

The Use of Event Related Potentials to Predict Amyloid PET Status Among Patients from a Memory Disorders Clinic

Anna Marin^{a,b,1,*}, Katherine W. Turk^{a,c,1}, Kylie Schiloski^a, Ana Vives-Rodriguez^a, Cheongmin Suh^a, Prayerna Uppal^a, Brigid Dwyer^{a,c}, Rocco Palumbo^a and Andrew E. Budson^{a,c}
^a*Center for Translational Cognitive Neuroscience, VA Boston Healthcare System, Boston, MA, USA*
^b*Department of Behavioral Neuroscience, Boston University Aram V. Chobanian & Edward Avedisian School of Medicine, Boston, MA, USA*
^c*Alzheimer's Disease Research Center, Boston University, Boston, MA, USA*

Accepted 4 June 2024

Abstract.

Background: Amyloid positron emission tomography (PET) scans provide *in vivo* evidence of Alzheimer's disease (AD); however, their high cost limits their use in standard clinical care. Event related potentials (ERPs) may represent an inexpensive and non-invasive additional method for detecting AD pathology.

Objective: We investigated whether ERPs, along with neuropsychological data, serve as predictors of amyloid PET status in patients with memory complaints.

Methods: Veterans aged 50–100 were recruited from a memory disorders clinic. Participants underwent a neuropsychological battery and an ERP auditory oddball protocol. Twenty-eight patients had a positive amyloid PET scan, and thirty-nine patients had a negative scan.

Results: ERP-P200 target amplitude and P200 standard latency were predictors of amyloid PET status. When submitting to ROC analysis, P200 standard latency exhibited the highest specificity and sensitivity in predicting amyloid PET positivity, correctly classifying the amyloid PET status for 86% of patients.

Conclusions: ERP-P200 measures are strong indicators of amyloid- β presence in patients from a memory disorder clinic. Increased P200 amplitude and decreased P200 latency in patients with a positive amyloid PET scan may be attributed to hyperactivation of perceptual bottom-up processes compensating for AD-related synaptic loss in the fronto-parietal networks.

Keywords: Alzheimer's disease, amyloid PET, biomarkers, event-related potentials

INTRODUCTION

Amyloid positron emission tomography (PET) scans provide *in vivo* evidence of Alzheimer's disease (AD) pathology; however, due to their high cost and limited insurance coverage, they are not widely used in clinical settings outside of the VA Healthcare System.¹ This situation may change

with the Medicare coverage of Amyloid PET scans for patients undergoing treatment with Lecanemab. Event related potentials (ERPs) are a time-locked measure of electrical activity that reflect both sensory and cognitive responses to stimuli with a high degree of temporal resolution. ERPs are an inexpensive and non-invasive method of measuring brain function that may support the detection of AD pathology *in vivo*² and thus help gauge the presence of neurodegenerative processes.³ ERPs may ultimately help in triaging patients who need further confirmatory testing with high-cost PET scans that

¹These authors contributed equally to this work.

*Correspondence to: Anna Marin, BA, 150 S. Huntington Ave., Jamaica Plain, MA 02130, USA. E-mail: amarin@bu.edu.

involve radiation or an invasive lumbar puncture procedure.

ERPs are helpful to further understand the interconnectivity between different brain systems and also to detect compensatory mechanisms which may help to overcome neurodegenerative mediated synaptic dysfunction.^{4,5} Synaptic dysfunction is a hallmark characteristic of AD pathophysiology and is directly associated with amyloid- β (A β) deposition in the brain.⁶ AD-related synaptic dysfunction affects a variety of brain networks as both the neocortex and hippocampus gradually lose their normal inputs and ability to function properly.⁷ *We hypothesized that ERPs, especially early and middle evoked potentials, may help predict the presence of amyloid beta deposition among patients in a memory disorders clinic.*

