Clinical Neurophysiology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Clinical Neurophysiology

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

Quantification of EEG reactivity in comatose patients

Mathilde C. Hermans [a,d], M. Brandon Westover [b], Michel J.A.M. van Putten [c], Lawrence J. Hirsch [d], Nicolas Gaspard [d,e,*]

[a] Department of Technical Medicine, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands
[b] Department of Neurology, Massachusetts General Hospital, Boston, MA 02114-2622, USA
[c] Department of Neurology and Clinical Neurophysiology, Medisch Spectrum Twente and Clinical Neurophysiology Group, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands
[d] Department of Neurology, Comprehensive Epilepsy Center, Yale University School of Medicine, PO Box 208018, New Haven, CT 06520-8018, USA
[e] Department of Neurology, Comprehensive Epilepsy Center, Université Libre de Bruxelles – Hôpital Erasme, Route de Lennik, 808, 1070 Bruxelles, Belgium

ARTICLE INFO

Article history:
Accepted 5 June 2015
Available online xxxx

Keywords:
EEG reactivity
Coma
Automated analysis
Quantitative EEG
tBSI

HIGHLIGHTS

- Quantitative EEG features are used to classify reactive and non-reactive EEGs.
- Probabilities based on quantitative EEG features reflect the degree of reactivity.
- Quantitative methods may increase reproducibility and objectivity of EEG reactivity assessment.

ABSTRACT

Objective: EEG reactivity is an important predictor of outcome in comatose patients. However, visual analysis of reactivity is prone to subjectivity and may benefit from quantitative approaches.
Methods: In EEG segments recorded during reactivity testing in 59 comatose patients, 13 quantitative EEG parameters were used to compare the spectral characteristics of 1-minute segments before and after the onset of stimulation (spectral temporal symmetry). Reactivity was quantified with probability values estimated using combinations of these parameters. The accuracy of probability values as a reactivity classifier was evaluated against the consensus assessment of three expert clinical electroencephalographers using visual analysis.
Results: The binary classifier assessing spectral temporal symmetry in four frequency bands (delta, theta, alpha and beta) showed best accuracy (Median AUC: 0.95) and was accompanied by substantial agreement with the individual opinion of experts (Gwet's AC1: 65–70%), at least as good as inter-expert agreement (AC1: 55%). Probability values also reflected the degree of reactivity, as measured by the inter-experts' agreement regarding reactivity for each individual case.
Conclusion: Automated quantitative EEG approaches based on probabilistic description of spectral temporal symmetry reliably quantify EEG reactivity.
Significance: Quantitative EEG may be useful for evaluating reactivity in comatose patients, offering increased objectivity.

1. Introduction

Accurate prediction of neurologic outcome is essential for the care of comatose patients, particularly to guide decisions about whether or not to continue supportive care. Prognostication in comatose patients is based on both clinical and electrophysiological parameters (Gaspard et al., 2014; Tjepkema-Cloostermans et al., 2015; Wijdicks et al., 2006). One of these variables is the reactivity of the EEG to external stimulation, which has emerged as an important predictor of improved outcome in a wide variety of clinical conditions, including traumatic and anoxic brain injury (Logi et al., 2011; Rossetti et al., 2010). In multimodal prediction of outcome after anoxic brain injury, the information provided by EEG reactivity complements the information derived from clinical examination and

* Corresponding author at: Department of Neurology, Comprehensive Epilepsy Center, Université Libre de Bruxelles – Hôpital Erasme, Route de Lennik, 808, 1070 Bruxelles, Belgium.

http://dx.doi.org/10.1016/j.clinph.2015.06.024
1388-2457/

somatosensory evoked potentials (SSEP) (Oddo and Rossetti, 2014; Rossetti et al., 2010).

EEG reactivity is generally regarded as presence of any change in frequency or amplitude of the EEG background pattern, detected after the application of an external stimulus (Young, 2000), although no consensus exists about the detailed characteristics or exact timing or duration of changes involved in a responsive EEG. External stimulation often includes applying auditory stimuli (i.e. shouting or clapping), somatosensory stimuli (i.e. applying pressure to the nail bed or supraorbital nerve) or visual stimuli (i.e. passive eye opening), but there is no consensus about the particular stimulus or stimuli that need to be applied.

In current practice, EEG reactivity in comatose patients is assessed by visual comparison of EEG segments before and after the time of stimulus administration. However, visual EEG analysis can be difficult and is prone to subjectivity (Gerber et al., 2008; Noirhomme et al., 2014; Young et al., 1997).

By providing an objective assessment of the EEG signal and detecting subtle changes in the signal that might escape visual assessment (Claassen et al., 2004; Vespa et al., 1997), quantitative EEG (qEEG) analysis can assist the interpretation of the EEG (Lodder and van Putten, 2013; Nuwer, 1997).

In the present study, several quantitative approaches that score differences in signal characteristics are compared with visual analysis of EEG reactivity by experts. With this work, we aimed to explore the possibility for automated quantification of EEG reactivity in comatose patients.

2. Methods

2.1. EEG data

We used EEG recordings of 70 consecutive comatose patients that were admitted to the ICU and underwent continuous EEG and stimulation to assess reactivity for neurological prognostication. Coma was clinically defined as the absence of meaningful clinical response to noxious stimulation (i.e. withdrawal or better response on the motor component of the Glasgow Coma Scale (Wijdicks et al., 2006)). At the time of EEG recording, some patients were sedated with midazolam, propofol or a combination of these and some were hypothermic.

The set of EEG recordings was selected from the digital ICU EEG databases of Yale-New Haven Hospital (New Haven, Connecticut, USA), Massachusetts General Hospital (Boston, Massachusetts, USA) and Medisch Spectrum Twente (Enschede, the Netherlands). The institutional review boards of participating institutions did not require informed consent and approved the research protocols under which the study was conducted.

Recordings included standard arrangements of 19 electrodes placed according to the international 10–20 system. Recordings were routinely performed with commercially available medical-grade EEG equipment, with a sampling frequency of 200, 256 or 512 Hz.

EEG reactivity was tested as part of routine clinical care at each institution using a variety of external stimuli. Stimuli were administered sequentially and included calling the subject's name, clapping hands, shaking the subject, administration of central or peripheral noxious stimuli (nostril tickle and/or nail bed pressure), and passive eye opening. The sequence of stimuli takes less than 30 s and could have been interrupted in case the subject showed a clinical response. We cannot exclude some discrepancy between the technicians' and the treating physicians' clinical assessment, but followed the treating physicians' opinion. One epoch of stimulation was selected for every patient. The time of onset of reactivity testing, i.e. start of stimulation, was determined using the notes in the original EEG file and confirmed by reviewing the corresponding video recording. Since the goal of this study was to develop a quantitative method for assessing reactivity that is competitive with the expert's opinion regarding reactivity, and acknowledging the variability in the way stimulation was provided to patients, rather than pursuing its use for prognostication, information about outcome was not collected. In addition, no attempt was made to select patients stimulated in any specific or uniform manner, or to ensure uniformity of testing conditions or methods of external stimulation.

