Chapter 7

Diagnostic properties of P300 analysis techniques in head injury patients: Comparison with imaging data

Jan Willem Elting, Natasha Maurits, Tom van Weerden, Joke Spikman, Jacques De Keyser, Joukje van der Naalt

Submitted
Abstract

Objective: In this study we evaluated diagnostic accuracy of conventional P300 analysis and source analysis in head injury patients. We compared P300 results with neuropsychological test data and imaging data.

Methods: In total, 21 healthy control subjects and 33 patients with head injury were included. Latency and amplitude parameters were obtained for both analysis methods. Neuropsychological evaluation included the Stroop test, the Paced Auditory Serial Addition Test (PASAT) and Rey’s Verbal Learning Test (VLT). Diagnostic accuracy was evaluated with Receiver Operating Curve (ROC) analysis using the neuropsychological test results as the golden standard. Magnetic Resonance Imaging (MRI) was performed between 3 and 9 months after the injury.

Results: For both P300 analysis methods, only amplitude parameters were correlated with neuropsychological test data. Diagnostic accuracy was better with source analysis when compared to the VLT (p=0.03) and the PASAT (p=0.04, only for those patients with imaging abnormalities). Correlation with MRI data showed a non-linear trend between contusional severity and P3A amplitude. Patients with normal MRI results and patients with severe diffuse MRI abnormalities had decreased mean P3A amplitude compared to controls. ERP task performance was worse in patients with normal MRI results when compared to controls and patients with MRI abnormalities. In patients with focal frontal or temporal injury, the presence of mediofrontal lesions was associated with P3A amplitude reduction, while orbitofrontal lesions tended to increase P3A amplitude.

Conclusions: Source analysis of P300 resulted in improved diagnostic accuracy in head injury patients. In view of the non-linear trend between contusional severity and P3A abnormalities, it is advisable to analyze P300 results in conjunction with imaging results. This approach may further facilitate the interpretation and diagnostic applicability of source analysis P300 results in head injury patients.
Introduction

The major long term consequences of traumatic brain injury are residual neuropsychological sequelae. These include fatigue, poor memory and concentration, and emotional symptoms such as depression and irritability, together called the postconcussive syndrome. Neuropsychological research has revealed reduced information processing speed as one of the most fundamental cognitive deficits, which may be essential in producing the attentional and memory disturbances. This is known as the theory of cognitive slowing, which states that there is a reduction of speed at which various cognitive operations can be executed after head injury. In attempts to elucidate the pathophysiology of these phenomena, Event Related Potentials (ERPs) have often been used, because their latency is thought to reflect basic central information processing time. In concordance with the cognitive slowing theory, the latency of the most widely used ERP, the P300 potential, is usually delayed in patients with head injury. This phenomenon has raised interest in using P300 latency as a diagnostic or screening test for cognitive dysfunction.

In patients with moderate to severe head injury, P300 latency is reported to be delayed, and often P300 amplitude is also reduced. In those studies where P300 latency is not significantly delayed, amplitude reduction is usually present. In mild head injury patients, the results are more variable. Increased latency, reduced amplitude or even no change in P300 parameters have all been reported. Furthermore, P300 abnormalities can sometimes be demonstrated when neuropsychological test results are normal. This suggests that P300 measurements may be even more sensitive than standard neuropsychological testing for detection of mild cognitive abnormalities after head injury.

