N400 Predicts Recovery from Disorders of Consciousness

Inga Steppacher, PhD, Simon Eickhoff, MSc, Todor Jordanov, PhD, Michael Kaps, MD, Wolfgang Witzke, PhD, and Johanna Kissler, PhD

Objective: Patients with the unresponsive wakefulness syndrome (UWS; formerly vegetative state) or in a minimally conscious state (MCS) open their eyes spontaneously but show no (UWS) or only marginal (MCS) signs of awareness. Because these states can become permanent, residual information processing capacities need to be determined, and reliable outcome predictors need to be found. We assessed higher-order cortical information processing in UWS or MCS in a large group of patients using electroencephalographic event-related potentials (ERPs) and determined their long-term prognostic value for recovery.

Methods: Cognitive ERPs elicited by sound (P300) and speech (N400) were used to assess information processing in 92 behaviorally unresponsive patients diagnosed as in the state of either UWS (n = 53) or MCS (n = 39). ERPs were assessed with a clinical standard evaluation method and a computerized method, the t-continuous wavelet transform. The patients’ clinical outcome was followed up between 2 and 14 years after discharge from the rehabilitation center.

Results: Within the first year of the disease, many patients showed an intact P300 and several also an N400, indicating considerable residual information processing. At clinical follow-up, about 25% of the patients recovered and regained communicative capabilities. A highly significant relationship between N400, but not P300, presence and subsequent recovery was found.

Interpretation: Results specify cognitive capabilities in disorders of consciousness, and determine their prognostic value. Specifically the N400 ERP is suggested as an important tool to assess information-processing capacities that can predict the likelihood of recovery of patients in UWS or MCS.

Due to advances in intensive care medicine, the number of patients surviving severe brain damage has increased in recent years. Such patients often fall into a coma, from which they may gradually progress to an unresponsive wakefulness syndrome (UWS; formerly called a vegetative state1–3) and later a minimally conscious state (MCS; as defined by Giacino et al4). UWS and MCS patients show spontaneous movements, have sleep–wakefulness cycles, and spontaneously open or close their eyes, yet show no (UWS) or only minimal (MCS) signs of awareness. These states may last for months or years—a challenging situation for families and medical staff. At present, the extent to which such patients are capable of processing environmental information is unclear, and prognostic indicators predicting their functional outcome are scarce and very unspecific (further information on outcome in patients with severe alterations of consciousness is available2,5,6). It has long been known that brain event-related potentials (ERPs) offer a means to assess information processing capabilities in the absence of overt behavior.7 Some studies have already demonstrated preserved ERPs in coma, UWS, and MCS,8–11 but due to the small number of investigated cases and frequent lack of outcome documentation, their clinical and prognostic significance is currently unclear. However, ERPs could be a powerful and efficient tool to assess information processing and perhaps even cortical correlates of command following in behaviorally...
unresponsive patients. Recently, Cruse et al showed for the first time that some patients who behaviorally seemed to be entirely unresponsive were able to modulate Mu-rhythm electroencephalographic (EEG) oscillations over the motor cortex following motor imagery instructions. In addition to their usefulness in immediately identifying cortically responsive patients, the presence of ERPs after brain damage may predict subsequent recovery of consciousness and communicative abilities and, if systematically assessed and successfully modulated, open new avenues for rehabilitation. In particular, endogenous ERP components that appear relatively late after stimulus onset in the processing stream and reflect more complex cognitive operations may have prognostic value. They depend on the synchronized activity of multiple brain regions, which may be critical for conscious awareness.

Here, we examine in a large number of UWS and MCS patients the prevalence and long-term predictive values of the P300 and the N400 endogenous ERPs. Both components reflect higher-order auditory processing. The P300 is elicited by the detection of irregularities in sequences of environmental sounds, whereas the N400 reflects the detection of semantic incongruities in spoken language.