EEG recordings with prominent artifacts prior to stimulation were excluded. EEG recordings with suppression-burst were also excluded from the analysis as spontaneous short-term variation in the duration and spectral content of the burst between the pre- and post-stimulation epoch may affect the performance of the detector.

In addition to the selected test cases, a verifications set of 34 2-min EEG clips was randomly selected from comatose patients with background better than suppression-burst between 12 am and 6 am, at a time no active stimulation was provided.

2.2. Visual scoring

Reactivity of each test case was assessed by three independent practicing, board certified, clinical neurophysiologists routinely involved in prognostication of patients with coma. These expert EEG readers evaluated responsiveness using standard visual analysis of the total EEG recording, under ordinary clinical conditions, without time constraints. During analysis, the expert readers were able to manipulate the 19 channel EEG by changing filters, montages, signal gain, and the amount of data shown on a single EEG review screen.

Before assessment, the expert readers agreed to use a reactivity score in which each case was classified as 'Reactive', 'Non-reactive' or 'Unclear'. The presence of stimulus-induced rhythmic, periodic, or ictal discharges (SIRPIDS) was coded separately but regarded as reactivity. Presence of a change in electromyographic (EMG) activity by itself without corresponding EEG changes was regarded as non-reactive. Besides this common set of rules, experts were free to determine what represented reactivity as they would in routine clinical practice. For purposes of model training, cases were categorized according to the majority (2/3) of the experts' opinion. Cases in which no consensus was reached were classified as unclear reactivity.

2.3. Quantitative analysis

For each EEG recording, a set of qEEG parameters was calculated. These qEEG features were based on different computational approaches, which have been described as potentially capable of quantifying changes in spectral characteristics of the EEG. In addition to the qEEG features describing EEG reactivity, parameters detecting EMG reactivity were implemented.

The quantitative analysis of EEG reactivity involved comparison of the EEG characteristics before and after administration of external stimulation. The pre-stimulation epoch that was selected as baseline included the EEG segment of 60 s prior to the documented time of the onset of reactivity testing. The subsequent 60 s segment starting at the onset of stimulation was selected as the post-stimulation epoch. The duration of the epochs was chosen to allow the detection of any EEG changes during the total period of stimulation, which may last up to 30 s, and the detection of delayed responses, which in our experience can occur several seconds after the stimulus. Quantitative analysis was performed for all derivations of a bipolar longitudinal montage.

EEG recordings were first exported in standard European Data Format (.edf), and then imported into the Matlab (Natick, MA)

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

computing environment for further analysis. Independent component analysis (ICA), as implemented by the *FastICA* tool version 2.5 for Matlab (Gävert et al., 2005) was used to identify and eliminate the component with the highest correlation with the electrocardiogram (ECG). Furthermore, components with high kurtosis were removed in order to minimize artifacts (Delorme et al., 2001) using a threshold of 15, which was established through observation of the data. Components were manually checked to prevent elimination of EEG signal. After detection of EMG reactivity (see below), a 4th order zero-phase high pass Butterworth filter with a cut-off frequency of 0.5 Hz was used to reduce baseline instability.

The power spectrum of the EEG segments was estimated with the Thomson's multi-taper method (Thomson, 1982) implemented in the Chronux toolbox (Bokil et al., 2007, 2010), which generates spectral estimates with an optimal balance between spectral resolution (bias) and variance. The frequency resolution, i.e. the intervals between spectral components, was 0.39 and 0.5 Hz, depending on the sampling rate. In all cases the spectral estimation was based on 3 slepian tapers, a moving window length of 1.5 s and a step length of 0.1 s, resulting in a spectral resolution bandwidth, i.e. the minimal proximity in which two spectral peaks are clearly distinguishable, of 2.67 Hz, independent of sampling rate. Analysis of EEG spectral characteristics was confined to 1–18 Hz, to mitigate the contamination by the electromyography (EMG) activity. Within this total frequency band (1–18 Hz or 1.2–18.4 Hz, depending on frequency resolution), we specified the delta (δ; 1–4 Hz or 1.2–4.3 Hz), theta (θ; 4–8 Hz or 4.3–8.2 Hz), alpha-band (α; 8–12 Hz or 8.2–12.1 Hz) and beta-band (β: 16–18 Hz or 16–18.4 Hz).

EMG activity of scalp muscles, when present, influences spectral characteristics of the EEG (Goncharova et al., 2003). This is potentially problematic, as the neurologic prognostic significance of EEG reactivity in comatose patients is conventionally thought to depend on measuring the potential of cortical activity to react to stimulation, whereas EMG activity may arise from extra-cortical activity. Elimination of EMG artifacts remains a challenge, despite the use of artifact reduction methods. As a result, EMG reactivity on stimulation might influence qEEG features, with the potential to produce false positive assessments of reactivity.

In the present study, the ratio of mean power in the 20–35 Hz and 1–4 Hz frequency bands (muscle activity ratio) and the ratio in mean 20–35 Hz power before and after stimulation (muscle reactivity ratio), were used to detect EMG reactivity of a certain quantity, as described in more detail in the Supplementary Material. Cases in which EMG reactivity was detected were excluded from analysis.

2.4. Temporal brain symmetry index

In order to quantify temporal changes in spectral characteristics of EEG, the temporal brain symmetry index (tBSI) was proposed in a recent study (van Putten, 2006). The tBSI is defined as the normalized difference between spectral estimates of two EEG epochs and thus provides a measure of temporal invariance or symmetry. In the current study, the tBSI is calculated to assess the difference between the pre-stimulation epoch and post-stimulation within an individual derivation:

$$\mathrm{tBSI} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=q}^{j=q+k}\left|\frac{Spost_{i,j} - Spre_{i,j}}{Spost_{i,j} + Spre_{i,j}}\right| \quad (1)$$

Here, $Spost_{i,j}$ and $Spre_{i,j}$ are the spectrogram of the post- and pre-stimulation epochs with $i = 1,2,\ldots N$ time samples and $j = q,..q+K$ frequency components starting at frequency q. The tBSI was separately calculated for the previously described δ, θ, α and β-bands and the total frequency band.

2.5. Two-group test

The two-group test (Bokil et al., 2007) is a method for comparing the power spectra of two time series. In this method, a Z-statistic for the differences between each frequency component of the two time series is calculated, and a p-value is calculated using the jackknife method to test the null hypothesis that the spectral components are identical (Bokil et al., 2007; Miller, 1968).

We used the implementation of this method in the function *two_group_test_spectrum* from the Chronux toolbox (Arvesen, 1969; Bokil et al., 2010) to obtain p-values for all the individual frequencies between 1 and 18 Hz. The mean p-value of all frequency components within a specific frequency band was used as reactivity scoring index, which we call hereafter the TGT:

$$TGT = \frac{1}{K}\sum_{j=q}^{j=q+k} P_j$$

with P_j is the p-value corresponding to a specific frequency component and $j = q,...,q+K$ are the frequency components within a particular frequency band. The TGT was calculated for frequency bands identical to the bands used in the tBSI, including the δ, θ, α and β-bands and the total frequency band.