Previous studies have used ERPs to elucidate cortical dysfunction in AD and identify specific ERP features that can better characterize the cognitive profile of AD patients.^{8–10} For example, ERP measures such as N200 and P300 have been found to strongly correlate with disease severity in AD patients,¹¹ while others, such as P200, have been found to correlate with sensory and attentional processes.¹² Although research is still limited, early sensory markers are also potentially useful markers of AD-related degeneration, even in individuals with pre-clinical disease who do not yet show cognitive impairment on neuropsychological testing.¹⁰ Others have used ERPs to help predict the presence of cerebrospinal fluid biomarkers of AD including tau and A β ₄₂⁸ and found that early sensory ERP measures show increased amplitude in mild cognitive impairment (MCI) due to AD compared to non-AD related MCI. In the current study, we explored the potential role of ERPs in characterizing amyloid status in a memory disorders clinic population. We compared the ERP profile of patients with memory complaints due to amyloid PET-proven AD to those of patients with memory complaints who had negative amyloid PET studies, and thus non-AD clinical diagnoses. Previous work has used ERPs in a research-based sample to differentiate cases of MCI due to AD confirmed by amyloid PET from cases of MCI due to non-AD etiologies.¹³ The current study is the first to assess the utility of ERPs as potential biomarkers of amyloid beta positivity in a real-world clinical memory disorders setting. We also investigated whether neuropsychological deficits are associated with electrophysiological changes, and whether ERPs and neuropsychological measures in combination were able to predict amyloid PET status.

METHODS

Participants and initial screening

Between July 2016 and May 2018, a total of 253 participants were consented as part of a larger study evaluating ERPs for the differential diagnosis of dementia in a memory disorders clinic. Participants aged 50–100 were recruited from a memory disorders clinic at VA Boston Healthcare System. All patients were either self-referred for memory complaints or referred to the clinic by their primary care physicians for memory problems. Participants were not excluded based on comorbid conditions or medications, so as to be a representative sample of patients evaluated in a memory disorders clinic. Disease severity was evaluated by the team of trained neurologists at the VA Boston memory disorders clinic (KWT, AEB). Patients met criteria for MCI or dementia as described by the National Institutes of Aging–Alzheimer’s Association.^{14,15}

Participants’ capacity to consent was assessed by asking them to repeat the content of the informed consent form back to us. If participants were unable to do so, we did not enroll them in the study.

Of those 253 participants in the study, 48 were discontinued or excluded because they (i) did not meet the post-EEG inclusion criteria due to low button press accuracy or artifacts pre-processing ($N = 35$), (ii) had poor audiometry testing ($N = 5$), or (iii) experienced problems comprehending and completing the task due to cognitive impairment or agitation ($N = 8$).

Through a review of medical history, we identified a subset of participants who had received a clinical 18F-AV-45 (Florbetapir) amyloid PET scan as part of routine clinical care in the VA Healthcare System. Amyloid PET results were determined by a trained neuroradiologist who had undergone the appropriate, tracer-specific training ($N = 67$; amyloid PET negative = 39, amyloid PET positive = 28) (Fig. 1).

Ethics approval

Initial human research study approval was granted by the Department of Veterans Affairs, VA Boston Healthcare System Institutional Review Board.

Cognitive testing

All participants were administered a neuropsychological battery that includes the Montreal Cognitive Assessment (MoCA),¹⁶ the Consortium to Establish