Despite these favorable characteristics, sensitivity and specificity in a clinical setting have been variable, which casts doubt on the clinical utility of the method. Suggestions for improvement have included standardization of measurement techniques, adequate artifact rejection, screening of individual trials for adequate signal to noise ratio, and identification of factors that contribute to the variability of ERP measurements. Factors of importance during measurement are stimulus characteristics and several subject biological-variables, which should be standardized for reliable P300 measurements and comparisons between patient groups. Another important factor that may explain part of the variable success of P300 testing in head injury patients, is the influence of injury severity as estimated by imaging studies. Mild head injury and focal frontal contusional brain injury do not produce similar ERP and cognitive deficits. Surprisingly, several studies have shown that patients with mild head injury have more abnormalities in ERP parameters, reaction time data and psychometric test scores compared to patients with verified contusional lesions in the frontal lobe on imaging studies. The cause of this phenomenon remains unclear. Thus, there seems to be a non-linear relation between the severity of the head injury and ERP abnormalities, with some of the milder injuries having more abnormalities than severe
injuries. This is an undesirable property for any diagnostic test, which may obscure the true diagnostic potential if left unnoticed. Therefore, improvements in diagnostic properties of P300 testing may be achieved by reducing variability and by recognizing different pathophysiological subgroups of patients. In a previous study we demonstrated that an important cause of P300 latency variability is the existence of overlapping P300 components. The main P300 components are P3A, an earlier and more frontocentrally located component, and P3B, a later and more centroparietally located component. We demonstrated that while conventional P300 analysis identifies P3A components in only a minority (± 25%) of subjects, source analysis shows that both P3A and P3B almost always contribute to the P300 complex in normal subjects, but to varying degrees. Furthermore, source analysis yielded a later mean P3B latency with a smaller standard deviation in a group of control subjects when compared with conventional P300 latency analysis. Theoretically this could lead to better diagnostic properties, if the latency distributions of normal subjects and patients should indeed show less overlap.

In this study, we examined a group of head injury patients and controls with both P300 methods. We also analyzed imaging data to identify subgroups of patients with different degrees of head injury severity. Our first aim was to compare diagnostic properties for both P300 methods, using neuropsychological test results as the golden standard. Our second aim was to compare the P300 results with imaging data, to identify more clearly the relation between damaged brain areas and abnormal P300 results.

**Methods**

**Patients and Subjects**

21 healthy control subjects (8 female, 13 male; age 30 +/- 8) and 33 patients with head injury (8 female, 25 males, age 29 +/- 11) were included in this study. None of the control subjects had a history of neurological illness or head injury. Glasgow Coma Scale (GCS) total scores ranged from 5-15 (mean 11 +/- 3,1) and post traumatic amnesia (PTA) duration ranged from 0-60 days (mean 14,2 +/- 14,2). In 2 patients the GCS score was not reliable due to effects of sedatives and/or anesthetics. In all patients a P300 recording and neuropsychological evaluation was performed between 3 and 9 months after the injury. All subjects and patients gave their informed consent. Approval for this research was obtained from the Ethics Committee of the University Medical Center Groningen.

**Neuropsychological evaluation**

**Stroop test:** As measure for cognitive speed and selective attention we used the Stroop test. This test consists of three subtasks: patients first have to read aloud 100 color names printed in black (Stroop word naming), then name the colors of 100 colored squares (Stroop color naming), and lastly name the colors of 100 words (printed in color) that are themselves color
names, with incongruent color names and ink-colors. The first two tasks reflect general cognitive speed, while the last task requires selective attention to suppress the color names. The time needed for the completion of each card was recorded, resulting in four parameters: Word naming score, color naming score, word-color naming score and the interference score. The interference score is calculated by taking the ratio between the color naming score and the Word-Color naming score. Raw test scores were corrected for age, sex and educational level.

**Paced auditory Serial Addition task (PASAT):** The PASAT is viewed as a measure for divided attention. We used a modification of the test developed by Gronwall and Sampson 28. Patients are asked to add every pair of successive numbers, which are presented at a fixed rate through a headphone, and to respond immediately. Five series of 61 numbers (60 additions) are given with interstimulus-intervals of 3.2, 2.8, 2.4, 2.0 and 1.6 seconds. The total test score consists of the sum of the number of correct additions in each series made by the patient. The raw test scores were also corrected for educational level. In this test, attention must be divided between listening, storing numbers in memory, calculations and responding with the correct answer.

**Rey’s Verbal Learning Test:** Memory was evaluated using the VLT 29. We used the Dutch version of this test, in which patients were required to learn 15 one-syllable words. The words are presented five times through a speaker. After each presentation the patient must name the words remembered. The direct recall score is the sum of the total number of words that are correctly remembered after each presentation. Delayed recall is the number of words remembered after 15 to 30 minutes, during which the patient has to perform some other non-memory or naming task. Test results were corrected for age, sex and educational level.