2.6. Relative entropy

Spectral entropy is a widely used feature to quantify the degree of regularity or organization in a signal, which is based on quantification of the uniformity of power in the power spectrum of a signal. The relative entropy (RE), or Kullback–Leibler divergence, compares the spread of power in the frequency spectra of two signals, thereby quantifying the change in signal organization (Kullbback and Leibler, 1951).

It has been suggested that RE based methods might be useful in various settings of EEG analysis, including the assessment of event related potentials (Rosso et al., 2001) and the detection of epileptic seizures (Quiroga et al., 2000). In the current study, the spectral RE is used to assess the similarity in degree of order between the pre- and post-stimulation EEG segment:

$$RE = -\sum_{j}^{K} S'postj * \log\left(\frac{S'post_j}{S/pre_j}\right) \quad (2)$$

In which $S'post_j$ and $S'post_j$ are respectively the 1-18 Hz mean normalized spectra of the 60 s post- stimulation and pre- stimulation epochs with $j = 1,2...,K$ frequency components.

2.7. Kolmogorov–Smirnov test

The Kolmogorov–Smirnov (KS) test assesses the difference in the distribution of a variable in two populations (Lilliefors, 1967; Massey, 1951). It has been suggested that the KS-test might be used to test the equality of the spectral distribution of two EEG samples, which can be used to verify the stationarity of the EEG (McEwen and Anderson, 1975). In the present study, a two-sample KS-test was used to test the null hypothesis that the distribution of the mean pre-stimulation spectrum is equal to the distribution of the mean post-stimulation spectrum. To obtain these mean spectra, the 1–18 Hz spectrogram corresponding to the pre-and post-stimulation segments is averaged over the time. The p-value of KS-test is used as quantitative score of EEG reactivity.

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

2.8. Peak comparison method

In a recent study, a method that automatically analyzes EEG reactivity in comatose patients was presented (Noirhomme et al., 2014). The method is based on the detection of changes in peaks in the spectra of multiple channels, and we will hereafter refer to this method as the Peak comparison method (PCM). The detection involves a comparison of spectral power of 1 s EEG segments immediately before and after stimulation, performed in multiple channels. Noirhomme et al. reported that this automated method provided substantial agreement with visual reactivity scoring, and a fairly high correlation between automated reactivity scoring and the patient outcome was found.

In the current study, a variation of the PCM is tested for its capability of detecting reactivity. In this variation, we used the preprocessed data of the longitudinal bipolar montage instead of a referential montage and limited the frequency domain to 1–18 Hz instead of >1 Hz. The computational design and all other settings were similar to the originally described approach. Details can be found in the Supplementary Material.

2.9. Statistical analysis

The ability of each qEEG feature to discriminate reactive and non-reactive EEGs was assessed using receiver operating characteristic (ROC) curves, calculated with 500 iterations of repeated random subsampling cross-validation (Hastie et al., 2009). The gold standard label (reactive vs. non-reactive) of each case was determined by classification according to the majority of experts, and only reactive and non-reactive cases were included in the assessment of accuracy; cases determined to exhibit 'unclear reactivity' were excluded.

Cross-validation was performed for all individual qEEG features (TBSI and BSI in the δ, θ, α, β and total frequency band, RE, KS-test and PCM), but also for several combinations of qEEG features. Feature combinations (FC) that were tested involved either 2–4 of the best performing individual features or 2 to 4 individual features together covering the 1–18 Hz frequency range.

In each iteration of cross-validation, classifier training was conducted on a randomly selected subset of N = 10 cases (training data), as sketched in Fig. 1. The classifiers were tested on the 10 remaining cases (test data) to avoid over-estimating performance due to over fitting to idiosyncrasies in the training data (Zhang, 1993).

In training classifiers, the feature values and labels of the training data were used to generate ROC-curves for every individual feature-channel combination. Classification consisted in comparing the feature value in that channel to a threshold. For every feature involved, the channel with the highest area under the curve was selected as the optimal channel, after which a reactivity probability model based on the feature values in these corresponding channels was defined. Hence, classifier training of individual features resulted in univariate model estimating the probability of reactivity using the feature value obtained in the optimal channel. Likewise, classifier training of feature combinations resulted in multivariate models involving multiple features each calculated in one channel (feature-channel set). With this approach, changes of activity in the different frequency bands – which typically have their own spatial distribution in the brain (Young, 2000) – were evaluated in the single channel in which the corresponding feature was most distinctive.

In testing classifiers, i.e. testing the probability model, an ROC for classification of the left-out testing data was obtained, involving classification of testing data according to the modeled reactivity probability values. The area under this curve (AUC) is reported as a measure of classifier performance. Classifier specificity and sensitivity were calculated by setting the classifier threshold to the value that achieved maximum classification accuracy.

A final reactivity probability model was obtained using the best performing feature or feature combination identified in the cross validation analysis, in which the optimal channel(s) and corresponding probability model were calculated using all reactive and non-reactive cases. Hence, no randomization was performed. With this final model, the probability of reactivity was calculated for each case, and the distribution of these probability values in the different case categories was evaluated. To verify to what extent the final probability model is calibrated to the opinions of individual raters, the agreement between the expert scores and the scores determined by the model was assessed for a range of classification thresholds. As a reference, the inter rater agreement among three experts was evaluated. Agreement was evaluated using percent agreement and Gwet's kappa value AC1 (Gwet, 2008).

3. Results

3.1. Cases

A total of 70 EEG recordings were collected. A burst suppression pattern was observed in 2 cases and EMG reactivity was detected in 8 cases, which were all excluded from analysis. One case was excluded due to prominent non-physiological artifacts.

Of the 59 included cases, 18 cases were categorized as reactive and 34 as non-reactive, according to the majority of expert opinions. In 9 reactive cases and 24 non-reactive cases, all experts agreed on the presence of absence of reactivity, while in the other 9 reactive and 10 non-reactive cases only two out of three experts agreed. A total of 7 cases were categorized as unclear, either due to total disagreement between experts or to multiple experts classifying the case as unclear. In total, the experts did not fully agree in 44% of the cases (26 out of 59). The inter-rater agreement for the included cases was 66%, and the corresponding Gwet's AC1 was 53%, indicating moderate overall agreement.

3.2. qEEG

3.2.1. Optimal channels

A complete summary of the AUC values of the qEEG features found in each individual is included the Supplementary Material (Supplementary Figure S1 and S2). The AUC values differed between channels. In addition, the channel in which the AUC was maximal varied between different qEEG features, and was dependent on the set of cases that was randomly selected as training data in the cross validation procedure. In general, features based on the δ-band showed a better performance in the frontal derivations, while performance was most optimal in parietal region for the θ-band, in the posterior temporal area for the α-band and in the occipital derivations for the β-band. For most features, AUC values were higher on the left hemisphere than on the right.

