Serveur Académique Lausannois SERVAL serval.unil.ch

Author Manuscript
Faculty of Biology and Medicine Publication

This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Published in final edited form as:

serval
serveur académique lausannois

Unil
UNIL | Université de Lausanne
Faculty of Biology and Medicine

CHUV

EEG synchronization measures are early outcome predictors in comatose patients after cardiac arrest

Frédéric Zubler[a,*], Andreas Steimer[a], Rebekka Kurmann[a], Mojtaba Bandarabadi[a], Jan Novy[b], Heidemarie Gast[a], Mauro Oddo[c], Kaspar Schindler[a], Andrea O. Rossetti[b]

a. Department of Neurology, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

b. Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

c. Department of Intensive Care Medicine, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

* Corresponding author:

Frédéric Zubler, MD, PhD
Sleep-Wake-Epilepsy-Center
Department of Neurology
Inselspital - Bern University Hospital
Freiburgstrasse 4
3010 Bern
Switzerland
E-mail: frederic.zubler@gmail.com

Abstract

Objective: Outcome prognostication in comatose patients after cardiac arrest (CA) remains a major challenge. Here we investigated the prognostic value of combinations of linear and non-linear bivariate EEG synchronization measures.

Methods: 94 comatose patients with EEG within 24h after CA were included. Clinical outcome was assessed at 3 months using the Cerebral Performance Categories (CPC). EEG synchronization between the left and right parasagittal, and between the frontal and parietal brain regions was assessed with 4 different quantitative measures (delta power asymmetry, cross-correlation, mutual information, and transfer entropy). 2/3 of patients were used to assess the predictive power of all possible combinations of these eight features (4 measures x 2 directions) using cross-validation. The predictive power of the best combination was tested on the remaining 1/3 of patients.

Results: The best combination for prognostication consisted of 4 of the 8 features, and contained linear and non-linear measures. Predictive power for poor outcome (CPC 3-5), measured with the area under the ROC curve, was 0.84 during cross-validation, and 0.81 on the test set. At specificity of 1.0 the sensitivity was 0.54, and the accuracy 0.81.

Conclusion: Combinations of EEG synchronization measures can contribute to early prognostication after CA. In particular, combining linear and non-linear measures is important for good predictive power.

Significance: quantitative methods might increase the prognostic yield of currently used multi-modal approaches.

Highlights

- Predicting neurological outcome after cardiac arrest remains a challenging task.
- Bivariate EEG synchronization measures can contribute to early prognostication.
- Further studies are needed to evaluate the place of quantitative EEG within multi-modal prognostic algorithms.

Keywords

quantitative EEG, synchronization, prognostication, anoxic-ischemic encephalopathy.

Abbreviations

AP = anterior-posterior axis, AUC = area under the ROC curve, CA = cardiac arrest, CC = cross-correlation, CPC = Glasgow-Pittsburg Cerebral Performance Categories, LR = left-right axis, MI = mutual information, qEEG = quantitative electroencephalography, RDP = relative delta power, ROC = receiver operating characteristic curve, TE = transfer entropy. TTM = targeted temperature management.

1. Introduction

Prognostication in comatose patients after cardiac arrest (CA) remains one of the biggest challenges for a neurologist in the intensive care unit (Rossetti et al., 2016). Current clinical decisions are based on a multi-modal approach including several clinical and para-clinical tests, one of the most important being Electroencephalography (EEG; Horn et al., 2014; Rossetti et al., 2016; Sandroni et al., 2014). Several EEG patterns during hypothermia or controlled normothermia (summarized as targeted temperature management, TTM) have been associated with an unfavourable outcome (Hofmeijer et al., 2015, 2014; Sivaraju et al., 2015; Westhall et al., 2016), whereas an early, continuous and reactive EEG may herald a favourable outcome (Hofmeijer et al., 2015; Sivaraju et al., 2015; Tsetsou et al., 2013). However, clinical EEG also has several limitations. Firstly, it requires the assistance of a trained specialist for interpretation (Spalletti et al., 2016; Taccone et al., 2014). Secondly, it lacks objectivity as inter-rater agreement remains imperfect despite the attempt to propose standardized interpretations (Foreman et al., 2016; Halford et al., 2015; Hirsch et al., 2013; Ng et al., 2015; Westhall et al., 2016). In particular, the classification of visual EEG patterns into unfavourable, intermediate, or favourable categories varies between studies (Hofmeijer et al., 2015; Sivaraju et al., 2015; Westhall et al., 2016, 2015). Computer/ algorithm-based analysis of EEG (quantitative EEG; qEEG) appears as a promising approach to circumvent these limitations.

Several univariate qEEG methods were used to assist visual interpretation, providing for instance compact representations of amplitude or spectrum of EEGs, allowing for rapid identifications of segments where the EEG signal changes. These methods have been used successfully for prognostication after CA (Moura et al., 2014; Oh et al., 2015; Rundgren et al., 2010). Various qEEG measures have been used to refine EEG patterns classifications, such as generalized periodic discharges (Ruijter et al., 2015), similarity of bursts in burst-suppression, burst-suppression ratio and epileptiform activity (Wennervirta et al., 2009), or reactivity (Noirhomme et al., 2014). Five different uni- and multivariate qEEG features were combined into a single index ("cerebral recovery index") to mimic the way neurologists visually interpret EEG, serving as "surrogate electroencephalographers" (Tjepkema-Cloostermans et al., 2013).

Bivariate synchronization measures are classical tools for quantitative EEG analysis in the context of epilepsy or neurodegenerative diseases, where they are often used to define functional networks (Bullmore and Sporns, 2009; Stam and van Straaten, 2012). However, they have only been applied in very few studies for prognostication after CA. Coherence, for instance, was one of the elements of the cerebral recovery index mentioned above. In another study, a bivariate measure based on similarity of the power spectrum of EEG signals was used to define a functional graph, the properties of which were different according to the clinical outcome (Beudel et al., 2014). A previous study has shown a potential value of combinations of bivariate measures for clinical assessment in a heterogeneous population of comatose patients due to various aetiologies (Zubler et al., 2016). Here we set out to investigate the value of combinations of bivariate quantitative EEG measures as an early prognostic marker in a prospectively collected cohort of comatose patients after CA, postulating that this approach would have a good performance in discriminating patients with good from those with unfavourable prognosis.

