ELSEVIER

Neurobiology of Aging 28 (2007) 194–201

NEUROBIOLOGY OF AGING

www.elsevier.com/locate/neuaging

Brain event-related potentials: Diagnosing early-stage Alzheimer's disease

Robert M. Chapman *, Geoffrey H. Nowlis ✠, John W. McCrary, John A. Chapman ✠, Tiffany C. Sandoval [1], Maria D. Guillily, Margaret N. Gardner, Lindsey A. Reilly

Department of Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, NY 14627, USA

Received 19 May 2005; received in revised form 30 November 2005; accepted 8 December 2005
Available online 20 January 2006

Abstract

A pattern of components from brain event-related potentials (ERPs) (cognitive non-invasive electrical brain measures) performed well in separating early-stage Alzheimer's disease (AD) subjects from normal-aging control subjects and shows promise for developing a clinical diagnostic for probable AD. A Number–Letter task elicited brain activity related to cognitive processes. In response to the task stimuli, brain activity was recorded as ERPs, whose components were measured by principal components analysis (PCA). The ERP component scores to relevant and irrelevant stimuli were used in discriminant analyses to develop functions that successfully classified individuals as belonging to an early-stage Alzheimer's disease group or a like-aged Control group, with probabilities of an individual belonging to each group. Applying the discriminant function to the developmental half of the data showed 92% of the subjects were correctly classified into either the AD group or the Control group with a sensitivity of 1.00. The two crossvalidation results were good with sensitivities of 0.83 and classification accuracies of 0.75–0.79. P3 and CNV components, as well as other, earlier ERP components, e.g. C145 and the memory "Storage" component, were useful in the discriminant functions.

Keywords: Event-related potentials (ERPs); Neurophysiology; Alzheimer's disease; Diagnostic; EEG; Principal components analysis (PCA); Discriminant analysis; P3; CNV; "Storage" component

1. Introduction

A valid and objective biological marker that can reliably distinguish between early-stage Alzheimer's disease (AD) patients and those with normal age-related cognitive deficits is critical to the advancement of both basic AD research and clinical intervention for patients with the disease [3,23,32]. Here we report event-related potential (ERP) brain measures that hold promise for developing a sensitive diagnostic test for AD. This brain marker may allow diagnosis of individuals in nascent stages of AD, consequently facilitating early intervention. Additionally, identifying specific ERP components and task conditions that discriminate AD individuals from like-aged controls may also provide better understanding of brain functions related to Alzheimer's disease.

Our approach uses cognitive ERP signs from non-invasive recordings of brain electrical activity as a diagnostic technique. Some researchers have taken steps in this direction by investigating ERPs in AD patients and studying group differences between AD and control subjects (e.g. [14,16,17,21,24–27]). Many of the results were based on the odd-ball paradigm and measured the P3 component using ERP peak or area measures. Our study employs a Number–Letter paradigm that varies information processing conditions and a more formal measurement method (principal components analysis [PCA]) to identify and measure a number of independent ERP components.

The Number–Letter paradigm generates a number of interesting ERP components. One is the P3, which is larger to relevant stimuli [4,6,18] and may be involved with

* Corresponding author. Tel.: +1 585 275 8679; fax: +1 585 271 3043.
E-mail addresses: rmc@cvs.rochester.edu (R.M. Chapman), tcsandov@ucsd.edu (T.C. Sandoval).
[1] Present address: University of California, San Diego, CLIN-JDP, La Jolla, CA 92093-0603, USA.
✠ Deceased.

0197-4580/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neurobiolaging.2005.12.008

memory modification and context updating [10,11]. Another component evident in this paradigm is the contingent negative variation (CNV) [18,33], which indicates another memory-dependent process: expectancy of an upcoming relevant stimulus [8]. The memory "Storage" component (C250), previously discovered in this paradigm [9], is larger when the stimulus should be stored in short-term memory; in a behavioral probe test, the C250 amplitude was predictive of recall a short time later [10]. Thus, testing with this paradigm allows the measurement of ERP components associated with stimulus relevancy, expectancy, and short-term memory storage, as well as other ERP components.

ERP component amplitudes vary with experimental conditions, and we use the term "component_condition" to refer to the amplitude of a component under a specified experimental condition. We apply discriminant analyses that combine weighted measures of a set of ERP component_conditions to construct functions that classify an individual as belonging to the AD group or Control group and give posterior probabilities of group membership. This classification of individuals provides a more precise evaluation than analyses of mean group differences. Significant mean group differences do not necessarily assure that a significant number of the individuals in those groups can be correctly classified. On the other hand, a significant number of correct classifications of individuals does assure that mean group differences are significant. Additionally, the sensitivity and specificity of this pilot test are assessed.