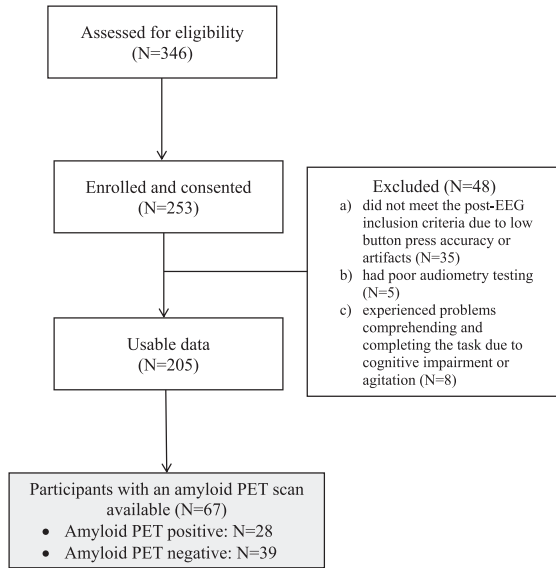


Fig. 1. CONSORT diagram.

a Registry in Alzheimer’s disease (CERAD) Word List Memory Test,¹⁷ Verbal Fluency Test (Category and Letter Fluency),¹⁸ Boston Naming Test-Short Form,¹⁹ Trail Making Tests A and B,²⁰ Mini-Mental State Examination (MMSE),²¹ Geriatric Depression Scale-Short Form (GDS),²² Geriatric Anxiety Inventory (GAI),²³ and the Center for Neurologic Study-Lability Scale (CNS-LS).²⁴ At the time of recruitment, all participants scored 13 or above on the MMSE.²¹

Event-related potential assessment procedure

Data were collected using a seven-electrode COGNISION™ rig in 67 older veterans from the VA Boston Healthcare system memory disorders clinic. EEG/ERP data were recorded while participants completed a three-tone auditory oddball task. The testing procedure was split into three parts: an initial audiometry test, then a 40 stimuli auditory-oddball practice test, followed by a four-hundred stimuli auditory-oddball full test lasting 20 min. All tones were presented through sound-isolating earbuds. Each subject was administered a standard, pure tone audiometry test using the COGNISION System Software. The results from audiometry testing were automatically used by the COGNISION® software to adjust the tone volume to a comfortable level for each participant, thus correcting for any hearing loss. Auditory tone stimuli were amplified to compensate subjects with <30 dB of hearing loss.

Next, participants underwent the auditory-oddball paradigm practice session. Practice session stimuli consisted of 1,000 Hz standard tones (80% frequency) and 2,000 Hz target tones (20% frequency) in pseudorandom order. Participants were instructed to press a button on the handheld set using their dominant hand each time they heard the higher-pitched tone. Participants were given feedback for errors during the practice test and were allowed to take the practice test up to three times to reach at least 80% correct responses. EEGs from the practice session were not used for analysis. The full task consisted of four-hundred frequent standard, infrequent target, and white noise distractor tones appearing with a frequency of 75, 15, and 10%, respectively, in pseudorandom order. The duration for each tone was 100 ms with rise and fall times of 10 ms. The inter-stimulus interval for all the tones was randomized between 1.5 and 2 s. Participants were instructed to press the button as soon as they heard target tones but not for other tones. If participants incorrectly pressed the button for the distractor tone at the start of the task, they were reminded once to press the button only for the same tone they had heard in the practice. No other corrections were given during the full task.

EEG testing and data acquisition

A seven-active electrode COGNISION® EEG rig (Cognision, Louisville, KY) was used to collect ERP data. EEG activity was recorded from sites Fz, Cz, Pz, F3, P3, F4, and P4 sites of the international 10–20 system²⁵ with two linked reference electrodes on each mastoid process (M1, M2) and one common electrode on the frontal bone (Fpz). The EEG rig was chosen for its ease of use which would allow results to have potentially increased clinical applicability as the rig can be used by clinicians without prior EEG/ERP expertise. The headset used for data collection has been validated to perform reliable ERP recording when skin contact impedance is < 70 k Ohms.^{2,26} Other work has provided evidence that higher impedance does not affect statistical power when in a cool and dry environment, such as an open office space.²⁷ Impedance was automatically checked at all electrodes and was kept below this limit throughout testing. Data were collected from –240 to 1,000 ms from stimulus onset, digitized at 125 Hz, and bandpass filtered from 0.3 to 35 Hz except for P3b measurements where the low-pass filter was reduced to 16 Hz. An automatic artifact threshold detection limited of +/- 100 uV was set for the tests. Trial sets