**Imaging**

All patients were evaluated with MRI between 3 and 12 months after the injury. T1 and T2 weighted sequences were acquired in 10 mm slices in the transaxial and coronal planes. MRI scoring was adapted from previously published criteria 30. The approximate localization of lesions was scored as orbitofrontal, laterofrontal, mediofrontal, premotor frontal, anterior/basal temporal, superior temporal, parietal or occipital. Lesion type was scored as 1: focal axonal injury when hemosiderin depositions were present on T2 flash-images, or as 2: focal atrophy on T1 or T2 weighted images, or as 3: white matter hyperintensities on T2 weighted images. Diffuse axonal injury was scored separately when hemosiderin depositions were present bilaterally in at least both frontal and temporal lobes. The size of the lesions was estimated by visual inspection and coded as less then 1 cm, 1-3 cm, or more than 3 cm in maximal diameter. For comparison with P3A amplitude, patients were grouped into 4 categories of contusional severity; 1: normal MRI, 2: unilateral frontal or temporal abnormalities, 3: bilateral frontal or temporal abnormalities, or unilateral frontal and temporal abnormalities, 4: diffuse abnormalities on MRI, defined as presence of bilateral abnormalities.
in the frontal and temporal lobes. This was aimed specifically at quantifying the amount of fronto-temporal damage, since this was considered to reflect increasing probability of damage to intracranial P3A generators. Occipital, parietal and basal ganglia lesions are not of influence in this classification. This is consistent with current knowledge on brain areas that contribute to the scalp recorded P3A 31.

**P300 recording and analysis**

Conventional analysis and source analysis of P300 were performed by different investigators, and the analysis was ‘blind’ in terms of subject characteristics. All recordings were performed under comparable standard test conditions. The entire recording and analysis procedure was described previously in our study on normal subjects 26. In short, we followed the recommendations on P300 recordings as outlined by Polich 21. A standard oddball paradigm was used, in which subjects were asked to silently count auditory target stimuli. EEG was recorded from the scalp using a 128-electrode cap. One electrode was placed on each earlobe for use as a linked ears reference in the conventional P300 analysis.

**Conventional P300 analysis:** Fz, Cz and Pz traces were produced using a linked ears reference. The largest positive peak occurring after the N1, P2 and N2 components that increases in amplitude from frontal to parietal scalp areas was identified as the P300 wave 32. In cases of bifurcated peaks, the second peak with a central/parietal maximum was selected for P300 latency and amplitude determination 33. In these cases, the earlier peak having a frontocentral maximum was termed P3A and was scored separately. Thus, the P3B component latency corresponds with P300 latency in this method.

**Source analysis:** The averaged data were exported to ASA software (ANT software BV, Enschede, the Netherlands) using an average reference. All P300 topographical maps were first visually inspected in ASA to identify P300 components fields, i.e. the frontocentral P3A component and the centroparietal P3B component. In order to model both P3A and P3B, a 2-dipole model was applied using x=60, y=0 and z=0 as a starting position for dipole 1, and x=−60 y=0 and z=0 as starting position for dipole 2. A model with 1 dipole was also applied, which can also identify both components if multiple peaks in dipole activity are found. A starting position of x=0, y=0 and z=0 was used. For both dipole models, rotating dipole models were used without any constraints. Dipoles were fitted over the interval from 250 ms until 500 ms after target stimulus onset, using the following criteria:

1. Dipoles need to explain at least 75% of the variance in individual EEG data, so no more than 25 % residual variance is accepted.
2. Component latency and amplitude are determined using dipole time course information in combination with topographical mapping. Dipole peak activity is attributed to P300 component activity only if a clear frontocentral P3A or centroparietal P3B positive field can be seen in the topographical maps.
3. For the 2-dipole method, crossed dipole solutions (the frontal dipole explains P3B and the posterior dipole explains P3A) and solutions in which two dipoles are equal (i.e. rotate exactly in the same manner and are in close proximity) are not accepted. Instead, the 1 dipole method is used for latency and amplitude determination.