2. Methods

2.1. Study population

A desirable test for Alzheimer's disease would detect abnormality in an early stage of the disease when "evidence of cognitive dysfunction will be different and less frank than in later stages" [1]. Therefore, we selected participants whose diagnoses were considered early in the course of AD. The participants were diagnosed by physicians who are AD specialists and who were blind to the results of our ERP tests. The AD participants met standard National Institute for Neurological and Communicative Disorders and Stroke and Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria [22]. As further inclusion criteria, the AD participants had Mini Mental State Exam (MMSE) scores of 22 or greater (out of 30; higher scores indicating better performance) [13]. The AD participants were from the Alzheimer's Disease Center at the University of Rochester, including the Geriatric Neurology and Psychiatry Clinic at Monroe Community Hospital. The control participants selected were normal for their age and demographically similar to AD participants, e.g. age (mean [S.D.] years, for AD: 75.8 [4.5] and for controls: 74.2 [4.8]). Exclusion criteria included clinical (or imaging) evidence of stroke, Parkinson's disease, HIV/AIDs, and reversible dementias, as well as treatment with benzodiazepines, antipsychotic, or antiepileptic medications. Of the 12 subjects in the AD group, there were a minumum of four subjects taking cholinesterase inhibitors to treat mild Alzheimer's disease, and at least three subjects were not.

The AD group had nine males and three females; the Control group had nine females and three males. Using a Bausch & Lomb Vision Tester, their acuity (corrective lenses allowed) in the better eye was 20/30 or better. All subjects could read the large, bright numbers and letters used as stimuli in our Number–Letter paradigm.

2.2. The Number–Letter paradigm

The Number–Letter paradigm employed a visual task with memory and other cognitive demands. On each trial it entailed discrimination between stimuli relevant and irrelevant to the task and memory storage of the first relevant stimulus in order to compare it with the second relevant stimulus. Two numbers and two letters flashed individually in random order at intervals of 750 ms preceded and followed by a blank flash [9]. On a number-relevant block, the participant compared the two numbers in each trial for numerical order, the letters being irrelevant to the task. On another block of trials, the numbers were irrelevant and the task was to compare the two letters for alphabetic order. At the end of each trial, the participant said "Forward", "Backward", or "Same" to indicate the numerical order of the two numbers on number-relevant blocks or the alphabetic order on letter-relevant blocks. The numbers and letters were randomly selected (1 to 6, A to F), and the sequences of numbers and letters in the four temporal intratrial positions were randomized (constraint of two numbers and two letters per trial). The stimuli were large (height of 5.3° visual angle), bright (55 cd/m^2), and presented sequentially in the middle of a computer monitor in order to make it easy for the participants to see them. The stimuli were brief (20 ms) to restrict the time at which each stimulus could be processed and to lower the influence of eye movements. Two blocks of 102 trials each were run; one was number-relevant and the other was letter-relevant. The sequence of stimuli in the two blocks of trials was the same so that any differences in the neural responses would not be attributable to differences in the physical stimuli, but rather to the different perceptual/cognitive processing of the same stimuli. Sequences were randomized for each subject. The relevance order was alternately assigned to each subject. The relevance order was balanced for both groups of subjects (half of the subjects received the number-relevant task first, while the other half had the letter-relevant task first). Using behavioral data from the two blocks of trials, the median behavioral performance was about the same for the number and letter tasks (97 and 94%, respectively) and for the first and second task (96 and 97%, respectively). Thus, the relevance order did not impact performance. The data was halved by odd and even trials for later crossvalidation testing.

The subject was instructed and given practice at doing the task for that block (usually only 10 or so trials were needed). Then, the subject performed the task while his/her EEG was recorded. The participants in this pilot set were capable of performing this task after the brief behavioral training. Only correct trials were included in the brain function measures. The median correct was 92% for the early-stage AD group and 97% for the like-aged Control group. The mean (and standard deviation) of the percentage of correct trials for the AD and Control groups were 90 (18.3) and 96 (2.6), respectively. One AD subject answered correctly only 40% of the time. To account for this discrepancy, extra blocks of trials were measured to ensure that this subject had roughly the same number of correct trials enter the analyses as the other subjects had.