3.2.2. qEEG features

The cross validated median AUC values and corresponding interquartile range (IQR) of the individual qEEG features are summarized in Table 1. Among all single qEEG features evaluated, those based on the tBSI showed the highest median AUC values, with tBSI calculations over the total band showing the best overall performance, followed by tBSIθ, tBSIβ and tBSIδ and tBSIα, respectively. Features based on the TGT were as a group better than features based on RE, the KS-test, and PCM. Table 1 shows a selection of multivariate probability models involving feature combinations,

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

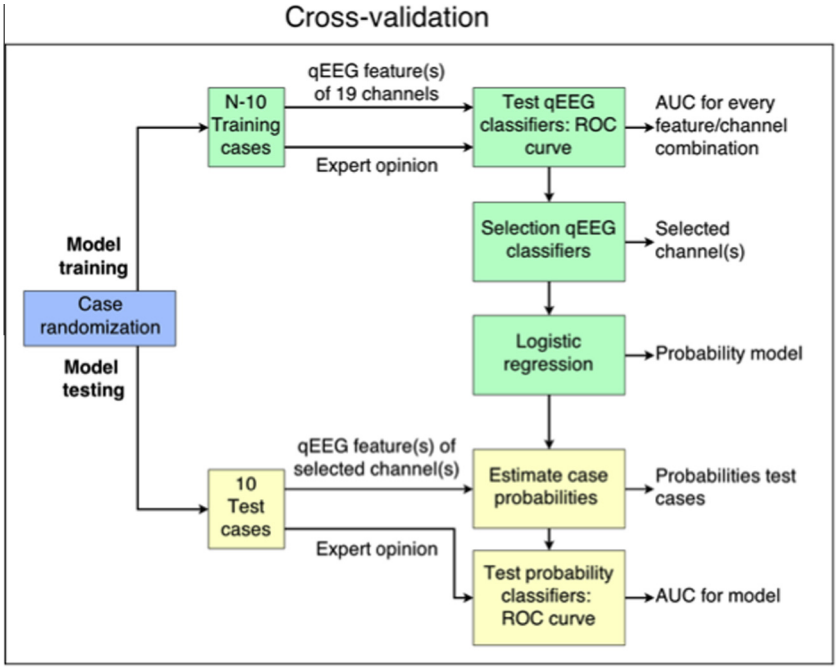

Fig. 1. Schematic overview of the cross-validation iteration, which is performed 500 times for purposes of classifier training and testing. Abbreviations: ROC, receiver operating characteristics; AUC, area under curve; qEEG, quantitative EEG.

Table 1
Accuracy of tested quantitative EEG features.

Feature	Median AUC (IQR)
tBSI total	0.94 (0.83–1.00)
tBSI δ	0.88 (0.76–0.96)
tBSI θ	0.92 (0.83–1.00)
tBSI α	0.88 (0.76–0.96)
tBSI β	0.89 (0.79–1.00)
TGT total	0.89 (0.79–1.00)
TGT δ	0.86 (0.75–0.95)
TGT θ	0.83 (0.72–0.92)
TGT α	0.83 (0.69–0.94)
TGT β	0.88 (0.80–0.94)
RE	0.79 (0.67–0.90)
KS-test	0.78 (0.65–0.88)
PCM	0.54 (0.53–0.56)
FC1 (tBSI total + TGT total)	0.94 (0.84–1.00)
FC2 (tBSI total + TGT total + RE + KS-test)	0.92 (0.83–1.00)
FC3 (tBSI δ + tBSI θ + tBSI α + tBSI β)	0.95 (0.86–1.00)
FC4 (TGT δ + TGT θ + TGT α + TGT β)	0.91 (0.81–1.00)
FC5 (tBSI δ + tBSI θ + tBSI α + tBSI β + TGT δ + TGT θ + TGT α + TGT β)	0.92 (0.81–1.00)

Abbreviations: FC, feature combination; tBSI, temporal brain symmetry index; TGT, two-group test; RE, relative entropy; KS-test, Kolmogorov–Smirnov test; PCM, peak comparison method; AUC, area under the curve; IQR, interquartile range.

with their AUC values obtained from the cross-validation process. No other multivariate models that were developed provided higher AUC values than the FC3 model.

3.2.3. Optimal qEEG model

Of the univariate models, the model based on the tBSI total features performed best, with a median AUC of 0.94 (IQR: 0.83–1.00 and 1–99% percentile range: 0.43-1.00). Overall, the probability model performing best as a classifier of reactivity was based on feature combination FC3 (combining individual tBSI values from the in the δ-, θ-, α- and β-bands). The median AUC of the FC3-based probability models found in the in 500-fold repeated random subset cross-validation, in which models were trained on a randomized set of 10 cases and tested on the left-out cases, was 0.95 (IQR 0.86–1.00 and 1–99% percentile range: 0.50–1.00). The ROC curve and corresponding classifier thresholds of the 500 fold iterations is show in Fig. 2A. In these FC3-models, the prediction of presence of reactivity was associated with a specificity of 86% (IQR: 67–100%) at 100% sensitivity and a sensitivity of 80% (IQR: 50–100%) at 87% specificity.

The FC3-based model that was estimated and tested using all cases classified as reactive or non-reactive by experts (final model) selected the F3-C3, C3-P3, P7-O1 and P3-O1 derivations as the optimal channels to obtain tBSI features of the δ, θ, α and β frequency bands, respectively, according to the AUC values of the individual features. Hence, the FC3 probability model included tBSI δ calculated in the F3-C3, tBSI θ in C3-P3, tBSI α in P7-O1 and tBSI β in P3-O1. This final model was associated with an AUC of 0.95, as calculated using all included reactive and non-reactive cases.

The distribution among different case categories and the probability function corresponding to the final model are summarized in Fig. 2B and C. The probability values – i.e. predicted probability of reactivity – of the reactive cases were higher than those of the non-reactive cases. In addition, the median probability of reactivity was higher in reactive cases for which inter-expert agreement was 3/3than in cases scores as reactive by 2/3 of the raters. The opposite was seen for non-reactive cases; the predicted probability of reactivity was lower for the group classified as non-reactive by all experts compared to those in which 2/3 raters classified the case as non-reactive. For cases in which reactivity was assessed as unclear by the experts, probability values varied between 0.02 and 1.00, but showed a median value of 0.50. In the 34 epochs of unstimulated comatose patients, the model yielded a probability values of 0–0.055, indicating lack of spontaneous reactivity.

Fig. 2D shows the agreement between the individual expert and the final FC3 probability model used as a classifier involving a specific probability threshold, i.e. a threshold that determines from which probability value a case is considered reactive. The maximal agreement was found at a probability threshold of 0.65, accompanied by a Gwet's AC1 of 65% for expert 1, and 70% and 66% for experts 2 and 3 respectively. Using this threshold, classification by the FC3 model was in agreement with the expert's opinion in