4. After both dipole methods have been tried, the one with the lowest residual variance is used for latency and amplitude determination.

Statistics

For comparison with neuropsychological tests results, both linear and non-linear regression analyses were performed using the curve-fit option in the SPSS software (version 10), also using the uncorrected test results. To investigate sensitivity and specificity of P300 results, Receiver Operating Curve (ROC) characteristics were used with the corrected neuropsychological test results operating as the golden standard. Values below the 10th percentile were considered abnormal. The area under the ROC curve (AUC-ROC) was used for evaluating general diagnostic accuracy, using non-parametric methods for estimating standard errors and using the method described by De Long et al. for comparisons between the P300 methods 34. AUC-ROC values were also tested against the value of 0.5, in which a significant result indicates better diagnostic accuracy than would be expected by chance. For comparison with imaging data, a Kruskall Wallis test was done, with Bonferroni post hoc corrections for multiple comparisons. Furthermore, linear and non-linear trends were tested across imaging subgroups. Further analysis of the influence of lesion localization, size, side and type was performed with a multiple regression analysis using a backward stepwise strategy and a probability of F for removal $\geq 0,1$ and probability of F for entry $\leq 0,05$.

Results

Neuropsychological test data

The neuropsychological data are summarized in table 1. Five patients were lost to follow up before neuropsychological evaluation could be completed. These were all patients with a mild contusion, with GCS total score $\geq 14$ and PTA no more than a few hours. In one patient the Stroop data were not available. In 10 patients the PASAT data were incomplete, because the patients could not complete the test series due to early fatigue.

P300 data

Task performance quantified as the percentage of correctly identified targets was not different between controls (median 100%, 90% central range 96,8-100%) and patients (median 100%,
Chapter 7

Raw test scores

<table>
<thead>
<tr>
<th>Item</th>
<th>Score (sec.)</th>
<th>ISI (sec)</th>
<th>Score (n correct)</th>
<th>Score (n correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word card</td>
<td>53.6 ± 16.1</td>
<td>3.2</td>
<td>48.1 ± 6.1</td>
<td>39-58</td>
</tr>
<tr>
<td>Color card</td>
<td>67.7 ± 15.3</td>
<td>2.8</td>
<td>46.7 ± 5.7</td>
<td>36-56</td>
</tr>
<tr>
<td>WC card</td>
<td>101 ± 29.0</td>
<td>2.4</td>
<td>42.5 ± 7.7</td>
<td>29-55</td>
</tr>
<tr>
<td>IF score</td>
<td>0.68 ± 0.10</td>
<td>2.0</td>
<td>39.6 ± 9.7</td>
<td>26-59</td>
</tr>
<tr>
<td>Total</td>
<td>207 ± 36.5</td>
<td>134-268</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scores corrected for age, sex and educational level

<table>
<thead>
<tr>
<th>Item</th>
<th>Mean ± SD (percentile)</th>
<th>Range</th>
<th>ISI (sec)</th>
<th>Median +/- iqr (decile)</th>
<th>Range</th>
<th>Item</th>
<th>Median +/- iqr (decile)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word card</td>
<td>19.6 ± 22.3</td>
<td>0-88</td>
<td>Total</td>
<td>2 ± 3.3</td>
<td>1-8</td>
<td>Direct</td>
<td>3 ± 6.8</td>
<td>1-10</td>
</tr>
<tr>
<td>Color card</td>
<td>20.3 ± 21.7</td>
<td>0-86</td>
<td>Recall</td>
<td>4 ± 5.0</td>
<td>1-9</td>
<td>IF score</td>
<td>54.7 ± 27.8</td>
<td>8-99</td>
</tr>
<tr>
<td>WC card</td>
<td>36.4 ± 30.1</td>
<td>1-99</td>
<td>Total</td>
<td>2 ± 3.3</td>
<td>1-8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IF score</td>
<td>54.7 ± 27.8</td>
<td>8-99</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Neuropsychological test data. The upper half of the table shows the uncorrected test scores. The lower half shows the data corrected for age, sex and educational level. ISI = interstimulus interval, IF score = interference score, WC card = word/color card. Iqr = interquartile range.