of a deviant tone and the immediately preceding standard tones (epoch sets) with artifacts exceeding the threshold were rejected in real time and immediately repeated. The entire epoch was repeated to reduce the likelihood of participants adapting to the rare stimuli. The real-time utility of online artifact detection allowed for trial repetition, ensuring that the task proceeded only when artifacts were resolved, however, if the task exceeded the maximum limit of 254 artifacts or 60 minutes duration, the session was terminated before completion. Eye blink artifacts were automatically removed as part of pre-processing in real time.

Trial averaging and extraction of ERP measures were automatically performed by the COGNISION[®] System software (Cognition). EEG data from each trial were baseline corrected using the pre-stimulus period and averaged according to stimulus. For standard tones, only the trials immediately preceding target and distractor stimuli were averaged. During data preprocessing, recordings that exceeded two times the root mean square value (RMS) for the EEG test data or with wrong button presses were rejected and excluded from averaging. ERP waves that averaged < 20 trials after preprocessing were eliminated from all analyses² in keeping with prior literature that recommends a minimum of 16 artifact free trials.²⁸

Peak amplitude of the ERP features was measured as the difference between the mean pre-stimulus baseline and maximum peak amplitude. Peak latency was defined as the time point corresponding to the maximum amplitude and was calculated relative to stimulus onset.²⁹ P50 amplitude and latency features were measured for the response to the standard and target trials for 24 to 72 ms. P200 amplitude and latency features were measured for the response to the standard and the target tone trials for 150 to 250 ms. Time windows were determined based on prior literature and by inspecting individual averages and group grand averages. The system selects each ERP parameter at the time window when the signal exhibits the maximal response and then determines its corresponding latency. This extraction procedure has been previously described.^{2,26,30}

The feature extraction algorithm used for the analysis extracted the peak amplitude and peak latency at each electrode site and averaged them all together, taking into account differences in electrode signal contribution and as a result generating two reference scores (peak amplitude and peak latency) for each ERP feature. Each ERP component is initially

extracted as an average of the maximal detected signal at the electrode that shows the greatest value in the time window of interest, and then all other electrodes are examined in the same time interval and averaged afterwards.² Summary measurements for each ERP feature were used in the analysis, as done in previous studies.^{2,26}

For the behavioral data collected during the auditory oddball paradigm, accuracy was calculated as the percent of correct responses to target tones (hits) minus button presses to non-targets (false alarms). To examine reaction time during the auditory oddball task, we performed geometric mean Reaction Time (RT) calculation. RT is defined as the time from the target stimulus onset to the release of the button press.³¹ Geometric mean RT was calculated to reduce the influence of any outlier reaction times.

After EEG testing, audiometry and button press accuracy results for each participant were further examined using the COGNISION[®] software to determine cutoffs for inclusion in data analysis. Audiometry scores were scored using hearing thresholds at 1,000 and 2,000 Hz for each ear. These two frequencies correspond to the stimuli used in the auditory oddball paradigm. Patients whose button press accuracy was below 35% on the 400-tone auditory oddball task were excluded from the analysis to preserve the reliability of the averaged ERP signals.^{2,30}

Statistical analyses

A power analysis was conducted using G-Power.³² Alpha was set to 0.05, and Beta to 0.90. Studies with similar methodologies found large effect sizes when reporting ERP differences.^{2,33} Using a two-tailed independent *t*-test, it was determined that for a large effect size (0.8), a sample size of 34 participants would be needed. Prior studies testing event-related potentials in older adults with and without confirmed AD pathology have used sample sizes of around 17 to 19 participants per group, which is in keeping with our estimated necessary sample size.⁸

Analyses were performed in SPSS (Version 27). Bivariate analyses were performed using independent *t*-tests and correcting for multiple comparisons using Bonferroni correction to evaluate demographic, neuropsychological, and electrophysiological differences between patients who had a positive versus a negative amyloid PET scan. In the results, any *p*-values that are not labeled as “Bonferroni-corrected” represented the uncorrected *p*-value.