Patients and Methods

Patients

The initial sample consisted of 175 patients with UWS (n=92) or MCS (n=83) treated at the Neurorehabilitation Hospital Kliniken Schmieder (Allensbach, Germany) between 1994 and 2005. These patients had undergone ERP measurement during their stay at the clinic’s Early Rehabilitation Unit. In line with the recommendations of the Multi Society Task Force on PVS, inclusion criteria for the ERP study were: (1) integrity of the hearing nerves as evident by normal brainstem auditory evoked potentials and no documented pre-existing hardness of hearing or deafness; (2) time since critical incident for hypoxic events not more than 3 months, for traumatic events not more than 12 months; and (3) regarding clinical state, patients must have documented to be in either UWS or MCS when the ERP measurement was recorded.

Follow-up data could be obtained from 53 UWS and 39 MCS patients (92 patients in total). These patients’ combined ERP and outcome data are included in this study.

Coma Remission Scale

The patients’ cognitive functioning was evaluated using the Coma Remission Scale (CRS). The CRS was developed in Germany specifically to monitor and protocol the improvements of coma, UWS, and MCS patients in early rehabilitation units and was used in all patients as part of the clinic routine.

The CRS consists of 6 subscales, testing for arousability, motor responses, reaction to acoustic stimuli, reaction to visual stimuli, and the ability to produce sounds or words. Each subscale is rated independently and points are given, reaching from 0 to a maximum of 3 to 5, depending on the subscale. A maximum of 24 points can be reached, indicating an awake, responsive patient, whereas zero indicates a deeply comatose patient.

For a translated version of the CRS, see the Supplementary Materials.

EEG Recording

EEG activity was recorded from 5 silver–chloride electrodes at Fz, Cz, Pz, C3, and C4 (according to the international 10-20 system). They were referenced to linked earlobes, and the ground was placed on the mid forehead. The electro-oculogram electrode was placed above and below the right eye. Data were acquired with a sampling rate of 256Hz, a 70Hz low-pass filter, and an additional 50Hz Notch filter, using SCAN software (NeuroScan, Charlotte, NC).

The acoustic stimuli were generated with the software STIM (NeuroScan) and were presented over earphones. The ERPs were averaged offline with Brain Electrical Source Analysis (BESA, Gräfelfing, Germany). Trials with blink or movement artifacts were excluded from further analysis.

Stimuli and Procedure

P300. For the P300 an oddball paradigm was used with five hundred 1,000Hz sine tones (nontargets), one hundred 1,500Hz sine tones (targets), and 100 novel environmental sounds. The environmental sounds were synthetically generated but resembled naturally occurring sounds such as squeaking, clicking, or slurping. Stimulus duration was 100 milliseconds; intensity was 90dB. Stimuli were presented in randomized order with an interstimulus interval of 1 second. The patients were told to count the higher tones. Acoustic deviance, such as the occurrence of novel environmental sounds, typically elicits a P300.

N400. The paradigm for the N400 consisted of 200 five-word sentences. One hundred of these ended with a word that fit the given context; 100 endings were senseless (for example: “My sister has four children” vs “Flowers need lots of sugar”). The 2 conditions were alternated in pseudorandomized fashion. An artificial delay of 1.5 seconds was inserted before the last word of each sentence to ensure a stable EEG baseline. Sentences were separated by a pause of 3 seconds. Again, stimulus intensity was held at 90dB. Patients were instructed to decide whether the last word made sense in the given sentence. Semantic deviance, such as a nonsense sentence ending, typically elicits an N400.

For both paradigms, instructions were given orally regardless of the patients’ presumed ability to follow them.