2.3. EEG recordings

While the participant was performing the letter or number comparison tasks, scalp electrodes recorded electrical brain activity. Data were recorded with monopolar electrodes (Electro-Cap) from the midline central area (CZ) to linked earlobes. EOG was recorded to monitor eye movements. Frequency bandpass of the Grass amplifiers was 0.1–100 Hz. Beginning 30 ms before each stimulus presentation, 155 digital samples were obtained at 5 ms intervals. Offline, the digital data were digitally filtered to pass frequencies below 60 Hz. The ERPs were based on correct trials and data not rejected for artifacts (mean artifact rejection rate was 1.7%). Artifact criteria were applied to the CZ and EOG channels in response to each stimulus, and those 750 ms epochs were excluded if either voltage range exceeded 200 μV or either baseline exceeded ±250 μV (baseline was mean of 30 ms pre-stimulus). These artifact criteria were designed to primarily reject data with eye blink artifacts. Artifact reduction also depended on averaging to minimize the effects of small eye movements.

2.4. Event-related potentials (ERPs)

To obtain ERPs, the EEG vectors (155 time points) were averaged separately for each of the stimulus conditions in this experimental design. For each subject there were 36 ERPs: [2 relevance (relevant, irrelevant) × 4 intratrial positions × 2 stimuli (Number, Letter) + 2 blanks (B1, B2)] × 2 halves of data (even/odd trials).

2.4.1. ERP components measured by principal components analysis (PCA)

ERP components were identified and measured by principal components analysis (PCA) [30]. This formal multivariate procedure has a number of advantages over peak and area measures. For a discussion of PCA applied to brain ERPs see [7], and for examples of PCA used with this Number–Letter paradigm see [5,8].

In order to provide a common measurement of ERPs for a wide variety of subjects [19], we used additional groups of subjects for the PCA step. Our four groups of twelve subjects each were AD, like-aged Control, Mild Cognitive Impairment, and Young. The 36 experimental conditions × 48 subjects constituted 1728 ERPs (observations), each with amplitudes at 155 time points. This data matrix of 1728 observations by 155 variables was submitted to a PCA [7] using the correlation matrix and Varimax rotation.

The PCA output the temporal waveforms of each ERP component (loadings) (Fig. 2) and the amount (i.e., amplitude or score) of each component for each observation (experimental condition by subject). The waveforms from the PCA were considered the basis set for the measurements of the ERP components' amplitudes (scores).

After the PCA was computed, the resulting ERP component scores were averaged so that means were obtained for relevant and irrelevant stimuli for odd and even trials for each of the subjects (collapsing across intratrial position and stimuli). The component scores for the blank condition were not considered in further analyses. The measures were renamed to reflect both the name of the component and the relevance condition under which it was elicited. This produced 16 component_conditions: one relevant and one irrelevant measure for each of the eight components in the half of the data being examined at a time. These component_conditions were input into the discriminant analyses [30].

2.5. Discriminant analysis

Discriminant analyses developed linear combinations of ERP component scores that revealed differences between ERPs belonging to the AD group as distinguished from the like-aged Control group. These combinations of scores (discriminant functions) then served as the basis for classifying individuals into these groups. The resulting classification accuracy (proportion of correct classifications) was evaluated for the developmental data and with two crossvalidation methods (the jackknifed or one-left-out and the test data). The classification accuracies were statistically assessed with Fisher's Exact Test. Sensitivity and specificity of the diagnoses were also computed.

A stepwise discriminant procedure selected a subset of the ERP component_conditions that produced a good discrimination model. The selected component_conditions (Table 1) were the variables entered into discriminant analyses to develop discriminant functions.

Only half of the data (odd trials) were used initially in our discriminant analysis, leaving the other half of the data (even trials) as test data for crossvalidation testing. The one-left-out crossvalidation procedure determines a discriminant function based on N-1 subjects and then applies it to classify the one subject left out. This was done for each of the 24 training observations (12 AD and 12 control subjects). Thus, the data used to develop the discriminant functions were different from the data being classified. This method achieves a "nearly unbiased estimate" [20].

Table 1
ERP variables (component_conditions) in discriminant analysis: ADs vs. like-aged Controls

1	P3_Rel	P3 component to relevant stimuli (peak 415 ms poststimulus)
2	CNV_Irr	CNV component to irrelevant stimuli (negative wave prestimulus and turns off at about 350 ms)
3	C145_Irr	C145 component to irrelevant stimuli (peak 145 ms poststimulus)
4	CNV_Rel	CNV component to relevant stimuli (negative wave prestimulus and turns off at about 350 ms)
5	P3_Irr	P3 component to irrelevant stimuli (peak 415 ms poststimulus)
6	C250_Irr	Memory "Storage" component to irrelevant stimuli (peak 250 ms poststimulus)
7	SW_Rel	Slow wave component to relevant stimuli (peak about 745 ms poststimulus)

3. Results

3.1. ERPs

ERP waveforms for relevant and irrelevant stimuli are overlaid and shown separately (displaced) for the AD and the like-aged Control groups (Fig. 1). The AD group had smaller ERPs and smaller ERP differences between the relevant and irrelevant stimuli than the like-aged Control group. These effects appear to involve ERP components in a wide variety of time regions.