90% central range 85.2-100% (p=0.24). ERP task performance was also calculated in subgroups of imaging abnormalities. Median performance was 100% (96.7-100) for controls, 97% (83-100) for patients with normal MRI results and 100 (96.0-100%) for patients with abnormal MRI results. Performance was different between groups of imaging abnormalities (KW-χ² 10.6: p=0.005), with post-hoc differences between controls and the normal MRI group (p=0.018) and between the normal and abnormal MRI group (p=0.021).

Using conventional analysis, mean P300 latency was delayed in patients (p=0.04) compared to controls, and amplitude was also lower (p=0.01). P3A components were visible very infrequently, precluding statistical comparison. With source analysis, P3B latency and amplitude were not different between the groups. However, P3A components could only be identified in 67% of patients compared to 100% in controls (p=0.001). When all absent P3A
values in the patient group were left censored with a value of 34 (lowest P3A value=35), P3A amplitude was lower compared to controls (p=0.002).
Relation between test results and P300 data

**Stroop test:** No relation could be demonstrated between color and word naming scores and latency parameters for both the conventional and the source analysis method. However, an inverse linear relation between P300 amplitude and the word naming score was found in the conventional P300 method \( (p=0.049) \), but this could not be demonstrated for the color naming score \((p=0.1)\). For the source analysis method, an inverse relation existed between target P3A amplitude and word naming score \((p=0.001; \text{figure 1})\), and color naming score \((p=0.011; \text{not shown})\). The interference score was not related to latency or amplitude parameters for both P300 methods.

**PASAT:** For the conventional P300 method, no relation between the PASAT total score and latency and amplitude measures was found. For the source analysis method, a non-linear relation between P3A amplitude and PASAT total score \((p=0.025)\) was present (figure 1), but source analysis latency parameters and P3B amplitude were not related to the PASAT score.

**VLT:** For both P300 methods, latency parameters were not related to the direct score or delayed recall score. P300 amplitude related with the direct score in a non-linear fashion \((p=0.026)\), but this could not be demonstrated for the delayed recall score. P3A amplitude for the source analysis method also had a non linear relation with the direct score \((p=0.055; \text{not shown})\), and with the delayed recall score \((p=0.011; \text{figure 1})\).
Receiver Operating Curve (ROC) analysis

Only P300 variables that were significantly related to neuropsychological test scores were used for ROC analysis. Figure 2 shows the ROC curves for the amplitude data of both P300 methods. Table 2 shows the results of the AUC area calculations and comparisons, and shows the sensitivity and specificity values for the point of least false-diagnostic classification, which is the point nearest to the left upper corner of the ROC figure (figure 2).

AUC-ROC values were different from the null hypothesis area under the curve (0.5) for the source analysis method when compared with the PASAT results and the VLT results (marginally significant). For the conventional method, no significant differences were found. AUC areas were not different between both P300 methods for the stroop and PASAT tests, but for the VLT a significant difference was found.
### Table 2. Receiver Operating Curve data.

AUC ROC means Area Under the Curve of the receiver operating curve ± Standard Error. Vs area 0.5 tests the null hypothesis of no diagnostic value (AUC-ROC=0.5). The sensitivities and specificities stated are derived from the point closest to the left upper corner of the ROC figure, i.e. the point of least false diagnostic classification (see figure 2). NS = not significant. ROC difference means the difference between the AUC ROC of the conventional and source analysis methods. nA= nano Ampere, μV= micro Volts.