2. Materials and Methods

2.1. Patients and Treatment

This cohort was recruited at the Department of Intensive Care Medicine of the Lausanne University Hospital (CHUV), Switzerland, and is part of a prospective registry containing details of neurological examination (brainstem reflexes, motor reaction, and presence of myoclonus), electroencephalographic features (reactivity, continuity, epileptiform activity),

somatosensory evoked potentials, and neuron-specific enolase. For details, please see (Oddo and Rossetti, 2014). The study was approved by the Ethical Committee of the Canton of Vaud. Waiver of consent was granted since the EEGs and clinical information were recorded as part of clinical routine. Consecutive comatose patients admitted in the CHUV Intensive Care Unit from September 2012 to February 2016 after cardiac arrest (CA) and not deceased after 48 hours were included. The detailed treatment protocol has been described elsewhere (Oddo and Rossetti, 2014; Rossetti et al., 2010). In short, all patients received TTM, either hypothermia (target temperature 33 °C) or, increasingly since July 2014, controlled normothermia (target temperature 36°C). TTM was induced as soon as possible using ice packs and ice-cold isotonic solutions, followed by the application of a surface cooling device with automatic temperature control maintained for 24 hours. During this time, sedation with midazolam (0.1 mg/kg/h) or propofol (less frequently; 2-3 mg/kg/h), and fentanyl (1.5 µg/kg/h) infusions was provided; vecuronium (0,1mg/kg), rocuronium (0.6-0.7 mg/kg) or cisatracurium (0.15-0.2 mg/kg) boluses were administered for shivering. Patients with myoclonus and/or electrographic status epilepticus were treated with intravenous seizure suppressive drugs (mainly levetiracetam and valproic acid). The decision to withdraw intensive care support was taken interdisciplinary after at least 72h, based on the occurrence of at least two of the following criteria: unreactive EEG background after TTM, treatment-resistant myoclonus and/or electrographic status epilpeticus, bilateral absence of N20 in somatosensory-evoked potentials after NT/HT, absence from at least one of the three principal brainstem reflexes (pupillary, oculocephalic, and corneal, examined after sedation weaning) (Rossetti et al., 2010). In particular, the EEG during TTM was not taken into account for these decisions.

The neurological outcome at 3 months was prospectively assessed through a semi-structured telephone interview using the Glasgow-Pittsburg Cerebral Performance Categories (CPC)(Booth et al., 2004). Good neurological outcome was defined as CPC 1 (complete recovery) or 2 (moderate disability); poor outcome was defined as CPC 3 (severe disability), 4 (vegetative / unresponsive wakefulness) or 5 (deceased).

2.2. EEG recordings

Video-EEGs (Viasys Neurocare, Madison, WI) recording was performed for 20-30 minutes during TTM with 19 electrodes according to the international 10:20-system, with reference

placed near FpZ. The sampling rate of most EEGs was 250Hz; three recordings with original sampling rate from 1000Hz were down-sampled to 250Hz. From each recording, five minutes (the 30 first 10-second epochs without artifact or patient stimulation, and with closed eyes) were exported for quantitative analysis. Concerning muscular artefacts the following rules were applied: the EEG was excluded if the amplitude of the muscle artefacts after band pass filtering (0.5-20Hz, see below) exceeded 10 % of the averaged peak-to-peak amplitude, as judged by visual analysis. In case of burst-suppression pattern with superimposed muscular artefacts, the EEG was excluded if more than 20% of epochs contained no burst (because of the very low signal-to-noise ratio during suppression epochs). Epoch selection and EEG exclusion were performed prior to quantitative analysis and blind to clinical outcome by two board-certified electroencephalographers (FZ and RK).

2.3 Quantitative EEG analysis

We used four bivariate qEEG measures to characterize the synchronization between the left and right parasagittal (left-right axis, LR), and between the frontal and parietal brain regions (antero-posterior axis, AP) (Zubler et al., 2016). A bipolar derivation was used to represent each brain region: (F3-P3) for the left hemisphere, (F4-P4) for the right hemisphere, (F3-F4) for the anterior brain region, (P3-P4) for the posterior brain region (P3-P4). EEG synchronization was computed between the corresponding bipolar derivations (between F3-P3 and F4-P4, and between F3-F4 and P3-P4), independently for each 10-second epoch, and then averaged over all epochs. Except for relative delta power, the EEG was band-passed filtered between 0.5 and 20 Hz prior to analysis, which was performed with Matlab 2014a and 2016a (Mathworks).

Relative Delta Power asymmetry. The first signal coupling measure was based on spectral analysis, namely on (the asymmetry of) relative delta power (RDP). RDP for each derivation was defined as the ratio of power in the delta band (0.5 - 4 Hz) to the total power in the frequency interval 0.5 - 20Hz, computed with the Matlab built-in function *pwelsh.m*. RDP was used to define two asymmetry indices in the left-right axis, $RDP_{LR} = |(RDP_L - RDP_R) / (RDP_L + RDP_R)|$, and in the anterior-posterior axis, $RDP_{AP} = (RDP_A - RDP_P) / (RDP_A + RDP)$. The vertical bar "|" denotes absolute value. In the LR-axis we considered the absolute value of the asymmetry, so that the hemisphere (left or right) with more delta power was not relevant. By contrast, in the antero-posterior axis, we considered the signed value, so that a frontal or

posterior delta dominance was not equivalent. This choice was motivated by the LR-symmetry and AP-asymmetry of power spectrum that is found in healthy subjects ("posterior rhythm").

Cross-Correlation. The second measure was the (zero-lag) cross-correlation, which quantifies the similarity up to scaling factor of two signals. Cross-correlation between the left and right (CC_{LR}) and between the anterior and posterior regions (CC_{AP}) was computed for each 10-second epoch with the Matlab function *corrcoef.m*. Since anti-correlation also implies synchronization mechanisms, we considered only absolute-valued CC (in both axes).

Amplitude-discretization. The next two measures are information theoretical measures, the calculation of which is based on statistical properties of the EEG signals. In order to estimate these statistical properties, we considered a discretized version of the signal. An amplitude discretization with 6 bins was used to estimate the probability distributions of the signal (Steimer et al., 2015). The amplitude discretization chosen for the present study is more robust to noise (especially low-amplitude muscular artefacts) than the bitstring symbolization we used previously (Zubler et al., 2016).