3.2. ERP components

ERP components were identified and measured by principal components analysis (PCA) (online supplementary data, Table 1). From the PCA, eight components, accounting for 95% of the variance, were retained. (The ninth component, a stimulus artifact with high loadings during the brief stimulus, was not used further in the analyses.) The temporal waveforms of the ERP components are in Fig. 2. In addition to the well-known P3 [4,6,18], contingent negative variation (CNV) [8,33], and slow wave (SW) [8,29], other ERP components, including relatively early ones peaking at 145 and 250 ms, were seen.

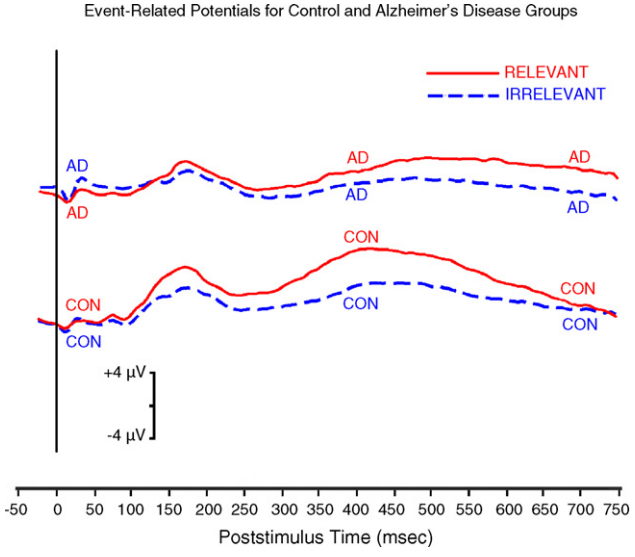

Fig. 1. Event-related potentials (ERPs) averaged over 12 Alzheimer's disease (AD) participants (early stage) and over 12 like-aged Control (CON) participants, graphically displaced. ERPs from the central-midline electrode (CZ) to relevant (solid red) and irrelevant (dashed blue) stimuli are superimposed.

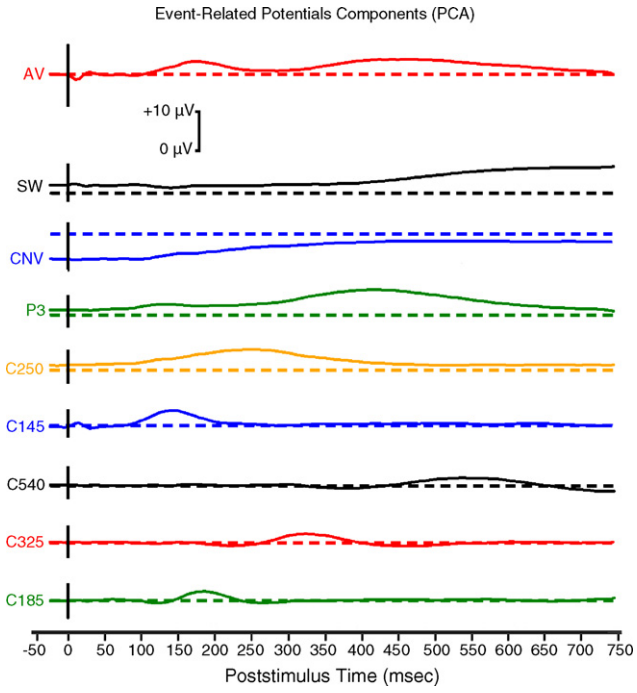

Fig. 2. Event-related potential components derived from principal components analysis across 48 subjects. The top waveform (AV) is the grand average ERP at electrode CZ for all experimental conditions and subjects. The remaining waveforms are the first 8 components from the PCA after Varimax rotation, in order of variance accounted for, from top to bottom. SW = slow wave. CNV = contingent negative variation (negative prior to stimulus). P3 = P300, maximum at 415 ms. C250, C145, C540, C325, and C185 are components with maxima at the poststimulus time (ms) given in their labels. Each component is offset here with its baseline shown as a dashed line. In these component waveforms, the metric has been restored by multiplying the loading at each time point by the standard deviation of the data set at the corresponding time point [7]. The amplitudes depicted are for a component score of 1.0.

The amplitudes of each component for each experimental condition by subject were another output from the PCA (component scores) [7]. These were the ERP component scores used in assessing differences between AD and control individuals.