<table>
<thead>
<tr>
<th></th>
<th>All patients</th>
<th>Patients with abnormal MRI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N= 27)</td>
<td>(N= 11)</td>
</tr>
<tr>
<td>% abnormal test score</td>
<td>41</td>
<td>39</td>
</tr>
<tr>
<td>Conventional P300 Amplitude</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC ROC vs area 0.5</td>
<td>0.65 ± 0.11 0.58 ± 0.14 0.42 ± 0.12</td>
<td>0.76 ± 0.12 0.63 ± 0.17 0.46 ± 0.15</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>73</td>
<td>57</td>
</tr>
<tr>
<td>Specificity</td>
<td>62</td>
<td>64</td>
</tr>
<tr>
<td>Cut off point(μV)</td>
<td>14,1</td>
<td>14,1</td>
</tr>
<tr>
<td>Source analysis P3A Amplitude</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC ROC vs area 0.5</td>
<td>0.66 ± 0.12 0.81 ± 0.11 0.72 ± 0.10</td>
<td>0.71 ± 0.13 1.0 ± 0.0 0.74 ± 0.11</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>64</td>
<td>86</td>
</tr>
<tr>
<td>Specificity</td>
<td>81</td>
<td>73</td>
</tr>
<tr>
<td>Cut off point(nA)</td>
<td>no P3A 135</td>
<td>no P3A</td>
</tr>
<tr>
<td>ROC difference</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>
Imaging results

In 2 patients MRI could not be performed due to claustrophobia. In 11 patients normal MRI results were obtained. Table 3 summarizes the MRI data obtained in the remaining 20 patients. A total of 58 lesions were identified, the majority of which were located in the frontal or temporal regions, which are known predilection sites for traumatic brain damage. Within the frontal lobe, the medial regions were most frequently involved, and within the temporal lobe, virtually always the basal temporal areas were involved, with only one patient having a lesion in the superior temporal gyrus. There was no predominance of one specific lesion type, nor was there a clear left/right difference in the distribution of the lesions.

<table>
<thead>
<tr>
<th>Localization</th>
<th>Type</th>
<th>Size (cm)</th>
<th>Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frontal</td>
<td>Axonal injury</td>
<td>22(70)</td>
<td>0-1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22(55)</td>
</tr>
<tr>
<td></td>
<td>Orbital</td>
<td>8(25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lateral</td>
<td>6(30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medial</td>
<td>14(40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Premotor</td>
<td>2(10)</td>
<td></td>
</tr>
<tr>
<td>Temporal</td>
<td>Focal cortical atrophy</td>
<td>19(55)</td>
<td>1-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17(60)</td>
</tr>
<tr>
<td></td>
<td>Basal</td>
<td>12(40)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Superior</td>
<td>1(5)</td>
<td></td>
</tr>
<tr>
<td>Parietal</td>
<td>White matter</td>
<td>17(45)</td>
<td>&gt;3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>19(55)</td>
</tr>
<tr>
<td>Occipital</td>
<td></td>
<td>2(5)</td>
<td></td>
</tr>
<tr>
<td>Basal ganglia</td>
<td></td>
<td>3(5)</td>
<td></td>
</tr>
<tr>
<td>Diffuse</td>
<td></td>
<td>4(20)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Type and distribution of supratentorial intraparenchymal lesions on MR imaging. A total of 20 patients had abnormalities on MRI. Numbers indicate the total number of observed lesions. Numbers in parentheses indicate the percentage of patients (relative to n=20) having that particular lesion.

Relation between imaging results and P300 data

Figure 3 shows the relation between the amount of frontal and temporal abnormalities and source analysis P3A amplitude. In all, 11 patients had no abnormalities in the frontotemporal regions, 6 had abnormalities in 1 area only, 9 had abnormalities in 2 or 3 frontal and/or temporal lobes, and 4 had diffuse axonal injury. One patient with only abnormalities in the basal ganglia was not included in this analysis. When controls were included in the analysis,
P3A amplitude was different between the groups (Kruskall Wallis; p=0.001), with post-hoc differences between controls and patients with normal MRI results (p=0.01) and between controls and patients with diffuse abnormalities (p=0.02), but not between other groups. When testing for trends in P3A amplitude with increasing imaging abnormalities, a cubic curve fitted the data best (p=0.002), illustrating the non-linear relation between contusional severity and P3A amplitude.

Figure 3. Relation between imaging abnormalities and source analysis P3A amplitude. A cubic curve with the equation $y = 217.4 -231.5x + 145.8x^2 -25.6x^3$ (p=0.002) was the best fitting curve, with category values (x-axis) ranging from 0 to 4. Labels: o=orbitofrontal, m=mediofrontal, l=laterofrontal, t=temporal.