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

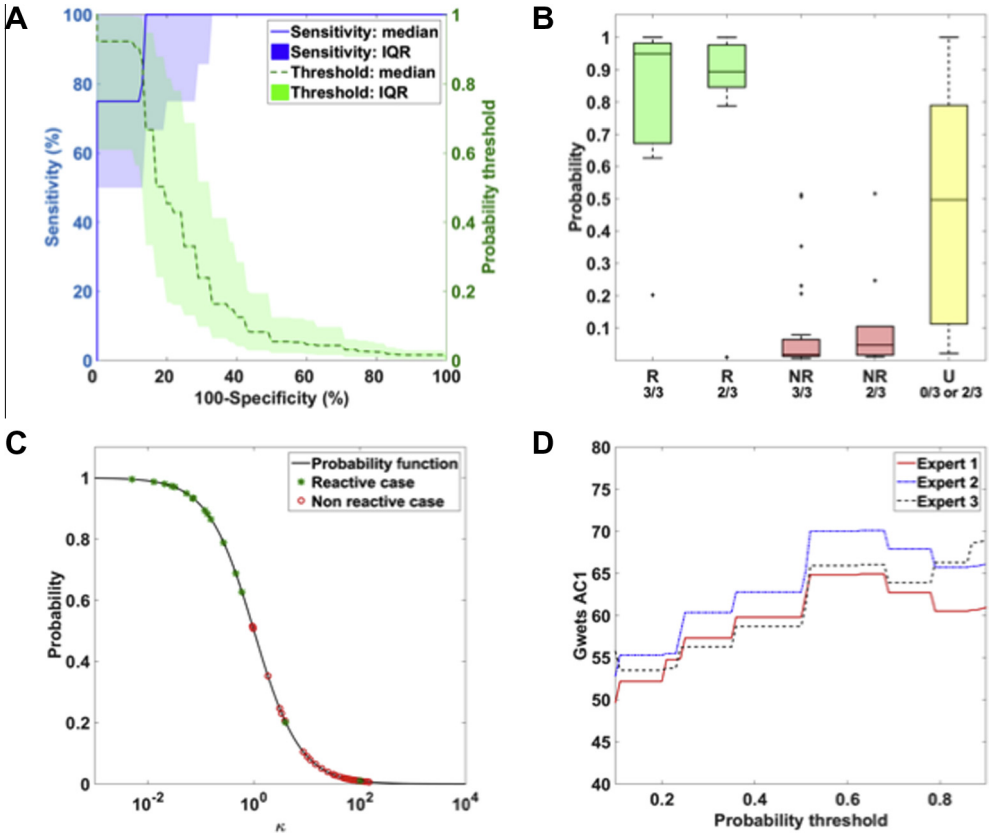

Fig. 2. (A) Receiver operating characteristic (ROC) curves corresponding to the probability function obtained the 500-fold cross-validation. The blue line reflects the median sensitivity of the 500 iterations, with its area reflecting the interquartile (IQR) range. The green line and area reflect the corresponding median probability threshold and IQR range used for classification. (B) Distribution of probabilities based on the FC3 model in corresponding to case categories. Cases were categorized according to the major expert score (R = Reactive, NR = Non-reactive, U = unclear) and the amount of experts that agreed on this score (3/3, 2/3 or 0/3). (C) Reactivity probability curve of final FC3 model. *k* reflects exponential function of the selected set of qEEG features, involved in the probability function $P = 1/1 + k$. Green stars and red circle reflect the cases scored as reactive and non-reactive respectively. (D) Agreement between categorization by the majority of experts' impressions and classifiers of the FC3 probability model involving a specific threshold.

50 out of 52 cases labeled as reactive or non-reactive. These results indicate that the FC3 final model was in substantial agreement with the individual opinion of all experts.

In Fig. 3A–D, the pre- and post-stimulation EEG recording and corresponding spectrogram of four different case examples are shown to illustrate the findings. The EEG recording of case presented in Fig. 3A was labeled as reactive by the experts, and was accompanied by a high probability of reactivity (close to 1) according to the FC3 model. Likewise, a case scored as non-reactive by experts (Fig. 3B) was accompanied by low probability values (close to 0). Hence, in both cases the probability model was in line with the opinion of experts, which was the case for all 50 cases labeled as reactive or non-reactive.

The reactivity probability values calculated by the FC3 model were in disagreement with the experts in two cases, of which one is presented in Fig. 3C. These cases were accompanied by a low probability of reactivity (<0.50), while the experts classified these cases as reactive. Strikingly, in both 'misclassified' cases, no clear spectral changes were observed in the spectrogram, supporting the estimated probability values.

Fig. 3D shows an example of a case in the unclear category, which was accompanied by an intermediate (close to 0.5) probability of reactivity.

4. Discussion

We have developed an algorithm to automatically assess EEG reactivity using several qEEG features. The algorithm relies on a probability model involving single or multiple features quantifying the spectral changes between a pre- and a post-stimulation epoch (spectral symmetry) and was tested against the consensus opinion of three expert EEG readers. The probability model showing best performance in the classification of reactivity was based on tBSI features in separate frequency bands and agreed with visual analysis by experts at least as well as experts agreed amongst themselves.

In the exploration of the performance of qEEG features as classifiers of reactivity, we found that the AUC values obtained from the ROC curves varied between the different features, indicating that some quantitative approaches reflect expert judgment regarding EEG reactivity better than others.

In general, among the univariate models, those based on tBSI emerged as the best classifiers of reactivity, showing median AUC values of 0.88-0.94, closely followed by TGT related features showing median AUC values of 0.83-0.89. Both methods compare the pre- and post-stimulation spectrum for all frequency components individually, and the high AUC values indicate that this approach might be efficient in characterizing EEG reactivity.

Classifier models based on the KS-test and RE related features achieved moderate median AUC values of 0.78 and 0.79 respectively. By design, both parameters are sensitive to alterations in the relative spectral distribution, but insensitive to changes in total power with a constant relative spectrogram, which might explain misclassification of several cases.

Classifiers based on the PCM method achieved low sensitivity and specificity and a median AUC of 0.54. During visual evaluation