Bivariate correlations were performed to assess the relationships between electrophysiological and neuropsychological measures. Logistic regression models were used to evaluate potential indicators of A β status. Regression models were submitted to receiver operator characteristic (ROC) curves to calculate the sensitivity and specificity of variables in correctly classifying A β PET status.

RESULTS

67 participants had an amyloid PET scan done (amyloid PET negative=39, amyloid PET positive=28). Neuropsychological data were collected for all 67 patients. On average, there was a 76-day gap (SD=86) between the neuropsychological/ERP session and the PET scan. Table 1 reports clinical diagnoses including disease stage, primary disease etiology post-PET scan, and psychiatric comorbidities of the two amyloid PET groups. Using Fisher's exact test we found a significant association between disease stage (dementia versus other stages) and PET status ($p=0.021$).

Behavioral data from the auditory oddball paradigm were also analyzed. Table 2 reports the demographic, neuropsychological, auditory oddball

behavioral, and ERP data for the total sample, the amyloid PET negative, and the amyloid PET positive subgroups. Differences between the positive and negative amyloid PET groups are also reported. Patients with a positive amyloid PET scan had a lower score on the MMSE ($M=24.0$, $SD=3.7$, min: 14, max: 30) compared to the amyloid PET negative group ($M=26.2$, $SD=2.2$; $p<0.005$). Amyloid PET positive patients were more impaired on the CERAD word list memory test and when adjusting for multiple comparisons using Bonferroni correction (five comparisons), showed impaired performance on the CERAD delayed recall sub-test (PET pos: $M=3.2$, $SD=1.6$; PET neg: $M=4.4$, $SD=2.0$; $p<0.015$). After adjusting for disease severity, we did not find differences in the MMSE ($p=0.129$) and the CERAD word list memory test ($p=0.170$) between the two groups.

Demographics and neurobehavioral characteristics

ERP characteristics

Significant differences were found at the ERP-P200 measure, with a greater target amplitude and decreased standard latency in the amyloid PET pos-

Table 1
aPET groups' diagnoses

	Amyloid PET+	Amyloid PET-
Disease stage	$N=28$	$N=39$
Subjective cognitive decline	0	4
Mild cognitive impairment (MCI)	19	31
Dementia	9	3
Unclear	0	1
Primary Etiology (Post aPET)	$N=28$	$N=39$
Alzheimer's disease	28	0
MCI due to other etiology	0	8
MCI due to mood disorder	0	6
Dementia with Lewy bodies	0	4
Subjective cognitive decline	0	4
Normal Aging	0	3
Possible chronic traumatic encephalopathy	0	3
Vascular dementia	0	2
Primary psychological disorder	0	2
Unclear	0	2
Corticobasal degeneration	0	1
Korsakoff's dementia	0	1
Frontotemporal dementia	0	1
Obstructive sleep apnea	0	1
Psychiatric comorbidities	$N=28$	$N=39$
None	15	9
Post-traumatic stress disorder	9	25
Unipolar depression	9	19
Bipolar depression	1	2
Anxiety	2	5
Schizophrenia	3	1