Clinical Follow-up

The clinical follow-up was a structured telephone interview. Depending on the patient, the interview took place 2 to 15 years after the initial incident (see Supplementary Tables 1 and 2 for detailed information). From the original sample of 175 patients (92 UWS, 83 MCS), 53 former UWS patients and 39 former MCS patients were successfully followed up. In the interview, the patients, or in most cases their caregivers or relatives, were asked to complete two questionnaires (the CRS and
the Barthel Index) about the patient’s current condition. If the patient had already died, the relatives where interviewed about the last cognitive and medical state the patients had been in, using the same questionnaires. Because relatives where not trained to administer the behavioral tests and because correctly diagnosing patients in UWS is a challenge even for experienced clinical professionals,14,15 for follow-up, no distinction between UWS and MCS was made. Although it would have been more precise to refer to former UWS patients as somewhat recovered if they improved to MCS, we felt that this distinction could not be made reliably on the basis of behavioral scores given by the relatives. We therefore considered patients as having recovered consciousness if they were able to functionally communicate (by any means), expecting that this is a boundary relatives are not likely to miss and could report reliably. Note that this represents a relatively conservative measure of recovery. Follow-up CRS scores derived from the telephone interview are available for inspection in the Supplementary Tables.

**ERP Analysis**

ERP components were analyzed with two different approaches.

**CLINICAL SCORING.** Artifact-free trials were averaged to form an ERP. In line with previous studies,16 the presence of the respective ERP component was assessed by visual inspection by two clinical neurophysiologists, and an ERP response was scored as present when both examiners, reviewing the waveforms independently, agreed on its existence. This approach resembles the current clinical routine at many institutions.

**T-CONTINUOUS WAVELET TRANSFORM.** The t-continuous wavelet transform (tCWT) algorithm17 was implemented to automatically detect and statistically compare ERP waveforms in a computerized and rater-independent manner. Individual trials are transformed into frequency space by means of the continuous wavelet transform. The average frequency content is statistically compared across 2 conditions by means of $t$ tests, yielding a tCWT scalogram (Figs 1 and 2). An ERP wave was considered significant if tCWT significance level exceeded $p=0.01$ (Bonferroni-corrected significance level).

Figure 1 shows tCWT scalograms and ERP tracings from two UWS patients for the P300 paradigm. Figure 2 illustrates tCWT scalograms and ERP tracings for the N400 paradigm from two different UWS patients.

**Statistical Outcome Prediction**

The statistical relationship between the presence of the two ERP components and the patients’ outcome was assessed from contingency tables using Fisher exact test. Clinically, patients were categorized as (1) no improvement, meaning the patients were still in UWS or MCS; or (2) recovered, meaning the patient was responsive and able to functionally communicate. The significance level...
was set at \( p = 0.05 \). Furthermore, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and likelihood ratio of clinical improvement were calculated. For all measures, 95% confidence intervals (CIs) are reported in brackets.

**Role of the Funding Sources**
The sponsors of this study had no influence on study design; data acquisition, analysis, and interpretation; or writing of the report. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit for publication.

**Results**

**Characterization of the Followed-up Patients**
Table 1 summarizes the groups’ demographic and clinical characteristics as obtained on the follow-up interview. More detailed individual information on the patients included in this study is given in Supplementary Tables 1 (MCS patients) and 2 (UWS patients).

**ERP Results**
For some patients, data for only one of the two paradigms were available (see Supplementary Tables 1 and 2), which accounts for different numbers of analyzed data sets for P300 and N400. Furthermore, not every paradigm could be analyzed by both methods (for example, occasionally not enough artifact-free epochs were left to yield a clear ERP for visual inspection, but the tCWT still handled those data [9 cases]). In other cases, only the averaged data used for the hospital discharge summaries were available, which conversely could not be analyzed by the tCWT (11 cases). As can be seen from the Supplementary Tables, restricting the analyses to only patients for whom data from both paradigms and assessment methods were available will not qualitatively alter the results.

In general, the ERPs’ predictive value did not differ in direction, regardless of the detection method used. The tCWT always detected more patients with significant ERPs than the human raters, such that associations with outcome may differ in magnitude, but the patterns are qualitatively analogous. For the respective contingencies, see Tables 2 and 3.

For examples of tCWT t-maps and averaged waveforms used for expert evaluation, see Figures 1 (P300 paradigm) and 2 (N400 paradigm).

**P300 and Outcome**

**P300 TCWT.** In the UWS group, a total of 48 data sets could be evaluated with the tCWT. Fisher exact test
found no predictive significance ($p=0.034$, sensitivity=0.58 [95% CI=0.27–0.84], specificity=0.76 [95% CI=0.50–0.93], PPV=63% [95% CI=31–89%], NPV=50% [95% CI=30–70%], and likelihood ratio=1.5).