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

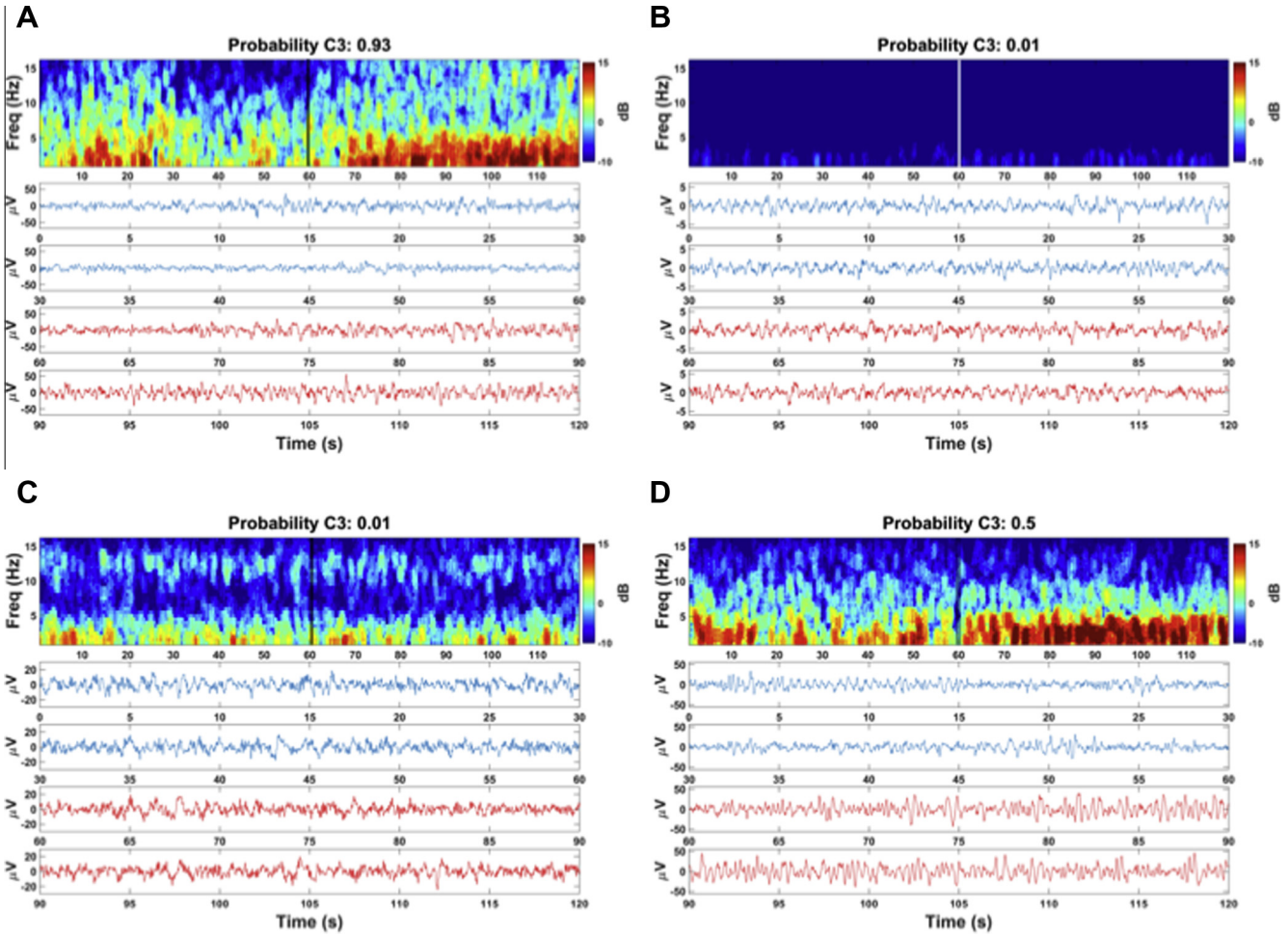

Fig. 3. Examples of EEG recording and corresponding probability of reactivity of cases with different expert labels: (A) Reactive case with corresponding high probability value. (B) Non-reactive case with corresponding low probability values. (C) Reactive case with a discrepant low probability value (misclassification) (D) Unclear case with median probability value. Top frames reflect the spectrogram of the pre- and post-stimulation epoch in Cz-Pz, showing the power (dB) of the frequency components (Freq) in time. Onset of stimulation is at 60s, as indicated by the double vertical lines in the spectrum. Lower curves display the Cz-Pz recording in time, where the blue line (0–60 s) reflects the pre-stimulation epoch and the red line the post-stimulation epoch (60–120 s). In the side panel, the probability calculated using the FC3 final model is shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the case spectra, it was seen that changes in the frequency or amplitude did not always start immediately after the onset of stimulation and that EEG reactivity was often not homogeneous throughout the total post-stimulation period. This might explain the low AUC values for classifiers based on the PCM method, as this approach uses only very short epochs, leading to a higher susceptibility to errors due to irregularities in the EEG and to delay in EEG reactivity.

The channels in which a maximal AUC was found most frequently varied between features based on either the δ, θ, α or β band. This suggests that the reactivity in individual frequency bands is more visible in specific locations of the brain than others. In general, models with good classification performance involved assessment of spectral changes in the frontal channels for the δ-band, in parietal channels for θ-band, in the posterior temporal area for the α-band and in the occipital region for the β-band. Strikingly, reactive and non-reactive cases could generally be distinguished better using EEG derivations situated on the left hemisphere. Although this is an intriguing finding, the differences in AUC between left and right derivations were relatively small and were not seen in all pairs of channels. Therefore, this asymmetry should be taken with caution and requires confirmation.

In general, qEEG models in which all frequency bands are involved showed higher AUC values compared to those involving one, two or three individual bands. This is not surprising since coma can be associated with a variety of EEG patterns involving activity in different frequency bands, depending on its cause (Sutter and Kaplan, 2012; Synek, 1990). The tBSI total model, in which all frequency components of the total frequency range contribute equally, was the best performing univariate model. Yet, the multivariate models FC1 and FC3 performed slightly better according to the AUC values, i.e. median, IQR and 1–99% percentile range, suggesting that it is beneficial to combine information across the δ, θ, α and β-bands in a specific ratio.

The probability functions based on the presented combination of qEEG features (FC3) showed best performance (median AUC value and IQR) and was accompanied with a specificity of 88% for 100% sensitivity. As shown in the illustrative examples of Fig. 3A and B, cases without obvious changes in the spectrograms generally result in low probabilities of reactivity estimated by FC3 while cases with evident spectral changes are assigned high probabilities. Cases in which the spectral changes are less clear were accompanied by variable probability values with an intermediate median value (close to 0.5), likely indicating that this group comprises cases showing 'weak' reactivity (Fig. 3D). Furthermore, the substantial agreement between the assessment by individual experts and the classification based on the FC3 probability model (Fig. 2D) indicates that the model agrees at least as well with the

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

individual experts as the experts agree amongst themselves and shows that the performance of the classifier was not biased by one or two specific reviewers.

Despite the promising AUC values and concordance with the experts' opinion, the probability functions of both individual features and combinations are still accompanied by imperfect specificity relative to overall expert consensus. According to the FC3 final model, two cases were classified as reactive by the experts but were accompanied by a discrepant low probability value (<0.50) indicating absence of reactivity. However, in retrospect evaluation of these cases, the experts did not notice evident spectral changes in the spectrograms (Fig. 3C). This makes it less likely that the final model provided false classification, and suggests that the EEG readers misclassified these cases.

The quantitative methods developed herein were evaluated for their performance as binary classifiers (reactive vs. non-reactive), which matches the current approach in which 'any' change in the EEG is regarded as reactive. Our results however indicate that reactivity is in fact not a binary phenomenon. Indeed, cases where there was a full agreement (3/3) between EEG readers showed more extreme probability of reactivity (i.e. values close to 1 for reactivity or 0 for non-reactivity) than cases where agreement was only partial (2/3) (Fig. 2B). Similarly, unclear cases (that were not used for training the classifier) were characterized by a wide range of probability values intermediate between cases classified as reactive or unreactive. This indicates that degrees of reactivity can be seen and interpreted variably by EEG readers. In practice, presence or absence of reactivity does not appear obvious in a significant minority of cases, as reflected by the substantial number of unclear cases in this study population and by the significant disagreement between experts. This inter-rater variability deserves to be investigated if reactivity testing is to become a reliable prognostic tool and motivated this work. It is likely that intermediary cases will benefit from quantification methods that provide an objective measure of reactivity, rather than depends on subjective assessment. To facilitate quantification adapted to the gradual character of reactivity, a non-binary measure or 'reactivity scale' would be most appropriate.