3.3. Discriminant analyses: developmental data

The seven ERP component_conditions selected (Table 1) were the variables entered into the discriminant analyses to develop a discriminant function (online supplementary data, Table 2). Applying the discriminant function to the

Table 2
Crossvalidation (one-left-out) classification of subjects as Alzheimer's disease (AD) or Control group using discriminant function of 7 ERP component_condition scores from the PCA

(A) Subject	From group	Classified into group	Correct classification	Posterior probabilities of classification into	
				AD	Control
80	A	A	+	0.9969	0.0031
81	A	A	+	0.9675	0.0325
82	C	C	+	0.0014	0.9986
84	A	A	+	0.9929	0.0071
90	A	A	+	0.9371	0.0629
93	C	C	+	0.0502	0.9498
AC	A	A	+	0.9987	0.0013
AE	C	C	+	0.0349	0.9651
AF	C	A		0.5941	0.4059
AG	C	A		0.9298	0.0702
AH	C	C	+	0.0000	1.0000
AJ	C	C	+	0.0332	0.9668
AK	A	C		0.0631	0.9369
AL	C	C	+	0.0016	0.9984
BA	A	A	+	0.6201	0.3799
BB	C	C	+	0.0101	0.9899
CM	A	A	+	0.9992	0.0008
CY	C	A		0.9989	0.0011
DN	A	A	+	0.9084	0.0916
DT	A	A	+	0.9937	0.0063
ET	A	A	+	0.6162	0.3838
FB	C	C	+	0.0175	0.9825
FD	A	C		0.3318	0.6882
FF	C	C	+	0.0168	0.9832

(B) ERP test diagnosis	Clinical diagnosis		Row total
	AD	Control	
T+	10	3	13
T−	2	9	11
Sum	12	12	24

+ Correctly classified.
(A) Classification results for each of the 24 subjects including posterior probabilities of group membership calculated by discriminant analysis using the ERP measures. (B) Summary of individual results for 12 subjects in each group (AD, Controls).
T+ indicates classified as AD by the ERP measures. T− indicates classified as not AD.

developmental half of the data (odd trials), 22 of the 24 subjects (92%) were correctly classified into the AD group or the Control group (sensitivity = 1.00; specificity = 0.83).

These results, which used the same set of data (odd trials) to develop and to apply the discriminant function, are promising. However, to avoid capitalizing on chance, it is important to crossvalidate the discriminant function.

3.4. Discriminant analyses: crossvalidations

The one-left-out crossvalidation results (Table 2) were correct for 19 of the 24 individuals (79% success), a statistically significant result by Fisher's Exact Test ($p < 0.01$). The sensitivity (0.83) and specificity (0.75) were relatively high.

A test-data crossvalidation further evaluated the diagnostic generality. The discriminant function developed with one set of data (odd trials) was tested on a different set of data (even trials) not used in creation of the function. The resulting classification success rates dropped slightly to 18 of 24 subjects (75%), compared to the 79% success rate in the one-left-out crossvalidation (Table 2), but were still statistically significant by a 2 × 2 test (Fisher's Exact Test, $p < 0.01$). The sensitivity for the AD group remained at 0.83, while the specificity dropped to 0.67 with one more control misclassified as AD.

4. Discussion

4.1. Diagnosing AD with discriminant analyses

For the purpose of developing a diagnostic procedure that discriminates between Alzheimer's disease and like-aged control individuals, ERP component scores under particular task conditions were used as the input variables in a discriminant analysis. The resulting discriminant functions (based on combinations of seven of these scores [Table 1]) served as the basis for classifying individuals into the two

groups. The resulting classification accuracy was excellent, in that 92% of the individuals in the developmental data were correctly classified with a sensitivity of 1.00 and specificity of 0.83.

These results used the same set of data (odd trials) to develop and to apply the discriminant functions. In additional analyses designed to better assess the generalizability of the results, the data used to develop the discriminant functions were different from the data used to test them. Two such cross-validation methods were used: the one-left-out and the test data (even trials). As expected, their classification accuracies were lower (0.79, 0.75) but remained statistically significant ($p < 0.01$).

Particularly relevant are the one-left-out results (Table 2), since in this crossvalidation method the subject being tested does not contribute to the development of the classification functions. Thus, this one-left-out procedure estimates the ability to generalize the results to new subjects. The sensitivity (0.83) and specificity (0.75) were relatively high for crossvalidation results.

