	99	31	14	33
	19	70	83	82	24	60	34
	20	72	72	98	5	6	23
	Mean	82	85	90	33	36	36