SUPPLEMENTAL MATERIAL

Words in the Supplemental Material: *1261 words*
Supplemental Tables: *2*

1. Mismatch Negativity Protocol

Patients took part in a mismatch negativity (MMN) protocol as described previously [1, 2]. A series of 16-bit stereo sinusoidal tones, sampled at 44.1 kHz, with a 10-ms linear amplitude envelope at onset and offset to avoid clicks was presented at 75 db loudness on in-ear stereo headphones (model ER-4P, Etymotic Research). Sounds were presented in three identical blocks of 500 stimuli for each recording. In each block there were 350 "standard" sounds (70% of the total) consisting of 1,000Hz tones with 100-millisecond duration and 0-millisecond interaural difference. The standard sounds were replaced pseudorandomly by three types of "deviant" sounds, which differed from the standard ones with respect to their pitch, duration, or location. There were 50 deviant sounds of each type in one block. Duration deviants were 1,000Hz, with 150-millisecond duration and 0-millisecond interaural difference. Pitch deviants were 1,200Hz tones with 100-millisecond duration and 0-millisecond interaural difference. Deviants in location were 1,000Hz tones, with 100-millisecond duration and 700- microsecond interaural difference, with the left ear leading. Sounds were presented at a fixed 750 ms inter-stimulus interval. We always recorded three blocks during the first day recording and three blocks during the second day recording, resulting thus in 1,500 presented stimuli per recoding.

2. EEG Acquisition and Preprocessing

Each patient had two EEG recordings at bedsite in the intensive care unit. The first recording took place within 24 hours after coma onset during TTM 36 and the second

recording at approximately 36-48 hours after coma onset after withdrawal of TTM, off sedation. In 12 patients the same clinical EEG system (Madison, WI) as describe previously [1, 2] was used. It had a sampling rate of 1000 Hz and consisted of 19 electrodes placed according to the international 10-20 system. In the remaining 50 patients an g.tec EEG system (i.e. g.HIamp, Guger Technologies, Graz, Austria) with a sampling rate of 1200 Hz and 62 active electrodes placed according to the 10-20 system was used. For the latter 50 patients, after having completed the auditory MMN task they also took part in a somatosensory stimulation protocol, which will be reported elsewhere. For the aim of the present study, we focused the analysis on the auditory MMN protocol for a selection of the same 19 EEG channels from the clinical and gtec electrode montages. Across all patients the impedances were kept <10kΩ and the data was referenced online to the Fpz electrode and in the course of preprocessing the average reference was computed. We preprocessed the EEG data offline using the same procedure as in [1, 3]. We extracted EEG responses to deviant sounds from the three experimental blocks and an equal number of responses to standard sounds.

3. Multivariate EEG Decoding

Single-patient EEG data was analyzed with a multivariate decoding algorithm based on EEG responses across the whole 19-channel montage [4]. This method can be used to quantify the differential responses to standard versus deviant sounds at the level of each single patient and recording. It has been previously used for decoding responses in healthy subjects [5, 6] and comatose patients [1, 2, 7]. This algorithm consists of modeling the distribution of single-trial EEG responses across all electrodes using a mixture of Gaussian models (GMM) in an n-dimensional space where n represents the number of electrodes [8, 9]. The models are computed through an expectation-maximization algorithm [10] for each patient and recording (first day, second day) separately, using only one part of the available data (training data set,

consisting of 90% of the artifact-free single trials). They are then fitted back to the single trials of the training data set by computing posterior probabilities [11]. These represent the probability of every time point and trial to be represented by the computed GMM models. Each trial in the training data set is decoded as being a response to a standard or a deviant sound according to which of the two models provide the highest posterior probability. The generalization of the decoding performance is then assessed by fitting them on the remaining 10% of the available single trials (test data set) and by assigning the test trials in one of the two experimental conditions (i.e., responses to standard vs. deviant sounds).