Mutual Information. Mutual information quantifies the uncertainty reduction about one signal if the value of another signal is known. Practically, normalized mutual information between signals x and y was computed as $[H(x) + H(y) - H(x,y)]/H(x,y)$, where $H(x)$ denotes the entropy of the probability distribution over the (6-bin) amplitude of signal x, and $H(x,y)$ the entropy of the joint distribution. It was computed between the left and right (MI_{LR}), and the frontal and parietal regions (MI_{AP}).

Transfer entropy is a directed measure of the information flow between two time series, which quantifies the uncertainty reduction about the future state of a signal (based on its current state), if information about the current value of another signal is also taken into account. Transfer entropy from x to y was also computed from block entropies as $[H(x,y) - H(y^f,x,y)+H(y^f,y)-H(y)] / [H(y^f,y)-H(y)]$ where y^f is signal y taken with a negative delay ("future of y"). We chose a delay of 6 time steps = 0.024 seconds, because 1/0.024 s = 42 Hz is approximately two times the fastest frequency of our filtered signals. For justification of the denominator in the computation of transfer entropy see (Marschinski and Kantz, 2002). Transfer entropy was computed in both axis in two directions, and used to define two asymmetry indices $TE_{LR} = |(TE_{L\ to\ R} - TE_{R\ to\ L}) / (TE_{L\ to\ R}+TE_{R\ to\ L})|$, and $TE_{AP} = (TE_{A\ to\ P} - TE_{P\ to\ A}) / (TE_{A\ to\ P}+TE_{P\ to\ A})$. As for RDP, we considered the absolute value in the left -right axis,

and the signed value in the anterior-posterior axis. Here, also, the justification was the asymmetry observed in healthy subjects (Lee et al., 2009).

2.4. Bayes classifier

The primary aim was to discriminate between good (CPC 1 or 2) and poor (CPC 3-5) functional outcome at three months.

These four synchronization measures computed along two axes provided us with a vector containing eight numerical values (in the following called "features") for each EEG (Fig. 1). Equivalently, an EEG can be represented by a single data point in an 8-dimensional feature space. Our goal was to identify regions of this space associated with a good outcome, and regions associated with a poor outcome (the two "classes"). To this end, we used a generative model based on a mixture of two Gaussians, that is, a superposition of two multi-variate Gaussians approximating the repartition of EEGs of each class (Bishop, 2006). Typically, for this type of classification problem, part of the data (the training set) is used to adjust parameters of the Gaussians ("training the model"), whereas the predictive power of the trained model is tested on the rest of the data (the test set). However, taking all eight features into considerations is not necessarily the optimal solution, since inclusion of irrelevant or redundant features can potentially weaken the model. Therefore, the training set was used not only for training, but also for identifying the best feature combination, using leave-one-out cross-validation (Bishop, 2006). That is, we selected randomly 2/3 of patients to form a training/cross-validation set. On these patients, the predictive power of all 2^8-1 (=255) possible feature combinations was assessed using the area under the ROC curve (AUC). The combination reaching the best AUC during leave-one-out cross-validation was then used in an "optimal" classifier. This classifier was trained on all patients of the training/cross-validation set, and then applied to the test set, namely to the 1/3 of EEGs that were never exposed to a classifier before. Of note, we used two Gaussians per classes because it is the smallest number allowing multimodal classification (for CC_{LR} and MI_{LR} extreme high and low values were both associated with unfavourable outcome). For the implementation details of the Bayes classifier, see Supplementary Material in (Zubler et al., 2016).

The mixture of Gaussians classifier attributed a probability of being associated with a good vs. bad outcome to each element. For practical use, however, a probability threshold for classifying a given subject as having a good or poor outcome has to be chosen, which

corresponds to a specific point (the so-called operating point) on the ROC curve(Noirhomme et al., 2015). Because of the unacceptable ethical cost of false positive prediction of poor outcome, we set the operating point to the highest threshold with specificity (and thus positive predictive value) of 1. At this point we computed the sensitivity, the negative predictive value and the accuracy with 95% confidence interval (CI) using bootstrapping (bias corrected and accelerated percentile method) with 1000 samples.

2.5. Statistics

Between groups comparisons were assessed with a Mann–Whitney U test for numerical data, and with a chi-squared test for categorical data (implemented in Matlab).

3. Results

3.1 Demographic information

Over the recruitment period, 114 consecutive patients were included in the CHUV registry; 18 patients had to be excluded because of low signal-to-noise ratio due to muscle or ECG artefacts on the EEG, resulting in the inclusion of 94 patients for this analysis (34 females). The mean age (± sd) of the analysed cohort was 60.9 years ± 16.1; 46 patients (49%) patients had a poor functional outcome at three months, of whom 3 had CPC 3 and 43 died. Differences between the patients with favourable and unfavourable outcome are summarized in Table 1. Note that for 13 patients (14%), the sedation at recording time was not recorded.

Excluded subjects had a higher proportion of EEGs with suppressed background (9/18 (50%) suppressed only, and 5/18 (18%) burst suppression vs. 5/94 (5%) and 10/94 (10.6%) in the included cohort respectively). They also differed significantly in term of functional outcome (15/18 = 83% patients had a poor outcome, p < 0.01).

3.2 Feature selection with cross-validation

To form the training/cross-validation set, 63/94 (67%) patients were randomly selected. The prognostic value (AUC) of all 255 possible features combinations was computed on this set using leave-one-out cross-validation. The 10 best feature combinations are listed in Table 2.

The best performance was obtained by the combination RDP_{AP}, CC_{LR}, MI_{LR}, and TE_{AP}, which reached an AUC of 0.84 (Fig. 2A). Several other combinations reached an almost equally good prediction value.

The feature CC_{LR} appeared in all 10 best combinations (this feature was even part of the 27 best combinations). CC_{LR} was also the synchronization measure with the best predictive power when considered alone (AUC = 0.72).