Altogether, our findings indicate that the quantitative method developed herein provides a proper representation of the expert opinion regarding reactivity, which seems efficient in distinguishing reactive cases from non-reactive cases. In addition, the numerical reactivity probability values might provide a more refined approach or 'scale of reactivity' that might quantity the degree of reactivity as a continuous rather than dichotomous variable.

Although the automated method performed at least as well as visual analysis, which has its own restraints and inaccuracies, the proposed quantitative methods might be accompanied by several limitations when used for prediction of clinical outcome.

First, given the potentially major implication of absence of reactivity, it is desirable that even the smallest or most subtle amount of EEG reactivity is detected. Therefore, the sensitivity of a quantitative parameter for the prediction of the presence of reactivity is critical. Further studies should investigate in greater detail the time course of EEG changes following stimulation and the optimal length of the pre- and post-stimulation epochs for visual and quantitative analysis. Second, non-stationarity of the background rhythm might distort the quantitative assessment, even though the use of 60 s epochs probably averages out most of the background variations and the results of the non-stimulated cases indicate low sensitivity for this phenomenon. In the setting of prognostication, evaluation of the background rhythm might be warranted to optimize the settings of the quantitative assessment of reactivity. Third, the presented models involve features that are based on one single pre-defined channel. Yet, it is plausible the various types of stimuli arouse different reactivity patterns in divergent brain areas or that functionality of the brain is inhomogeneous due to partial injury. With the presented quantitative approach, it is possible that reactive patterns that are merely present in brain areas distant from the selected channels are missed by the model which may lead to unjust quantification. Therefore, it needs to be explored whether use of multiple channels in different areas or separate models for different stimuli types increases sensitivity even more, without affecting specificity. Last, the calculated models give more importance to specific frequency domains than others, in which definition of the boundaries of these frequency bands is imperative. In the current analysis, the frequency range was limited to 1–18 Hz, and the β range that is mostly defined as 14–30 Hz was truncated to minimize effects of EMG activity. Whether this design has clinical implications has to be demonstrated, and further exploration of the most optimal frequency components with most clinical value is warranted.

Another issue that should be taken into account is that the currently proposed method may be inefficient in case artifacts, in particular EMG, are present in the EEG. To prevent erroneous quantification of EEG reactivity, development of methods reducing the influence of artifacts, burst suppression or EMG reactivity is desirable. It might also be beneficial to switch to different quantitative approaches in case burst-suppression is detected. In case of burst-suppression, one might calculate the qEEG features using the non-suppressed periods only. An additional parameter assessing changes in length or frequency of burst and inter-burst intervals could also be implemented.

Although we validated our algorithm using 500 iterations of cross-validation, further studies are required to demonstrate the generalizability of these findings from a small sample to a larger independent set of cases. Another limitation of the present study is that agreement between experts scores based on the total EEG expert recording was only moderate, which is in line with the observation that experts do not always agree on the interpretation of EEG findings(Gerber et al., 2008; Mani et al., 2012). This finding questions the validity of using expert scoring as a gold standard. To strengthen the reliability of expert scores as gold standard in further research, one could consider increasing the number of raters. In addition, utilization of a carefully defined guideline seems to improve agreement between raters in the assessment of various EEG characteristics (Gaspard et al., 2014). Thus it is highly recommended to define a uniform standard for the assessment of EEG reactivity. Furthermore, it is suggested to involve alternative domains to present the EEG, i.e. as a time-frequency or temporal symmetry representation, as this might support visualization of the EEG characteristics and contribute to an objective interpretation (van Putten, 2008). The finding that the inter-rater agreement is far from optimal underlines the subjectivity of visual analysis and stresses the importance of the development of objective quantitative methods as aimed in this present study.

Ultimately, the most important aspect of qEEG features is how well they can predict neurological outcomes of patients, regardless of how well they correlate with the opinion of experts. Therefore, future studies comparing qEEG findings and clinical outcome are desirable.

Yet, there is currently no standardized protocol for the assessment of EEG reactivity. The temporal, spatial and morphological characteristics of the EEG arousal response are not known to be influenced by the method of stimulation (Fischgold et al., 1959). It remains to be demonstrated if different arousal patterns have different prognostic implications and only limited prognostic information might be obtained in cases of hypothermia and sedation. These factors should be investigated in subsequent studies, after which development of a proper testing protocol is warranted to enable proper evaluation of the result of the EEG reactivity test.

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024

In addition in the interpretation of EEG reactivity test, it needs to be considered that a single assessment of EEG reactivity may underestimate the reactive capability of the brain, as intermittent reactivity may occur. Accordingly, absence of a change in the EEG after stimulation does not always imply a bad prognosis, i.e. when this is accompanied by a favorable background pattern representing an 'active' or 'already stimulated' cerebral state prior to reactivity testing. Likewise, not all changes in the EEG are inherently prognostically favorable, e.g. SIRPIDs or the induction of bursts (Alvarez et al., 2013). It should be noted that burst-suppression, and especially burst-suppression with identical bursts in postanoxic coma, is strongly associated with poor outcome (Hofmeijer et al., 2014; Young, 2000; Zandbergen et al., 1998), suggesting that EEG reactivity testing might be superfluous when such a pattern is observed. To ensure correct interpretation of the (quantitative) results of the reactivity test, visual verification of the underlying rhythm during reactivity testing is essential. Besides, it is imperative to obtain multiple epochs per patient to ensure that the outcome is reproducible.

Furthermore, for purposes of prognostication, the outcome of the EEG reactivity test should be interpreted in the light of other test results and findings such as clinical examination, EEG background rhythm, and somatosensory evoked potentials (Oddo and Rossetti, 2014; Rossetti et al., 2010; Tjepkema-Cloostermans et al., 2015; Wijdicks et al., 2006), which all have some predictive value. In this, it is relevant to make a distinction between patients with traumatic brain injury and postanoxic encephalopathy, as the mechanisms involved in cerebral damage and prognoses are quite different.

5. Conclusion

We have shown that EEG features quantifying spectral changes can detect EEG reactivity. Binary classifiers based on these features agreed with the visual assessment of reactivity by experts at least as well as experts agreed among themselves. In addition, the numerical measures provided by the probability model might provide a more refined representation or scale of the quantity of reactivity. Although further validation is needed, these results suggest that quantitative EEG is a useful tool to support visual analysis, potentially improving the objectivity of the EEG reactivity test and assisting in the prediction of clinical outcome in comatose patients.

Conflicts of interest

M.H. was the recipient of a grant from the Dutch Heart Organization. M.B.W. received research support from the National Institute of Health (NIH-NINDS, 1K23NS090900-01), the, and the Andrew David Heitman Neuroendovascular Research Fund.

M.vP. is a co-founder of Clinical Science Systems (www.clinicalscienceystems.com). L.J.H. received support for investigator-initiated studies from UCB-Pharma, Upsher-Smith, Lundbeck., Eisai and Sunovion and consultation fees from Lundbeck, Upsher-Smith, Neuropace, and Allergan. Besides, L.J.H. received honoraria for speaking from Natus and Neuropace. L.J.G. was author of chapters of UpToDate-Neurology and co-author of the book "Atlas of EEG in Critical Care", by Hirsch and Brenner, 2010. N.G. is a Clinical Master Specialist of the Belgian Fund for Scientific Research – FNRS and received honoraria from UpToDate.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.clinph.2015.06.024.