It is interesting to consider the few individuals who were misclassified in the crossvalidation results (Table 2). First, some of the misclassifications might be due to clinical misdiagnosis of these individuals since we were dependent on clinical judgments. In fact, some of the controls may have Alzheimer's disease and vice versa. If this ERP test is sensitive to very early stages of AD, then one might expect it to sometimes classify a control as AD before clinical symptoms appear. Three of the controls (Table 2) were misclassified as AD. Second, some of the individuals may be difficult to clearly classify by our ERP measures. The posterior probability of membership in each group is available from the discriminant analysis, in addition to the binary decision based on the higher probability. Table 2 shows that most of the correctly classified individuals had much higher probability of belonging to one group over the other (e.g., Subject 81 had 0.97 probability of belonging to the AD group and only 0.03 probability of belonging to the Control group). On the other hand, some misclassified individuals had probabilities closer to 0.50 (e.g., Subject AF had 0.59 probability of belonging to the AD group and 0.41 probability of belonging to the Control group). These posterior probabilities add a quantitative measure to the decision and could indicate that the evidence is "too close to call" one way or the other with those individuals.

For example, if one decided that only probabilities greater than 0.70 be used to diagnose an individual, in the cross-validation results (Table 2) 4 of the 24 individuals would be considered "too close to call". In the remaining individuals, this would reduce the error rate from 20.8 to 15.0%, and increase the sensitivity from 0.83 to 0.89 and the specificity from 0.75 to 0.82.

The success of discriminating AD from controls using brain ERP measures was not due simply to failure of the AD group to perform the Number–Letter task, since only ERP data from correct trials were included in the analyses and the task was performed with high accuracy (the median percent correct was better than 90% in both groups).

Though at least 33% of our AD subjects were taking cholinesterase inhibitors, we were not expecting this treatment to affect the ERP components. Based on the literature, significant P3 amplitude changes related to cholinergic treatment were not found in AD subjects, although modest P3 latency effects were reported [28,34]. Our ERP measures are more like amplitude measures and hence not likely to be very sensitive to cholinergic treatment. If there were ERP effects, one might expect the changes to be more in the direction of ADs looking like controls, and therefore make it more difficult for our measures to discriminate AD from control individuals. The presumed drug effect operates in the wrong direction to account for the AD effect, which was statistically significant. Thus, the literature and the expected direction of cholinergic effects make it unlikely that our reliable ERP differences between AD and controls are due to cholinergic treatment.

Also, there was a gender imbalance between the AD and Control groups (75% versus 25% male). We studied this potential confound through additional analyses that made gender enter the discriminant function either before or after the ERP measures listed in Table 1. With the ERP measures already in the discriminant model, gender did not make an important contribution to the discrimination of AD subjects from control subjects (the average squared canonical correlation increased only 0.681 to 0.688). When gender was entered into the model first and partialled out of the ERP measures, the set of seven ERP measures increased the discriminative power more than 2.5 times (average squared canonical correlation rose from 0.25 to 0.688, a figure nearly the same as the ERP measures alone, 0.681). Thus, the ERP measures played the major role. These additional analyses support that the imbalance in gender do not provide a worrisome confound to the ERP conclusions.

Overall, these analyses indicate that this pattern of seven ERP measures did well in classifying whether individuals belong in the AD or Control groups, and they offer a promising method for diagnosing early-stage Alzheimer's disease. The multivariate methods described here would be easy to use in an automated procedure. The component scores for new ERPs would be computed by using the PCA results already obtained in the test development. Then, the discriminant functions would classify an individual, with computed probabilities, to the AD or Control group, based on their ERP component scores.

4.2. ERP components

A select set of ERP measures was important in discriminating ADs from controls (Table 1). Each ERP measure was the amplitude (score) of a particular ERP component in response to a particular experimental condition (component_condition). For example, the amplitude of the P3 component to relevant stimuli was the first ERP measure

selected for the discriminant function. Consideration of the ERP components and conditions that were important for discriminating ADs from controls may generate ideas not only about improving diagnostic tests, but also about understanding mechanisms that differ in AD.

The discriminatory power of rather early ERP components (C145 and C250, peaking at 145 and 250 ms poststimulus) has not previously been explored for AD. Since both of these early ERP components were smaller for the AD group than for the Control group, this illustrates that it is not only late cognitive components that are different in AD. C145 may reflect perceptual processing of the stimuli [31], suggesting that very early aspects of neural processing may be affected in AD. C250 is the ERP component we have called the "Storage" component, because it has a larger amplitude when a stimulus is stored in short-term memory [2,9,10]. Finding that the median C250 amplitude is smaller for the AD group than the Control group suggests that AD deficits may include storage in short-term memory.

Another problem for AD individuals may lie in their cognitive processing of stimuli that are relevant to the task in which they are engaged. P3, an ERP component with a positive maximum at 415 ms poststimulus, is well-known to be larger in response to stimuli that are task relevant in normal adults [4,6,8,10]. A similar effect of stimulus relevance on P3 was also found in the AD group. However, the P3 amplitudes were considerably smaller in the AD group than in the like-aged Control group (median P3 scores to relevant stimuli of 0.02 and 0.64, respectively). Reduced P3 amplitudes in AD groups have been reported for other tasks [12,15,24,27].