3.3 Performance on the test set

The combination that reached the highest AUC during cross-validation (RDP_{AP}, CC_{LR}, MI_{LR}, and TE_{AP}) was used to design a final classifier, which was trained on all 63 patients of the cross-validation group. This classifier was then applied to the remaining 31 (33%) patients forming the test set, i.e. the group of patients who were never involved in feature selection or classifier training. The ROC of the classifier applied to the test set is shown in (Fig. 2B); the AUC was 0.81.

At the largest threshold with specificity and positive predictive value (PPV) of 1.0 for poor outcome, the sensitivity was 0.54 (CI: 0.27-0.83). The negative predictive value (NPV) was 0.76 (CI: 0.54-0.88). This operating point also corresponded to the highest accuracy of all points on the ROC curve (0.81, CI: 0.65-0.94).

To verify whether the results of the final classifier depended on the initial random partition of patients into cross-validation vs. test set, we tested the performance of the same feature combination on 100 different random partitions (63 patients for training, 31 for testing). The mean AUC (± sd) for these 100 trials was 0.82 ± 0.07.

4. Discussion

In this study analysing a prospectively collected cohort, we demonstrate that combinations of four synchronization measures derived from EEG recorded during early postanoxic coma, under temperature management and sedation, and computed between two bipolar derivations (the left and the right hemisphere, and the anterior and posterior regions), represent early predictors of clinical outcome.

In a previous study using a similar approach, but on a heterogeneous population not including the present cohort of patients, recorded in "real-world" conditions (EEG only upon request

from the treating physician, different timing of EEG, different sedation levels), the optimal feature combination for the subgroup of patients with post-hypoxic encephalopathy was very similar to the one found in the present study, namely CC_{LR}, MI_{LR} , and TE_{AP} (3rd best combination in the present study, Table 2). For the whole population (comatose patients with different aetiologies), by contrast, the best combination was different (Zubler et al., 2016). Even if the present study includes data from a single center, the reproducibility in feature selection between two different studies conducted in different hospitals, strongly supports the robustness of our method for prognostication in CA comatose patients.

This result is relevant because it corroborates the potential role of quantitative methods in this setting. It also further supports the role of EEG in the early phase (during TTM, under sedation), for which there is a growing body of evidence (Hofmeijer et al., 2015, 2014; Oddo and Rossetti, 2014; Sivaraju et al., 2015; Tzovara et al., 2016), but which has not yet been incorporated in official recommendations (Sandroni et al., 2014).

4.1. Sensitivity and specificity for practical application

The similarity of the ROC-curves on the cross-validation and the test-set (Fig. 2) are suggestive for a good generalization ability of the classifier.

Both curves had a segment along the x = 0 line (specificity of 1.0 for unfavorable outcome): for a sensitivity of 0.5, the specificity is 0.93 for the cross-validation set, 1.0 for the test set.

By contrast, both ROC curves failed to reach the y = 1 line (100% sensitivity for poor outcome) before the last point. In the test set, this was due to a single subject who was incorrectly classified as having a good outcome. The EEG of this subject was continuous and reactive (Patient C, Fig. 1C), and the somatosensory evoked potentials were present, whereas the brainstem reflexes were absent, so that this case would have been be very challenging also for experienced clinicians.

In summary, the specificity is better than the sensitivity (for poor outcome). For clinical application this property is reassuring, since it is of utmost importance to avoid false-positive predictions of poor prognosis. Of note, the ROC curves resulting from an early multimodal prognostication approaches on an overlapping cohort exhibited similar properties (Oddo and Rossetti, 2014). At least two reasons could explain this phenomenon. First, it could be that current diagnostic methods, including scalp EEG, fail to detect the gravity of post-hypoxic

encephalopathy in specific cases. Another possibility is that the death of patient occurs due to another cause than the first hypoxic brain damage (such as an infection, a multi-organ failure, or a second cardiac event), which might not be foreseeable at the time of the early EEG recording.

Nevertheless, our approach has an overall good accuracy. On the test set, at specificity and PPV of 1.0 for poor outcome, 81% subjects of the test set were correctly classified (accuracy). Moreover, at this operating threshold, the NPV for poor outcome was 0.76, meaning that a little more than 3/4 subjects who were classified as good outcome did indeed have a good outcome (PPV for good outcome). By comparison, in the present cohort, the PPV for poor outcome for classical EEG features, namely absence of EEG reactivity and "highly malignant pattern" (Westhall et al., 2016) was 0.96 and 1.0 respectively, whereas the NPV for these features was 0.68 and 0.60. Our PPV for good outcome also compares favorably with other recently described good outcome predictors such as motor reaction to pain (Rossetti et al., 2016).

4.2. Linear vs. non-linear measures

CC_{LR} and MI_{LR} were the measures that appeared the most frequently in effective feature combinations (Table 2). CC_{LR} was also the measure with best predictive power when considered alone (AUC 0.72). Extreme high and low values of inter-hemispheric cross-correlation were both associated with unfavourable outcome. The former (increased correlation) is found typically in burst suppressions and generalized periodic discharges (GPDs), which are associated with bad outcome (Westhall et al., 2016). On the other hand, one can postulate that very low inter-hemispheric correlation can be caused inter alia by severe diffuse neuronal lesions, which would also be associated with poor functional outcome.

MI_{LR} can be viewed as the non-linear pendant of CC_{LR}. Considered alone, it had a low predictive power, barely above chance level (AUC 0.56). It is commonly accepted that non-linear measures are less effective than linear ones in detecting linear relations in non-stationary signals (such as burst-suppression patterns) (Pereda et al., 2005).

Nonetheless, MI_{LR} was associated to CC_{LR} in 9/10, and in 24/27 best combinations. More generally, the best combination with exclusively linear or non linear measures was only found

at rank 53 (MI_{LR} and TE_{AP}). The reason could be that taken together, these measures allow differentiating linear (detected by both type of measures) from non-linear (only detected by non-linear measures) effects. Interestingly, the presence of purely non-linear effects have been considered a marker of pathology in EEG activity (Andrzejak et al., 2011; Schindler et al., 2016).