Contingency tables for the relationship between P300 and outcome are presented in Table 2.

### N400 and Outcome

**N400 TCWT.** In the UWS group, using the tCWT, N400–outcome associations were significant (Fisher exact test $p=0.034$, sensitivity=0.58 [95% CI=0.27–0.84], specificity=0.76 [95% CI=0.50–0.93], PPV=63% [95% CI=31–89%], NPV=50% [95% CI=30–70%], and likelihood ratio=1.5).

**N400 VISUAL INSPECTION.** In the UWS group, N400–outcome associations led to a highly significant Fisher exact test result ($p=0.0001$, sensitivity=0.60 [95% CI=0.26–0.88], specificity=0.97 [95% CI=0.86–1.00], PPV=86% [95% CI=42–100%], NPV=90% [95% CI=76–97%], and likelihood ratio=22.80).

Finally, in the MCS group, the significant association between N400 and outcome could be observed again (Fisher exact test $p=0.0033$, sensitivity=0.40 [95% CI=0.19–0.63], specificity=1.00 [95% CI=0.82–1.00], PPV=100% [95% CI=63–100%], NPV=61% [95% CI=42–78%], and likelihood ratio>100.0).

The contingencies for the association between N400 and outcome are presented in Table 3.

Unlike the P300, the N400 showed, across syndromes and both methods of analysis, a clear relationship with the patients’ outcome, which means it was associated with better clinical results.

### Discussion

This study examined the prevalence of the P300 and N400 late endogenous ERPs in a large sample of UWS or MCS patients and assessed these ERPs’ predictive value for the patients’ long-term clinical outcome several years later. Both P300 and N400 ERPs could be

---

**TABLE 1: Patient Characteristics and Outcome Data**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>MCS, Mean (SD), n=39</th>
<th>UWS, Mean (SD), n=53</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients’ age, yr</td>
<td>45.0 (16.9)</td>
<td>44.5 (14.5)</td>
</tr>
<tr>
<td>Duration of stay at the rehabilitation unit, days</td>
<td>119.4 (82.6)</td>
<td>105.6 (115.9)</td>
</tr>
<tr>
<td>Time until follow-up, yr</td>
<td>8.1 (3.7)</td>
<td>8.4 (2.9)</td>
</tr>
<tr>
<td>Mean time between event and ERP measurement, mo</td>
<td>6.8 (8.5)</td>
<td>1.9 (1.6)</td>
</tr>
<tr>
<td>Gender distribution, M/F</td>
<td>23/16</td>
<td>42/12</td>
</tr>
<tr>
<td>Type of injury, TBI/hypoxia/others</td>
<td>22/6/11</td>
<td>21/19/13</td>
</tr>
<tr>
<td>Most common cause of death</td>
<td>Pneumonia</td>
<td>Pneumonia</td>
</tr>
</tbody>
</table>

Patients were considered recovered if they reached the ability to functionally communicate. Type of injury, others=combined causes of condition (TBI and hypoxia), tumors, stroke, or encephalitis. None of the parameters differed significantly between the two patient groups.

ERP=event-related potential; F=female; M=male; MCS=minimally conscious state; SD=standard deviation; TBI=traumatic brain injury; UWS=unresponsive wakefulness syndrome.
### TABLE 2: Numbers of Patients and Distribution of P300 Occurrence and Outcome

<table>
<thead>
<tr>
<th>P300</th>
<th>tCWT</th>
<th>Human Raters</th>
<th>tCWT</th>
<th>Human Raters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UWS Patients</td>
<td>MCS Patients</td>
<td>UWS Patients</td>
<td>MCS Patients</td>
</tr>
<tr>
<td></td>
<td>Recovered</td>
<td>Not Recovered</td>
<td>Total</td>
<td>Recovered</td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>25</td>
<td>33</td>
<td>16</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>12</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>37</td>
<td>48</td>
<td>19</td>
</tr>
</tbody>
</table>

P300 data from UWS and MCS patients analyzed with Fisher exact test. P300 identified by tCWT (left columns) and human raters (right columns). Patients were considered recovered if they had reached the ability to communicate again.