In patients with focal contusional abnormalities in frontal and temporal areas, a remarkable inhomogeneity was observed, with some patients having low or absent P3A components, while others had normal P3A component amplitude. In an attempt to find an explanation, a backward regression analysis was performed, using only patients with focal abnormalities on MRI in the analysis (n=15). The dependent variable was P3A amplitude and independent variables included presence or absence of lesions in orbital, lateral, medial, and temporal areas, presence or absence of axonal, focal atrophy and white matter lesions, side (unilateral
or bilateral lesions) and cumulative size of all lesions within a patient. In the final model, only the absence or presence of mediofrontal abnormalities and cumulative lesion size were included. (Adjusted $R^2=0.39$ $p=0.02$). Presence of lesions in medial frontal areas was associated with decreased P3A amplitude, while orbitofrontal lesions tended to increase P3A amplitude (figure 3).

**Post Hoc ROC analysis**

In view of the non linear relation between contusional severity and P3A amplitude, the ROC analysis was repeated in the subgroup of patients with abnormalities on MRI. The results can be found in the right columns of table 2. While similar ROC areas were obtained for both analysis methods compared to the Stroop test and the VLT, the ROC area for the source analysis method vs. PASAT results increased to 1, resulting in a significant difference with the conventional method ROC area ($p=0.04$).

**Discussion**

The main goal of this study was to evaluate any differences in diagnostic accuracy between conventional P300 analysis and source analysis in head injury patients, with neuropsychological test results operating as the golden standard. Source analysis showed better diagnostic accuracy for detecting memory deficits, and also for detecting divided attention disturbances in the subgroup of head injury patients with imaging abnormalities. In line with findings from a previous study, the delay in P300 latency with the conventional method could not be demonstrated with source analysis. Instead, amplitude reduction of the P3A component was present, indicating that the delay found with conventional analysis is a “pseudodelay”. Therefore, P300 amplitude reductions are the most relevant abnormality, which is further illustrated by the fact that only amplitude parameters were correlated with neuropsychological test results in this study.

When comparing the amplitude parameters and neuropsychological test results for both methods, significant correlations were present with measures of cognitive speed (Stroop test) for both methods. However, only for the source analysis method significant correlations with measures of divided attention (PASAT) and memory were found. This can be interpreted as evidence that the P3A component of the source analysis method reflects the amount of attentive resource activation better than conventional P300 amplitude. Source analysis P3A could reflect contributions of mediofrontal and possibly also temporal areas to attention 31, 35, whereas conventional P300 reflects all areas involved in the task. Since impaired attentive resource activation may lead to impaired memory function, the relation between decreased
Chapter 7

P3A amplitude and poor memory task results may be indirect, secondary to impaired attention, or direct, reflecting hippocampal damage, or both.

To examine the relation between anatomical abnormalities and P300 results, imaging data were analyzed. Comparison with imaging results revealed different results for varying degrees of injury severity. Only the patient groups with very severe imaging abnormalities and those without imaging abnormalities had significant reductions of mean P3A amplitude, while focal abnormalities in the frontal or temporal regions affected P3A amplitude in an inhomogeneous way; either low or absent P3A or normal P3A amplitude was found. It remains intriguing why patients with mild head injury without imaging abnormalities have lower amplitude P3A components compared to controls and (although not statistically significant) patients with more severe injury. Apparently, patients with mild head injury allocate less attentional resources to the task than some patients with focal brain injury.