An alternative explanation is that combining different measures better encompass the spectrum of lesions (bilateral hemispheric and/or impaired arousal system) resulting in coma in anoxic-ischemic encephalopathy. The best performance was obtained with a combination of all four measures, containing linear (RDP, CC) and nonlinear measures (MI, TE), along two different axes (left-right and anterior-posterior). That is, a large variety of information was used. It is thus likely that additional measures would further improve the prognostic yield. Also, incorporating univariate measures for assessing amplitude and continuity, as was done for the cerebral recovery index (Tjepkema-Cloostermans et al., 2013) would allow including more low-voltage EEGs.

Another difference between the quantitative measures is that CC, MI, and TE are defined in the time domain, whereas RDP characterizes the signals in the frequency domain. RDP is a relatively "basic" measure (and more complex measures exist, such as coherence). We have nevertheless included RDP, first because of the importance of generalized or focal delta slowing in EEG, and second because it has been used successfully in hypoxic (vascular) brain lesions (Claassen et al., 2004). RDP is not a classical measure of synchronization implying a convergence of time series (Jiruska et al., 2013), but rather a reflection of the (a)symmetrical functioning of the brain.

4.3. Strengths and Limitations

This study was based on a relatively large cohort from a prospective registry, implying a high quality of ascertainment. Importantly, recordings were done on clinical, not on specifically performed high-density EEGs. In 13 patients the sedation amount was missing, this value is however unlikely to be systematically biased. As such, no significant differences in sedation or EEG timing between patients with favourable or unfavourable outcome can account for the present findings.

Excluded patients, by contrast, had a higher ratio of poor functional outcome. The reason was that EEGs with very low-voltage segments, which are more likely to be excluded because of low signal-to-noise ratio, are generally associated with unfavourable outcome . However, several EEGs with suppressed background (with or without superimposed discharges or bursts) were included in the analysis. One example is reproduced in (Fig. 1B). This subject was part of the test set, and was correctly attributed a maximal probability of 1.0 for unfavourable outcome. Of note, the exclusion of these recordings will not necessarily lessen the potential value of synchronization measures in future multi modal approaches. Since these patterns are easily recognized by visual analysis and straightforward to interpret ("highly malignant" EEG patterns (Westhall et al., 2016)), they are not those for which quantitative measures are most needed.

"Self fulfilling prophecy" is always at play in clinical studies (Rossetti et al., 2010; Zandbergen et al., 2006). However, importantly, the present analysis was not available to clinicians at the time of patient management, strongly reinforcing the validity of the results.

It is not excluded that some parameters of our model are sub-optimal. We tested 255 different features combinations, and therefore could not assess systematically several other parameters because of combinatorial explosion.

Most EEGs have been recorded between the 12th and 24th hour after CA. We postulate that an earlier analysis would be less accurate, for outcome prediction, in the same way that "malignant" pattern such as burst-suppression are less specific for poor outcome if registered within 12 hours after CA (Cloostermans et al., 2012). However, the influence of timing on the results of the quantitative analysis remains open, and further studies (e.g. using continuous EEG) are needed to clarify this point.

4.4 Conclusion

Our findings demonstrate that EEG bivariate synchronization measures are reliable predictors of neurological outcome in the very early phase after cardiac arrest. However, despite promising results, quantitative methods have not yet been included to the decision process concerning intensive care withdrawal in comatose patients after CA. To this end, computational approaches will have to demonstrate the utility of qEEG *in addition* to current clinical prognosticators, in particular its redundancy or complementarity with visual analysis

based on classifications such as the American Clinical Neurophysiology's Terminology (Hirsch et al., 2013) . The growing usage of continuous EEG in the critically-ill patients initiated in the recent years, in view of the amount of data generated, and the time cost of classical frame-by-frame analysis, should represent an incentive to include quantitative EEG analysis in future large prospective studies.

References

Andrzejak, R.G., Chicharro, D., Lehnertz, K., Mormann, F., 2011. Using bivariate signal analysis to characterize the epileptic focus: The benefit of surrogates. Phys. Rev. E 83. doi:10.1103/PhysRevE.83.046203

Beudel, M., Tjepkema-Cloostermans, M.C., Boersma, J.H., van Putten, M.J.A.M., 2014. Small-World Characteristics of EEG Patterns in Post-Anoxic Encephalopathy. Front. Neurol. 5. doi:10.3389/fneur.2014.00097

Bishop, C.M., 2006. Pattern recognition and machine learning, Information science and statistics. Springer, New York.

Booth, C.M., Boone, R.H., Tomlinson, G., Detsky, A.S., 2004. Is This Patient Dead, Vegetative, or Severely Neurologically Impaired?: Assessing Outcome for Comatose Survivors of Cardiac Arrest. JAMA 291, 870. doi:10.1001/jama.291.7.870

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198. doi:10.1038/nrn2575

Claassen, J., Hirsch, L.J., Kreiter, K.T., Du, E.Y., Sander Connolly, E., Emerson, R.G., Mayer, S.A., 2004. Quantitative continuous EEG for detecting delayed cerebral ischemia in patients with poor-grade subarachnoid hemorrhage. Clin. Neurophysiol. 115, 2699–2710. doi:10.1016/j.clinph.2004.06.017

Cloostermans, M.C., van Meulen, F.B., Eertman, C.J., Hom, H.W., van Putten, M.J.A.M., 2012. Continuous electroencephalography monitoring for early prediction of neurological outcome in postanoxic patients after cardiac arrest: a prospective cohort study. Crit. Care Med. 40, 2867–2875. doi:10.1097/CCM.0b013e31825b94f0

Foreman, B., Mahulikar, A., Tadi, P., Claassen, J., Szaflarski, J., Halford, J.J., Dean, B.C., Kaplan, P.W., Hirsch, L.J., LaRoche, S., 2016. Generalized periodic discharges and "triphasic waves": A blinded evaluation of inter-rater agreement and clinical significance. Clin. Neurophysiol. 127, 1073–1080. doi:10.1016/j.clinph.2015.07.018

Halford, J.J., Shiau, D., Desrochers, J.A., Kolls, B.J., Dean, B.C., Waters, C.G., Azar, N.J., Haas, K.F., Kutluay, E., Martz, G.U., Sinha, S.R., Kern, R.T., Kelly, K.M., Sackellares, J.C., LaRoche, S.M., 2015. Inter-rater agreement on identification of electrographic seizures and periodic discharges in ICU EEG recordings. Clin. Neurophysiol. 126, 1661–1669. doi:10.1016/j.clinph.2014.11.008