The SW component, a slow wave ERP component with a late maximum at approximately 745 ms, was considerably larger to relevant stimuli for the AD group than for the Control group (median scores of 0.53 and 0.13, respectively). Furthermore, the AD group showed a much larger SW amplitude difference between relevant and irrelevant stimuli (0.34) than did the Control group (0.06). This result illustrates that not all of the ERP components are smaller for the AD group than for the Control group. This is surprising if one expects AD effects to always be in the direction of smaller amplitude measures of brain function. The amplitude of SW has been interpreted to be larger when a stimulus requires more processing [29]. Perhaps this larger ERP component may be showing a compensatory effect in the AD group.

The CNV component also played a prominent role in discriminating ADs from controls. The CNV has been interpreted to represent expectation of a relevant stimulus. The AD group showed a larger component score for CNV to irrelevant stimuli than the Control group (medians scores of 0.58 and −0.14, respectively). This suggests that AD individuals may have problems in anticipating important stimuli.

An interesting aspect of the ERP measures that were selected for discriminating ADs from controls (Table 1) is that brain responses to irrelevant stimuli, as well as to relevant stimuli, were helpful. Of the seven ERP measures selected, ERP components P3, CNV, C145 and C250 were in response to task-irrelevant stimuli and P3, CNV and SW were in response to task-relevant stimuli. It is tempting to consider that a major problem for individuals with Alzheimer's disease lies in their processing of stimuli that are irrelevant to the task in which they are engaged, not just their processing of relevant stimuli.

It is premature to consider these interpretations based on ERP components as definitive, but they do suggest intriguing ideas for subsequent studies that seek to improve diagnostic tests based on ERPs and to understand better the specific nature of Alzheimer's disease deficits.

5. Conclusion

We have demonstrated that the analyses of a pattern of components from brain event-related potentials (ERPs) show promise as a diagnostic tool for detecting individuals as having probable Alzheimer's disease at an early stage. Two crossvalidation methods support the utility of these cognitive non-invasive electrical brain measures as a potential diagnostic test. The success of our method is fostered by using a cognitive task that involves a number of processes. We developed a pattern of ERP component_conditions that discriminates individuals with Alzheimer's disease from like-aged controls by measuring separable brain ERP components with a formal procedure (principal components analysis) and combining these measures in a discriminant analysis. The posterior probabilities of each individual belonging to either the AD group or the like-aged Control group could provide quantitative measures that offer additional precision and context to a diagnosis. Additionally, the ERP components found to be useful in the discrimination may present a means of further understanding the specific nature of Alzheimer's disease.

Acknowledgments

We thank: Pierre N. Tariot, the Geriatric Neurology and Psychiatry Clinic University of Rochester Medical Center, Monroe Community Hospital, the Alzheimer's Disease Center, especially Charles Duffy, Randolph Schiffer, Mark Mapstone, and Richard Kurlan, for their strong support of our research; William Vaughn and Robert Emerson for their technical contributions; Rafael Klorman for critical discussions; Courtney Vargas, Kim Phillips, Jeannie Padowski, and Dustina Holt for technical help; and the many voluntary participants in this research. This research was supported by the National Institute of Health grants P30-AG08665, R01-AG018880, and P30-EY01319.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neurobiolaging.2005.12.008.

References

[1] Barrett G. Clinical application of event-related potentials in dementing illness: issues and problems. Int J Psychophysiol 2000;37:49–53.

[2] Begleiter H, Porjesz B, Wang W. A neurophysiologic correlate of visual short-term memory in humans. Electroenceph Clin Neurophysiol 1993;87:46–53.

[3] Bozoki A, Giordani B, Heidebrink JL, Berent S, Foster NL. Mild cognitive impairments predict dementia in nondemented elderly patients with memory loss. Arch Neurol 2001;58:411–6.

[4] Chapman RM. Evoked responses to relevant and irrelevant visual stimuli while problem solving. Proc Am Psych Assoc 1965:177–8.

[5] Chapman RM, Latent components of average evoked brain responses functionally related to information processing. In: International Symposium on Cerebral Evoked Potentials in Man, pre-circulated abstracts. Brussels: Presses Universitaires de Bruxelles; 1974, p. 38–42.

[6] Chapman RM, Bragdon HR. Evoked responses to numerical and non-numerical visual stimuli while problem solving. Nature 1964;203:1155–7.

[7] Chapman RM, McCrary JW. EP component identification and measurement by principal components analysis. Brain Cogn 1995;27:288–310.