Decoding performance is measured as the area under the receiver operator characteristic curve (AUC, [12]) and it is computed for standard versus each type of deviant sound. The GMM model's parameters are optimized by repeating this whole procedure 10 times by splitting the data in training and test data sets in a way that the 10 test data sets never overlap. All AUC values reported here correspond to the mean value across all three contrasts (i.e., responses to standard sounds vs. deviant sounds in pitch, duration, or location). Full details about this algorithm have been reported elsewhere [4]. Here, we applied this algorithm as in our previous studies [1, 2] based on the same auditory MMN paradigm [13] in a new cohort of comatose patients treated with TTM 36 during the first day recording and after withdrawal of temperature control on the second day recording [14]. Outcome prediction was based on the change of decoding performance from Day 1 (AUC_{DAY1}) to Day 1 (AUC_{DAY2}) and specifically on the percentage change in AUC values: $100 \times (AUC_{DAY2} - AUC_{DAY1}) / AUC_{DAY1}$. Significance of outcome prediction results was assessed with 95% confidence intervals (CIs) based on a binomial distribution. Unpaired t tests for normally distributed continuous data were used for contrasting differences between patient's quantitative descriptors (**Tables 2 and 3**; for example age). Fisher exact or chi-square tests were used as needed for categorical data.

4. Outcome-Prediction Based on Reactivity and Added Value of Auditory Discrimination Progression

EEG reactivity was routinely assessed in the vast majority of the patients of the included cohort. This allows a direct comparison of the predictive power based on reactivity and that based on the progression of auditory discrimination. **Table S1** shows a full description of outcome prediction assessment from these two measurements and of their combination. The best results for predicting good outcome is derived from the combined score of presence of a reactive EEG on Day 2 and a positive progression of auditory discrimination as shown both in the increase absolute value of the positive predictive value (PPV) and in the significance of the specificity of the combined test.

Table S1: Prognostic values when including patients with comorbidities based on auditory discrimination, reactivity and their combination. Please note that the sample is restricted to 59 patients (30 Survivors, 29 Non-Survivors) as total number of patients who were also tested for reactivity on Day2 to allow comparisons between outcome prediction results. Values above chance level are highlighted in red.

	Progression in Auditory Discrimination	**Reactive EEG on Day 2**	**Reactive EEG and Progression in Auditory Discrimination**
PPV (95%CI)	**0.58** (0.37-0.78)	**0.73** (0.57-0.86)	**0.82** (0.57-0.96)
Sensitivity (95%CI)	**0.47** (0.28-0.66)	**1.00** (0.88-1.00)	**0.47** (0.28-0.66)
Specificity (95%CI)	**0.66** (0.46-0.82)	**0.62** (0.42-0.79)	**0.90** (0.73-0.98)
NPV (95%CI)	**0.54** (0.37-0.71)	**1.00** (0.81-1.00)	**0.62** (0.46-0.76)
Accuracy (95%CI)	**0.56** (0.38-0.63)	**0.81** (0.93-1.00)	**0.68** (0.43-0.67)

5. Outcome-Prediction Based on Auditory Discrimination on Day1

Based on the average decoding values from Survivors and Non-Survivors from the first and second day of coma (Figure 1), one might argue that the outcome prediction results were mainly driven by a high decoding performance of Non-Survivors on Day 1 (see Figure 1-B). We tested this hypothesis separate control analyses for both the 'All Patients' (n=60) and the 'No Epileptiform Features' sample (n=46). We specified a threshold for each of the analyzed patient samples to allow splitting patients according to a high first-day AUC (i.e. above specified threshold) and a low first-day AUC (i.e. below threshold). The AUC threshold was specified in such a way that the number of Survivors with high AUC was identical to the number of Survivors with an increase in our main analysis. Thus the $AUC_{THRESH} = 0.607$ for the 'All Patients' sample and $AUC_{THRESH} = 0.609$ for the 'No Epileptiform Features' sample. Results of outcome prediction based on these data from the first day were all non-significant for PPV, sensitivity, specificity, negative predictive value, and overall test accuracy (see Table S1 for an overview). Thus, outcome prediction based on decoding performance from the first day after CA was not predictive of coma outcome.