MCS = minimally conscious state; tCWT = t-continuous wavelet transform; UWS = unresponsive wakefulness syndrome.

### TABLE 3: Numbers of Patients and Distribution of N400 Occurrence and Outcome

<table>
<thead>
<tr>
<th>N400</th>
<th>tCWT</th>
<th>Human Raters</th>
<th>tCWT</th>
<th>Human Raters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UWS Patients</td>
<td>MCS Patients</td>
<td>UWS Patients</td>
<td>MCS Patients</td>
</tr>
<tr>
<td></td>
<td>Recovered</td>
<td>Not Recovered</td>
<td>Total</td>
<td>Recovered</td>
</tr>
<tr>
<td>Yes</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
<td>27</td>
<td>32</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>35</td>
<td>47</td>
<td>18</td>
</tr>
</tbody>
</table>

N400 data from UWS and MCS patients analyzed with Fisher exact test. N400 identified by tCWT (left columns) and human raters (right columns). Patients were considered recovered if they had reached the ability to communicate again.

MCS = minimally conscious state; tCWT = t-continuous wavelet transform; UWS = unresponsive wakefulness syndrome.
identified in a considerable proportion of patients, but specifically the N400 elicited by semantic deviance in spoken language predicted a favorable clinical outcome.

Two assessment methods were used, the clinical routine and the computerized tCWT. The former was used because it is still the clinical standard of ERP evaluation, the latter to confirm and compare the subjective human rater–based method with a standardized and objective algorithm.

P300 was detected in 20 (visual inspection) to 69% (tCWT) of the UWS patients and in 30 (visual inspection) to 83% (tCWT) of MCS patients. The N400 was identified in 16 (visual inspection) to 32% (tCWT) of UWS patients and in 21 (visual inspection) to 41% (tCWT) of MCS patients. In previous reports, the percentage of patients having a P300 ranged from 18% to 100%. This wide range has been ascribed to variations in stimulus complexity or emotionality, heterogeneity in patients' etiology, current diagnosis, and different ERP analysis techniques. For the N400, present numbers are similar to those of Kotchoubey et al and Schoenle and Witzke, who found an N400 in 14 to 39% of UWS patients, with the rates depending on assessment criteria and stimulus complexity. In line with the present pattern, across studies, more patients are identified with late ERPs when computerized analysis techniques are used, suggesting, perhaps contrary to common intuition, that clinicians use more conservative assessment criteria than current implementations of automated algorithms. This may be due to human raters taking into account overall data quality, discounting local differences even if they might be statistically significant, if they occur for example in the context of continuous high-voltage discharges. However, it is important to note that in this study ERP assessment and clinical follow-up were carried out completely independently. ERP analysis often predated the clinical interview by several years.

Our most important result is a clear association between N400 presence and a favorable outcome, whereas P300 was not statistically predictive of outcome. This pattern is robust across analysis methods despite variations in overall detection rates, and even held true in an additional analysis where patients were not divided by diagnosis (UWS and MCS) but by etiology (traumatic brain injury [TBI] vs non-TBI patients). Again, for both patient groups and analysis methods, the P300 showed no and the N400 clear prognostic value. Although some studies suggest an association between P300 and a favorable outcome, this finding is not consistent (for review, see Vanhaudenhuyse et al). PPVs reported in UWS range from 33 to 81%; NPVs vary between 0 and 100% for fully comatose and UWS patients together. This wide range could be related to the large variation in P300 detection rates, which will also affect predictive values, explaining the heterogeneity of the literature. In general, data still confirm the Multi-Society Task Force on PVS investigation’s statement: “The presence of P300 evoked responses is not necessarily correlated with the outcome.”