Hirsch, L.J., LaRoche, S.M., Gaspard, N., Gerard, E., Svoronos, A., Herman, S.T., Mani, R., Arif, H., Jette, N., Minazad, Y., Kerrigan, J.F., Vespa, P., Hantus, S., Claassen, J., Young, G.B., So, E., Kaplan, P.W., Nuwer, M.R., Fountain, N.B., Drislane, F.W., 2013. American Clinical Neurophysiology Society's Standardized Critical Care EEG Terminology: 2012 version. J. Clin. Neurophysiol. 30, 1–27. doi:10.1097/WNP.0b013e3182784729

Hofmeijer, J., Beernink, T.M.J., Bosch, F.H., Beishuizen, A., Tjepkema-Cloostermans, M.C., van Putten, M.J.A.M., 2015. Early EEG contributes to multimodal outcome prediction of postanoxic coma. Neurology 85, 137–143. doi:10.1212/WNL.0000000000001742

Hofmeijer, J., Tjepkema-Cloostermans, M.C., van Putten, M.J.A.M., 2014. Burst-suppression with identical bursts: A distinct EEG pattern with poor outcome in postanoxic coma. Clin. Neurophysiol. 125, 947–954. doi:10.1016/j.clinph.2013.10.017

Horn, J., Cronberg, T., Taccone, F.S., 2014. Prognostication after cardiac arrest. Curr. Opin. Crit. Care 20, 280–286. doi:10.1097/MCC.0000000000000085

Jiruska, P., de Curtis, M., Jefferys, J.G.R., Schevon, C.A., Schiff, S.J., Schindler, K., 2013. Synchronization and desynchronization in epilepsy: controversies and hypotheses: Synchronization in epilepsy. J. Physiol. 591, 787–797. doi:10.1113/jphysiol.2012.239590

Lee, U., Kim, S., Noh, G.-J., Choi, B.-M., Hwang, E., Mashour, G.A., 2009. The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. Conscious. Cogn. 18, 1069–1078. doi:10.1016/j.concog.2009.04.004

Marschinski, R., Kantz, H., 2002. Analysing the information flow between financial time series: An improved estimator for transfer entropy. Eur. Phys. J. B 30, 275–281. doi:10.1140/epjb/e2002-00379-2

Moura, L.M.V.R., Shafi, M.M., Ng, M., Pati, S., Cash, S.S., Cole, A.J., Hoch, D.B., Rosenthal, E.S., Westover, M.B., 2014. Spectrogram screening of adult EEGs is sensitive and efficient. Neurology 83, 56–64. doi:10.1212/WNL.0000000000000537

Ng, M.C., Gaspard, N., Cole, A.J., Hoch, D.B., Cash, S.S., Bianchi, M., O'Rourke, D.A., Rosenthal, E.S., Chu, C.J., Westover, M.B., 2015. The standardization debate: A conflation trap in critical care electroencephalography. Seizure 24, 52–58. doi:10.1016/j.seizure.2014.09.017

Noirhomme, Q., Brecheisen, R., Lesenfants, D., Antonopoulos, G., Laureys, S., 2015. "Look at my classifier's result": Disentangling unresponsive from (minimally) conscious patients. NeuroImage. doi:10.1016/j.neuroimage.2015.12.006

Noirhomme, Q., Lehembre, R., Lugo, Z. d. R., Lesenfants, D., Luxen, A., Laureys, S., Oddo, M., Rossetti, A.O., 2014. Automated Analysis of Background EEG and Reactivity During Therapeutic Hypothermia in Comatose Patients After Cardiac Arrest. Clin. EEG Neurosci. 45, 6–13. doi:10.1177/1550059413509616

Oddo, M., Rossetti, A.O., 2014. Early Multimodal Outcome Prediction After Cardiac Arrest in Patients Treated With Hypothermia*: Crit. Care Med. 42, 1340–1347. doi:10.1097/CCM.0000000000000211

Oh, S.H., Park, K.N., Shon, Y.-M., Kim, Y.-M., Kim, H.J., Youn, C.S., Kim, S.H., Choi, S.P., Kim, S.C., 2015. Continuous Amplitude-Integrated Electroencephalographic Monitoring Is a Useful Prognostic Tool for Hypothermia-Treated Cardiac Arrest PatientsCLINICAL PERSPECTIVE. Circulation 132, 1094–1103. doi:10.1161/CIRCULATIONAHA.115.015754

Pereda, E., Quiroga, R.Q., Bhattacharya, J., 2005. Nonlinear multivariate analysis of neurophysiological signals. Prog. Neurobiol. 77, 1–37. doi:10.1016/j.pneurobio.2005.10.003

Rossetti, A.O., Oddo, M., Logroscino, G., Kaplan, P.W., 2010. Prognostication after cardiac arrest and hypothermia: A prospective study. Ann. Neurol. NA–NA. doi:10.1002/ana.21984

Rossetti, A.O., Rabinstein, A.A., Oddo, M., 2016. Neurological prognostication of outcome in patients in coma after cardiac arrest. Lancet Neurol. 15, 597–609. doi:10.1016/S1474-4422(16)00015-6

Ruijter, B.J., van Putten, M.J.A.M., Hofmeijer, J., 2015. Generalized epileptiform discharges in postanoxic encephalopathy: Quantitative characterization in relation to outcome. Epilepsia 56, 1845–1854. doi:10.1111/epi.13202

Rundgren, M., Westhall, E., Cronberg, T., Rosén, I., Friberg, H., 2010. Continuous amplitude-integrated electroencephalogram predicts outcome in hypothermia-treated cardiac arrest patients: Crit. Care Med. 38, 1838–1844. doi:10.1097/CCM.0b013e3181eaa1e7

Sandroni, C., Cariou, A., Cavallaro, F., Cronberg, T., Friberg, H., Hoedemaekers, C., Horn, J., Nolan, J.P., Rossetti, A.O., Soar, J., 2014. Prognostication in comatose survivors of cardiac arrest: An advisory statement from the European Resuscitation Council and the European Society of Intensive Care Medicine. Intensive Care Med. 40, 1816–1831. doi:10.1007/s00134-014-3470-x