Table S2: Prognostic value for good outcome based on AUC_{DAY1} for patients treated with TTM 36. Results are shown separately for analyses across all patients and across subgroup of patients without epileptiform features.

	All Patients **n = 60**	**No Epileptiform Features** **n = 46**
PPV (95%CI)	**0.51** (0.34-0.69)	**0.50** (0.32-0.68)
Sensitivity (95%CI)	**0.24** (0.09-0.45)	**0.14** (0.02-0.43)
Specificity (95%CI)	**0.49** (0.32-0.66)	**0.57** (0.37-0.76)

NPV (95%CI)	**0.26** (0.10-0.48)	**0.11** (0.01-0.35)
Accuracy (95%CI)	**0.40** (0.29-0.55)	**0.39** (0.23-0.51)

References

[1] Tzovara A, Rossetti AO, Juan E, Suys T, Viceic D, Rusca M, et al. Prediction of awakening from hypothermic post anoxic coma based on auditory discrimination. Annals of neurology. 2016.

[2] Tzovara A, Rossetti AO, Spierer L, Grivel J, Murray MM, Oddo M, et al. Progression of auditory discrimination based on neural decoding predicts awakening from coma. Brain. 2013;136:81-9.

[3] Naccache L, Puybasset L, Gaillard R, Serve E, Willer JC. Auditory mismatch negativity is a good predictor of awakening in comatose patients: a fast and reliable procedure. Clin Neurophysiol. 2005;116:988-9.

[4] De Lucia M, Tzovara A. Decoding auditory EEG responses in healthy and clinical populations: A comparative study. Journal of neuroscience methods. 2015;250:106-13.

[5] Bernasconi F, De Lucia M, Tzovara A, Manuel AL, Murray MM, Spierer L. Noise in brain activity engenders perception and influences discrimination sensitivity. J Neurosci. 2011;31:17971-81.

[6] De Lucia M, Tzovara A, Bernasconi F, Spierer L, Murray MM. Auditory perceptual decision-making based on semantic categorization of environmental sounds. Neuroimage. 2012;60:1704-15.

[7] Cossy N, Tzovara A, Simonin A, Rossetti AO, De Lucia M. Robust discrimination between EEG responses to categories of environmental sounds in early coma. Front Psychol. 2014;5:155.

[8] Tzovara A, Murray MM, Plomp G, Herzog MH, Michel CM, De Lucia M. Decoding stimulus-related information from single-trial EEG responses based on voltage topographies. Pattern Recognition. 2012;45:2109-22.

[9] Tzovara A, Murray MM, Michel CM, De Lucia M. A tutorial review of electrical neuroimaging from group-average to single-trial event-related potentials. Developmental neuropsychology. 2012;37:518-44.

[10] Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data via the EM Algorithm. Journal of the Royal Statistical Society Series B (Methodological). 1977;39:1-38.

[11] Bishop CM. Neural networks for pattern recognition. Oxford, UK: Oxford University Press; 1995.

[12] Green D, Swets J. Signal detection theory and psychophysics. New York, NY: John Wiley and Sons; 1966.

[13] Todd J, Michie PT, Schall U, Karayanidis F, Yabe H, Näätänen R. Deviant matters: duration, frequency, and intensity deviants reveal different patterns of mismatch negativity reduction in early and late schizophrenia. Biological psychiatry. 2008;63:58-64.

[14] Tsetsou S, Novy J, Pfeiffer C, Oddo M, Rossetti AO. Multimodal Outcome Prognostication After Cardiac Arrest and Targeted Temperature Management: Analysis at 36 degrees C. Neurocritical care. 2017.