Although this could be interpreted as evidence for subtle cognitive deficits, motivational factors and depression could also play a role in mild head injury patients, as is suggested by the reduced ERP task performance. Another important factor may be “diagnostic threat”; having attention called to a history of prior head injury can result in negative expectations, which results in diminished neuropsychological test performance, as was demonstrated in a few recent studies. These results are consistent with some earlier studies, were similar phenomena for mild head injury patients and frontally damaged patients were demonstrated, with mild head injury patients having lower P3 amplitudes compared to controls and patients with frontal contusions, combined with decreased performance on the ERP task. However, our data show evidence for location specific alterations in P3A amplitude, suggesting that P3A amplitude decrease is caused by mediofrontal damage in severely injured patients, while lesions in the orbitofrontal region may have the opposite effect. Another finding reported in the literature in head injury patients, is increased P3A amplitude in response to novel sounds (i.e. novelty P3A) employed in a 3-tone paradigm, which was interpreted as increased distractibility. This suggests some form of disinhibited stimulus processing, which could in theory be linked to damage in specific frontal sites, such as the orbitofrontal cortex. In recent experiments, patients with orbitofrontal damage showed increased P3 amplitude in response to novel sounds, while patients with damage to the dorsolateral frontal areas did not. To explain these phenomena, a “dynamic filtering” role for the orbitofrontal cortex was proposed, in which the orbitofrontal cortex gates or filters neural activity associated with arousing events, especially if the stimulus has some emotional context. As there is evidence that the same brain processes underlie novelty P3A and target P3A, it is possible that orbitofrontal lesions result in disinhibition of target processing as well.

It is reasonable to assume that attentional networks that are involved in generating P300 activity are influenced by both anatomical factors and psychological factors. Hypoactivity of these networks in mild head injury patients results in reduced P3A amplitude, and is probably at least in part a consequence of behavioural and psychological factors. Focal brain injury
could cause hypoactivity if positioned in the medial frontal areas, or cause an increase in network activity, if positioned in the orbitofrontal regions resulting in normalization, or possibly even an increase in P3A amplitude. P300 testing is therefore probably especially relevant in head injury patients with attentional disturbances due to medial frontal lesions. In a clinical setting, the orbitofrontal phenotype is encountered more often, which is characterized by disturbances in social behavior and emotional disinhibition. In these patients, P300 testing is less useful for quantifying cognitive disturbances, however paradigms with novel stimuli may be useful for assessing disinhibition and distractibility.

One important consequence of the non-linear relation that we found between the amount of damage on MRI and P3A amplitude is the need for careful patient selection for P300 testing in clinical settings. When applied to the whole spectrum of head injury severity, the finding of abnormal amplitude cannot be interpreted in an unambiguous way; it may result either from mild head injury, without evidence for brain damage on imaging studies, or from medio-frontal contusional damage. Based on the results from this study, it would be appropriate to consider any P300 abnormality in combination with some or several other measures of injury severity (i.e., MRI) before conclusions are drawn. Diagnostic properties were better in severely injured patients, which suggests that clinical use of P300 testing is especially useful in this subgroup of patients.

From the clinician’s point of view, an essential question would be: does the assessment of P300 have additional value over neuropsychological evaluation in providing prognostic information in patients with mild to moderate head injury? Is it possible to use P300 as a diagnostic test to predict whether patients will experience cognitive disability interfering with work? Early after injury, in patients with cognitive disability, an extensive neuropsychological assessment is hampered by complaints of extreme fatigability. In general, an assessment of several hours is more than a patient can cope with early after injury. The use of P300 results as a quick screening instrument could facilitate the decision of referring patients earlier for intensified rehabilitation including cognitive rehabilitation therapy in the process of resuming work and daily activities.

Some limitations of this study need to be mentioned here. Due to the limited number of subjects in this study, power was only sufficient for demonstrating large differences between the methods. Therefore, a comparison of diagnostic accuracy between head injury groups of different injury severity was not performed. Furthermore, neuropsychological test data could not be completed in all patients, and this may have had a significant impact on the results, especially for the PASAT.

In conclusion, source analysis resulted in better correlation with measures of divided attention and memory than conventional analysis, which increases diagnostic accuracy. However, careful patient selection is necessary, since there is a non-linear decrease in P300 amplitude with increasing imaging abnormalities, because behavioural and psychological factors may play a role in mild head injury. Damage to the mediofrontal areas is associated with a
decrease in P3A amplitude, while orbitofrontal lesions may have the opposite effect. Applying P300 testing to subgroups of patients with different degrees of injury severity based on imaging results may further facilitate the interpretation and diagnostic applicability of P300 results in patients with head injury.

References