Schindler, K., Rummel, C., Andrzejak, R.G., Goodfellow, M., Zubler, F., Abela, E., Wiest, R., Pollo, C., Steimer, A., Gast, H., 2016. Ictal time-irreversible intracranial EEG signals as markers of the epileptogenic zone. Clin. Neurophysiol. 127, 3051–3058. doi:10.1016/j.clinph.2016.07.001

Sivaraju, A., Gilmore, E.J., Wira, C.R., Stevens, A., Rampal, N., Moeller, J.J., Greer, D.M., Hirsch, L.J., Gaspard, N., 2015. Prognostication of post-cardiac arrest coma: early clinical and electroencephalographic predictors of outcome. Intensive Care Med. 41, 1264–1272. doi:10.1007/s00134-015-3834-x

Spalletti, M., Carrai, R., Scarpino, M., Cossu, C., Ammannati, A., Ciapetti, M., Tadini Buoninsegni, L., Peris, A., Valente, S., Grippo, A., Amantini, A., 2016. Single electroencephalographic patterns as specific and time-dependent indicators of good and poor outcome after cardiac arrest. Clin. Neurophysiol. 127, 2610–2617. doi:10.1016/j.clinph.2016.04.008

Stam, C.J., van Straaten, E.C.W., 2012. The organization of physiological brain networks. Clin. Neurophysiol. 123, 1067–1087. doi:10.1016/j.clinph.2012.01.011

Steimer, A., Zubler, F., Schindler, K., 2015. Chow–Liu trees are sufficient predictive models for reproducing key features of functional networks of periictal EEG time-series. NeuroImage 118, 520–537. doi:10.1016/j.neuroimage.2015.05.089

Taccone, F., Cronberg, T., Friberg, H., Greer, D., Horn, J., Oddo, M., Scolletta, S., Vincent, J.-L., 2014. How to assess prognosis after cardiac arrest and therapeutic hypothermia. Crit. Care 18, 202. doi:10.1186/cc13696

Tjepkema-Cloostermans, M.C., van Meulen, F.B., Meinsma, G., van Putten, M.J., 2013. A Cerebral Recovery Index (CRI) for early prognosis in patients after cardiac arrest. Crit. Care 17, R252. doi:10.1186/cc13078

Tsetsou, S., Oddo, M., Rossetti, A.O., 2013. Clinical Outcome After a Reactive Hypothermic EEG Following Cardiac Arrest. Neurocrit. Care 19, 283–286. doi:10.1007/s12028-013-9883-5

Tzovara, A., Rossetti, A.O., Juan, E., Suys, T., Viceic, D., Rusca, M., Oddo, M., Lucia, M.D., 2016. Prediction of awakening from hypothermic postanoxic coma based on auditory discrimination: Awakening from Postanoxic Coma. Ann. Neurol. 79, 748–757. doi:10.1002/ana.24622

Wennervirta, J.E., Ermes, M.J., Tiainen, S.M., Salmi, T.K., Hynninen, M.S., Särkelä, M.O.K., Hynynen, M.J., Stenman, U.-H., Viertiö-Oja, H.E., Saastamoinen, K.-P., Pettilä, V.Y., Vakkuri, A.P., 2009. Hypothermia-treated cardiac arrest patients with good neurological outcome differ early in quantitative variables of EEG suppression and epileptiform activity*: Crit. Care Med. 37, 2427–2435. doi:10.1097/CCM.0b013e3181a0ff84

Westhall, E., Rosén, I., Rossetti, A.O., van Rootselaar, A.-F., Wesenberg Kjaer, T., Friberg, H., Horn, J., Nielsen, N., Ullén, S., Cronberg, T., 2015. Interrater variability of EEG interpretation in comatose cardiac arrest patients. Clin. Neurophysiol. 126, 2397–2404. doi:10.1016/j.clinph.2015.03.017

Westhall, E., Rossetti, A.O., van Rootselaar, A.-F., Wesenberg Kjaer, T., Horn, J., Ullén, S., Friberg, H., Nielsen, N., Rosén, I., Åneman, A., Erlinge, D., Gasche, Y., Hassager, C., Hovdenes, J., Kjaergaard, J., Kuiper, M., Pellis, T., Stammet, P., Wanscher, M., Wetterslev, J., Wise, M.P., 2016. Standardized EEG interpretation accurately predicts prognosis after cardiac arrest. Neurology 86, 1482–1490. doi:10.1212/WNL.0000000000002462

Zandbergen, E.G.J., Hijdra, A., Koelman, J.H.T.M., Hart, A.A.M., Vos, P.E., Verbeek, M.M., de Haan, R.J., for the PROPAC Study Group*, 2006. Prediction of poor outcome within the first 3 days of postanoxic coma. Neurology 66, 62–68. doi:10.1212/01.wnl.0000191308.22233.88

Zubler, F., Koenig, C., Steimer, A., Jakob, S.M., Schindler, K.A., Gast, H., 2016. Prognostic and diagnostic value of EEG signal coupling measures in coma. Clin. Neurophysiol. 127, 2942–2952. doi:10.1016/j.clinph.2015.08.022

Figure legends

Figure 1: Three examples of EEGs from the test set, with their corresponding feature vectors. The analysed four bipolar derivations corresponded to the left, right, anterior and posterior brain regions. Scale bar corresponds to one second (horizontal) and 50 μV (vertical). The value of the eight synchronization measures is represented as a histogram, on a normalized scale, 1 representing the highest value observed within the 94 patients (1, RDP_{LR}; 2, RDP_{AP}; 3, CC_{LR}; 4, CC_{AP}; 5, MI_{LR}; 6, MI_{AP}; 7, TE_{LR}; 8, TE_{AP}). (A) 62-year-old woman; the classifier (correctly) assigned a probability of 0.0 for poor outcome; the patient recovered (CPC 1). (B) 56-year-old male; the classifier (correctly) assigned a maximum probability of 1.0 to poor outcome. The patient died at the hospital. (C) 60-year-old male. The classifier (wrongfully) associated a high probability for good outcome, whereas the patient deceased. However, the EEG was continuous and reactive, and the somatosensory evoked potentials were present, whereas the brainstem reflexes were absent, so that this case would have been very challenging even for experienced clinicians.