References

Alvarez V, Oddo M, Rossetti AO. Stimulus-induced rhythmic, periodic or ictal discharges (SIRPIDs) in comatose survivors of cardiac arrest: characteristics and prognostic value. Clin Neurophysiol 2013;124(1):204–8.
Arvesen JN. Jackknifing U-statistics. Ann Math Stat 1969;40(6):2076–100.
Bokil H, Purpura K, Schoffelen J-M, Thomson D, Mitra P. Comparing spectra and coherences for groups of unequal size. J Neurosci Methods 2007;159(2):337–45.
Bokil H, Andrews P, Kulkarni JE, Mehta S, Mitra PP. Chronux: a platform for analyzing neural signals. J Neurosci Methods 2010;192(1):146–51.
Claassen J, Hirsch LJ, Kreiter KT, Du EY, Sander Connolly E, Emerson RG, et al. Quantitative continuous EEG for detecting delayed cerebral ischemia in patients with poor-grade subarachnoid hemorrhage. Clin Neurophysiol 2004;115(12):2699–710.
Delorme A, Makeig S, Sejnowski T. Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. In: Proceedings of the International workshop on ICA. San Diego, CA; 2001.
Fischgold H, Mathis P. Obnubilations, comas et stupeurs. Masson Paris; 1959.
Gaspard N, Hirsch LJ, LaRoche SM, Hahn CD, Westover MB. Interrater agreement for Critical Care EEG Terminology. Epilepsia 2014;55(9):1366–73.
Gävert H, Hurri J, Särelä J, Hyvärinen A. FastICA for Matlab, version 2.5 [Internet]. 2005 [cited 2014 Sep 5]. Available from: http://research.ics.aalto.fi/ica/fastica/.
Gerber PA, Chapman KE, Chung SS, Drees C, Maganti RK, Ng Y, et al. Interobserver agreement in the interpretation of EEG patterns in critically ill adults. J Clin Neurophysiol 2008;25(5):241–9.
Goncharova II, McFarland DJ, Vaughan TM, Wolpaw JR. EMG contamination of EEG: spectral and topographical characteristics. Clin Neurophysiol 2003;114(9):1580–93.
Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008;61(1):29–48.
Hastie T, Tibshirani R, Friedman J, Hastie T, Friedman J, Tibshirani R. The elements of statistical learning. New York: Springer; 2009.
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(5):947–54.
Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat 1951;22(1):79–86.
Lilliefors HW. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. J Am Stat Assoc 1967;62(318):399–402.
Lodder SS, van Putten MJAM. Quantification of the adult EEG background pattern. Clin Neurophysiol 2013;124(2):228–37.
Logi F, Pasqualetti P, Tomaiuolo F. Predict recovery of consciousness in post-acute severe brain injury: the role of EEG reactivity. Brain Inj 2011;25(10):972–9.
Mani R, Arif H, Hirsch LJ, Gerard EE, LaRoche SM. Interrater reliability of ICU EEG research terminology. J Clin Neurophysiol 2012;29(3):203–12.
Massey FJ. The Kolmogorov–Smirnov test for goodness of fit. J Am Stat Assoc 1951;46(253):68–78.
McEwen JA, Anderson GB. Modeling the stationarity and gaussianity of spontaneous electroencephalographic activity. IEEE Trans Biomed Eng 1975;22(5):361–9.
Miller RG. Jackknifing variances. Ann Math Stat 1968;39(2):567–82.
Noirhomme Q, Lehembre R, del Rosario Lugo Z, Lesenfants D, Luxen A, Laureys S, et al. Automated analysis of background EEG and reactivity during therapeutic hypothermia in comatose patients after cardiac arrest. Clin EEG Neurosci 2014;45(1):6–13.
Nuwer M. Assessment of digital EEG, quantitative EEG, and EEG brain mapping: report of the American Academy of Neurology and the American Clinical Neurophysiology Society. Neurology 1997;49(1):277–92.
Oddo M, Rossetti AO. Early multimodal outcome prediction after cardiac arrest in patients treated with hypothermia. Crit Care Med. LWW 2014;42(6):1340–7.
Quiroga RQ, Arnhold J, Lehnertz K, Grassberger P. Kulback–Leibler and renormalized entropies: applications to electroencephalograms of epilepsy patients. Phys Rev E 2000;62(6):8380–6.
Rossetti AO, Oddo M, Logroscino G, Kaplan PW. Prognostication after cardiac arrest and hypothermia: a prospective study. Ann Neurol 2010;67(3):301–7.
Rosso OA, Blanco S, Yordanova J, Kolev V, Figliola A, Schürmann M, et al. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. J Neurosci Methods 2001;105(1):65–75.
Sutter R, Kaplan PW. Electroencephalographic patterns in coma: when things slow down. Epileptologie 2012;29(8):201–9.
Synek VM. Value of a revised EEG coma scale for prognosis after cerebral anoxia and diffuse head injury. Clin EEG Neurosci 1990;21(1):25–30.
Thomson DJ. Spectrum estimation and harmonic analysis. Proc IEEE 1982;70(9):1055–96.
Tjepkema-Cloostermans MC, Hofmeijer J, Trof RJ, Blans MJ, Beishuizen A, van Putten MJAM. Electroencephalogram predicts outcome in patients with postanoxic coma during mild therapeutic hypothermia. Crit Care Med 2015;43(1):159–67.
Van Putten MJAM. Extended BSI for continuous EEG monitoring in carotid endarterectomy. Clin Neurophysiol 2006;117(12):2661–6.
Van Putten MJAM. The colorful brain: visualization of EEG background patterns. J Clin Neurophysiol 2008;25(2):63–8.
Vespa PM, Nuwer MR, Juh C, Alexander M, Nenov V, Martin N, et al. Early detection of vasospasm after acute subarachnoid hemorrhage using continuous EEG ICU monitoring. Electroencephalogr Clin Neurophysiol 1997;103(6):607–15.

Wijdicks EFM, Hijdra A, Young GB, Bassetti CL, Wiebe S. Practice parameter: prediction of outcome in comatose survivors after cardiopulmonary resuscitation (an evidence-based review) report of the quality standards subcommittee of the American Academy of Neurology. Neurology 2006;67(2):203–10.

Young GB. The EEG in coma. J Clin Neurophysiol 2000;17(5):473–85.

Young GB, McLachlan RS, Kreeft JH, Demelo JD. An electroencephalographic classification for coma. Can J Neurol Sci 1997;24(4):320–5.

Zandbergen EGJ, de Haan RJ, Stoutenbeek CP, Koelman JHTM, Hijdra A. Systematic review of early prediction of poor outcome in anoxic-ischaemic coma. Lancet 1998;352(9143):1808–12.

Zhang P. Model selection via multifold cross validation. Ann Stat 1993;21(1):299–313.

Please cite this article in press as: Hermans MC et al. Quantification of EEG reactivity in comatose patients. Clin Neurophysiol (2015), http://dx.doi.org/10.1016/j.clinph.2015.06.024