[8] Chapman RM, McCrary JW, Bragdon HR, Chapman JA, Latent components of event-related potentials functionally related to information processing. In: Progress in Clinical Neurophysiology, vol. 6. New York: Basel; 1979, p. 80–105.

[9] Chapman RM, McCrary JW, Chapman JA. Short-term memory: the "storage" component of human brain responses predicts recall. Science 1978;15:1211–4.

[10] Chapman RM, McCrary JW, Chapman JA. Memory processes and evoked potentials. Can J Psychol 1981;35:201–12.

[11] Donchin E. Presidential address, 1980. Surprise!. . .Surprise? Psychophysiology 1981;18:493–513.

[12] deToledo-Morrell L, Evers S, Hoeppner TJ, Morrell F, Garron DC, Fox JH. A 'stress' test for memory dysfunction. Electrophysiologic manifestations of early Alzheimer's disease. Arch Neurol 1991;48:605–9.

[13] Folstein FM, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.

[14] Ford JM, Askari N, Mathalon DH, Menon V, Gabrieli JD, Tinklenberg JR, et al. Event-related brain potential evidence of spared knowledge in Alzheimer's disease. Psychol Aging 2001;16:161–76.

[15] Ford JM, Roth WT, Isaacks BG, Tinklenberg JR, Yesavage J, Pfefferbaum A. Automatic and effortful processing in aging and dementia: event-related brain potentials. Neurobiol Aging 1997;18(2):169–80.

[16] Golob EJ, Starr A. Effects of stimulus sequence on event-related potentials and reaction time during target detection in Alzheimer's disease. Clin Neurophysiol 2000;111:1438–49.

[17] Goodin DS. Clinical utility of long latency 'cognitive' event-related potentials (P300): The pros. Electroenceph Clin Neurophysiol 1990;76:2–5.

[18] Hillyard SA, Picton TW, Electrophysiology of cognition. In: Handbook of Physiology, vol. 5. Bethesda: Wilkins; 1987, p. 519–84.

[19] John ER, Easton P, Prichep LS, Friedman J. Standardized varimax descriptors of event related potentials: basic considerations. Brain Topogr 1993;6:143–62.

[20] Lachenbruch PA. Discriminant Analysis. New York: Hafner; 1975.

[21] Marsh JT, Schubart G, Brown WS. PET and P300 relationships in early Alzheimer's disease. Neurobiol Aging 1990;11:471–6.

[22] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34:939–44.

[23] O'Connor DW, Fertig A, Grande MJ, Hyde JB, Perry JR, Roland MO, et al. Dementia in general practice: the practical consequences of a more positive approach to diagnosis. Br J Gen Pract 1993;43:185–8.

[24] Olichney JM, Hillert DG. Clinical applications of cognitive event-related potentials in Alzheimer's disease. Phys Med Rehabil Clin N Am 2004;15:205–33.

[25] Pfefferbaum A, Ford J, Kraemer HC. Clinical utility of long latency 'cognitive' event-related potentials (P300): The cons. Electroenceph Clin Neurophysiol 1990;76:6–12.

[26] Polich J, Herbst KL. P300 as a clinical assay: rationale, evaluation, and findings. Int J Psychophysiol 2000;38:3–19.

[27] Polich J, Ladish C, Bloom FE. P300 assessment of early Alzheimer's disease. Electroenceph Clin Neurophysiol 1990;77:179–89.

[28] Reeves RR, Struve FA, Patrick G, Booker J, Nave D. The effects of donepezil on the P300 auditory and visual cognitive evoked potentials of patients with Alzheimer's disease. Am J Geriatr Pyschiatry 1999;7:349–52.

[29] Ruchkin DS, Sutton S, Kietzman ML, Silver K. Slow wave and P300 in signal detection. Electroenceph Clin Neurophysiol 1980;50:35–47.

[30] SAS/STAT User's Guide, Release 6.03. Cary: SAS Institute Inc; 1988, p. 359–448.

[31] Silverstein S, Schwarzkopf S, Nowlis G, Chapman J, Nuernberger S, Schenkel L, et al. Electrophysiological correlates of configural and automaticity effects in perceptual organization. Psychophysiology 1997;34:S81.

[32] Smith CW, Byrne EJ, Arie T, Lilley JM. Diagnosis of dementia. II – Diagnostic methods: A survey of current consultant practice and review of the literature. Int J Geriatr Psychiatry 1992;7:323–9.

[33] Walter WG, Cooper R, Aldridge VJ, Mccallum WC, Winter AL. Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. Nature 1964;203:380–4.

[34] Werber E, Gandelman-Marton R, Klein C, Rabey JM. The clinical use of P300 event related potentials for the evaluation of cholinesterase inhibitors treatment in demented patients. J Neural Transmission 2003;110:659–69.