Figure 2: Prognostic performance for clinical outcome at 3 months for the combination RDP_{AP}, CC_{LR}, MI_{LR}, and TE_{AP} using a Bayes classifier with mixture of Gaussians. (A) Receiver operating characteristic (ROC) curve obtained on 63 patients using leave-one-out cross-validation. The area under the ROC curve (AUC) was 0.84. The gap between the first (0,0) and second (0,0.36) point of the ROC curve is due to the fact that 12 patients were (correctly) attributed a maximum probability of 1.0 for bad outcome. (B) ROC curve obtained by training the classifier on 63 patients, and testing it on the remaining 31 patients. The AUC was 0.81. Here also there is a gap between the first two points of the curve, due to 3 patients with (correctly) estimated probability of 1.0 for bad outcome. The classifier failed to reach the maximum sensitivity before the last point (1,1), due to a single patient being incorrectly attributed a very high probability for good outcome (Patient in Fig. 1C).

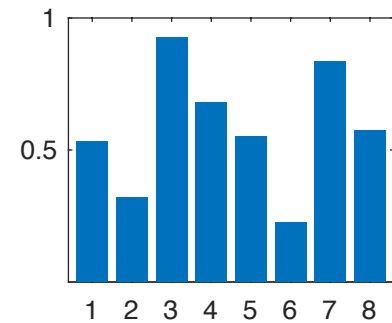## Tables

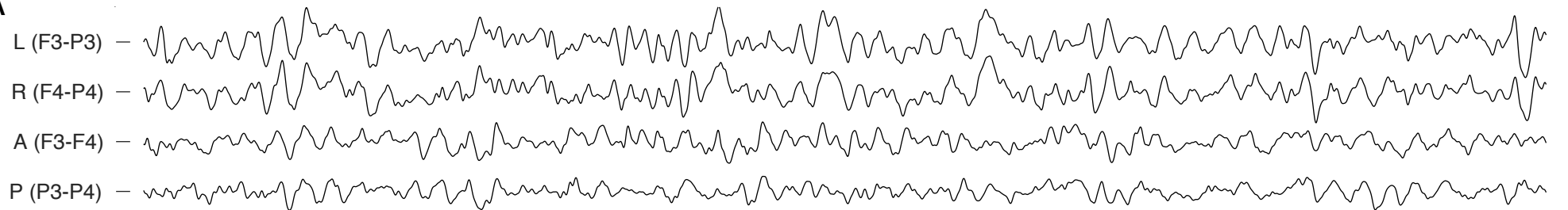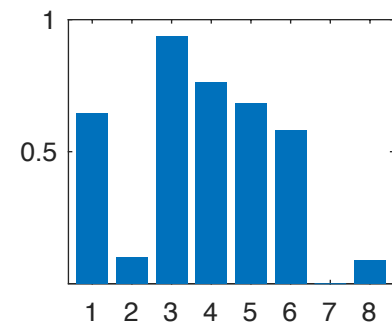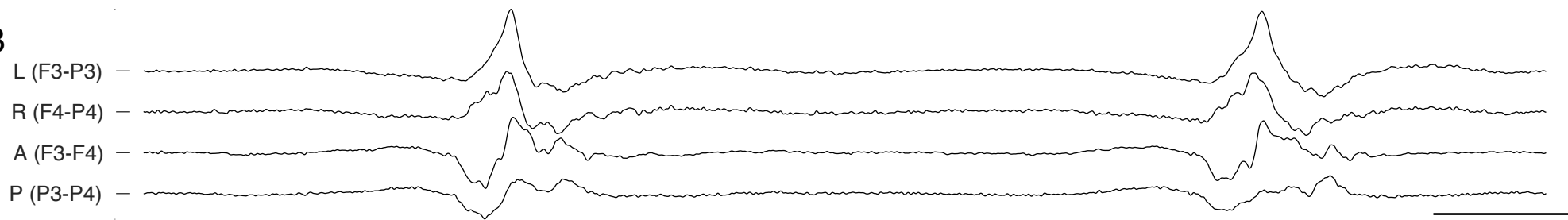Table 1: Clinical characteristics of patients.

	Good outcome (CPC 1,2)	Poor outcome (CPC 3-5)	p-value
N	48	46	
Female gender	10	13	0.40
Age ($\pm$ sd) [y]	56.7 ($\pm$ 16)	65 ($\pm$ 15)	0.01
Non-cardiac aetiology	6	16	0.01
Asystole or pulseless electrical activity on site	8	24	<0.01
Latency of EEG recording ($\pm$ sd) [h]	18.8 ($\pm$ 4.3)	20.5 ($\pm$ 6.2)	0.07
EEG pattern: non-reactive highly malignant pattern[1]	1 (2%) 0	24 (52%) 14 (30%)	<0.01 <0.01
Patients sedated with propofol*	14/44 (32%)	12/37 (32%)	0.95
Mean propofol dosis*($\pm$ sd) [mg/kg/h]	1.79 ($\pm$ 0.95)	1.80 ($\pm$ 0.97)	0.98
Patients sedated with midazolam*	41/44 (93%)	31/37 (84%)	0.47
Mean midazolam dosis* ($\pm$ sd) [mg/kg/h]	0.12 ($\pm$ 0.04)	0.13 ($\pm$ 0.04)	0.45

Overview of the demographics and clinical information for patients with favourable and unfavourable outcome (* values concerning sedation are based on the 81/94 patients for whom they were documented). 1 as defined in (Westhall et al., 2016)

Table 2: Feature combinations with highest predictive power for outcome.

RDP_{LR}	RDP_{AP}	CC_{LR}	CC_{AP}	MI_{LR}	MI_{AP}	TE_{LR}	TE_{AP}	AUC
	x	x		x			x	0.8444
	x	x	x	x	x			0.8182
		x		x			x	0.8141
x		x		x				0.8066
x	x	x		x			x	0.8061
	x	x					x	0.8020
		x		x		x	x	0.8010
x		x		x		x	x	0.8005
x		x		x			x	0.7995
	x	x	x	x	x		x	0.7980

The table lists the 10 feature combinations with highest predictive power for clinical outcome, assessed with the area under the receiving operating characteristic curve (AUC) on 2/3 of patients using cross-validation. Rows represent individual combinations; a cross (x) indicates that the corresponding synchronization measure is included in the combination.