A much brighter picture emerges for the N400 recorded from largely the same patients. Only one of the two previous studies examining N400 in UWS or MCS patients looked at prognosis. Kotchoubey reported a tendency for N400 presence to correlate with a better short-term outcome in a mixed group of UWS and MCS patients. Our study confirms this notion over a much longer follow-up period, differentiating two large groups of MCS and UWS patients and, compared with P300, demonstrating long-term prognostic value to be specific to N400.

With a likelihood ratio >100 for MCS patients and up to 22 for UWS patients, N400 presence is a predictor that can reach the best possible diagnostic evidence (likelihood ratio >10 is conclusive diagnostic evidence). Moreover, it is a better predictor than other previously discussed predictors also calculated in our sample, such as age (p=0.08; likelihood ratio=1.61) and cause (p=0.67; likelihood ratio=1.15).

Both methods used indicate that N400 presence after severe brain injury specifically predicts a favorable clinical outcome in UWS and MCS. The considerable overlap in 95% confidence intervals for the computed measures indicates substantial qualitative similarities in the two methods. The tCWT usually yields higher numbers of patients who show a N400, but visual inspection leads to a higher correspondence between the patients who showed an N400 at initial assessment and those who recover. It must remain unresolved which method better resembles the truth regarding how many patients actually have an N400, but for clinical use the N400 assessments provided by expert visual inspection currently seem to be a more robust indicator of recovery. In our cohort, visual inspection is the more accurate, although more conservative method to assess the presence of N400 as it relates to recovery of individual patients.

N400 presence depends on the synchronized activity of multiple brain regions and therefore reflects the functional integrity of some of the brain’s connections. To produce an N400, a patient’s language network, consisting of widespread tempororothalamic loops, must be intact. Several positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies have shown that UWS patients tend to activate only primary sensory cortices, whereas MCS patients also
activate secondary and association cortices. Nonetheless, Di et al recently showed that some ostensible UWS, but more MCS than UWS patients displayed a N400. Furthermore, PET and fMRI studies were able to reveal that intact or recovered connectivity, as needed to produce an N400, is accompanied by an improvement of cognitive function and the patients’ return to consciousness.

N400 circuitry is not necessarily identical with the connections and regions that give rise to consciousness. However, the present data suggest that if the connections and regions needed to produce a N400 are still intact, there is a chance that the regions and connections subserving consciousness are not too badly injured either or will recover over time. This does not hold to the same extent for the P300. N400 generators may be physically and functionally closer to the consciousness network than networks subserving automatic sound discrimination. Conversely, people with at least partly intact semantic processing may be better able to cognitively benefit from incidental or intentional language stimulation in their environment. This suggests a direction for specific intervention. These residual capacities should be engaged as much as possible by targeted direct communicative language stimulation.

In sum, in the present study, which is to our knowledge currently the largest study of its kind, the N400 ERP is found to be able to predict a patient’s favorable outcome. N400 assessment can help to evaluate an unresponsive patient’s long-term prognosis in UWS and MCS. Language processing in UWS and MCS can easily and objectively be quantified, because EEG is portable and without contraindications. Furthermore, it could easily be implemented in the clinical routine, because other types of EEG-recordings are standard examinations in these patient groups. This could fundamentally change cognitive assessment and outcome prediction for UWS and MCS patients. It may also open a window of opportunity for unobtrusive early intervention and training, with the potential of improving the rehabilitation outcomes of initially unresponsive or low-responsive patients.

Acknowledgment

This work was supported by the University of Konstanz Zukunftskolleg program, Hannelore-Kohl-CNS Foundation 2007013 (K.J), Kliniken Schmieder 015 (K.J), Lur-ija Institute, and state of Baden-Württemberg’s Ministry for Science and Education 31.655.042/Kissler/12.

We thank T. Schertler and S. Klepper for their help with data collection and analysis and all the patients and their relatives contributing to this study.

Potential Conflicts of Interest

Nothing to report.

References

16. Guérin JM, Verougstraete D, de Tourchaninoff M, et al. ERPs obtained with the auditory oddball paradigm in coma and altered