Table 2
Demographics

Behavioral data		Total		PET+ (N=28)		PET- (N=39)		P
		Mean	SD	Mean	SD	Mean	SD	
Demographics	Age	68.5	6.6	68.6	6.1	68.4	7.0	0.903
	Y.o.E	13.9	2.5	14.4	2.3	13.6	2.6	0.179
Neuropsychological battery	MMSE*	25.2	3.2	24.0	3.7	26.2	2.2	0.002
	CERAD.E.	15.2	4.0	13.8	3.7	15.9	4.0	0.030
	CERAD.DR.**	3.9	1.9	3.2	1.6	4.4	2.0	0.010
	CERAD.R.	8.8	1.5	8.5	1.5	8.9	1.4	0.244
	CERAD.FPs.	0.5	1.2	0.8	1.5	0.3	0.9	0.053
	CERAD.CR.	8.2	1.9	7.7	2.1	8.7	1.7	0.036
	GDS	5.1	3.5	5.4	3.5	4.8	3.5	0.574
	GAI	6.4	5.9	6.6	5.9	6.3	6.0	0.883
	TMTA (time)	57.0	34.2	61.5	34.7	53.8	33.5	0.373
	TMTA (errors)	0.2	0.5	0.3	0.7	0.1	0.3	0.048
	TMTB (time)	118.4	60.6	107.1	49.7	123.5	65.0	0.423
	TMTB (errors)	0.8	1.2	1.2	1.5	0.6	1.0	0.122
	Letter fluency	30.1	10.2	31.2	10.6	29.2	10.0	0.436
	Category fluency	33.7	11.2	31.7	12.0	35.3	10.6	0.197
	BNT	13.3	1.5	13.2	1.7	13.4	1.5	0.662
	Auditory Oddball Task	BPA	88.0	16.1	87.5	17.5	88.4	15.2
FA		1.9	3.1	2.5	4.0	1.4	2.2	0.141
Geometric Mean RT		449.1	129.2	436.6	136.2	465.3	119.9	0.372

Y.o.E., Years of Education; CERAD E., CERAD Encoding Sum; CERAD DR., CERAD Delayed Recall; CERAD R., CERAD Recognition; CERAD FPs., CERAD False Positives; CERAD CR., CERAD Corrected Recognition; GDS, Geriatric Depression Scale; GAI, General Anxiety Inventory; TMT, Trail Making Test; BNT, Boston Naming Test; BPA, Button Press hits; FA, False Alarms; RT, Reaction time. Significant between-group differences: * $p < 0.05$ and ** $p < 0.015$ for CERAD P -values adjusted using Bonferroni correction for multiple comparisons.

itive group compared to the amyloid PET negative group.

A qualitative review of the grand average waveforms generated for the target and standard tone was first completed. Visually identified differences between the amyloid PET positive and negative groups were observed around the 150–250 ms time period for the standard and target tones (Fig. 2). Visually, the grand average P200 amplitude was greater in the PET positive group compared to the PET negative group at the target tone, and decreased for the P200 amplitude in the PET positive compared to the PET negative group at the standard tone. This is especially visible in the area of the Cz electrode, where the P200 is best visualized.²

We then performed a quantitative analysis of the ERP data. P200 target amplitude was higher in the PET positive group ($M = 5.8 \mu\text{V}$, $SD = 4.1$) compared to the PET negative group ($M = 3.9 \mu\text{V}$, $SD = 2.4$; *uncorrected* $p = 0.029$; *Bonferroni-corrected* $p = 0.116$, Cohen's $d = -0.607$). No difference was found for P200 target latency between the two groups (PET negative: $M = 205.1$ ms, $SD = 20.8$; PET positive: $M = 202.9$ ms, $SD = 22.0$; $p = 0.707$; Cohen's $d = 0.103$). P200 standard amplitude did not differ between the two PET groups (PET negative:

$M = 4.5 \mu\text{V}$, $SD = 2.3$; PET positive: $M = 4.6 \mu\text{V}$, $SD = 2.6$; $p = 0.905$; Cohen's $d = -0.033$). P200 standard latency was shorter in the PET positive ($M = 198.9$ ms, $SD = 22.1$) compared to the PET negative group ($M = 214.4$ ms, $SD = 20.9$; *uncorrected* $p = 0.010$; *Bonferroni-corrected* $p = 0.040$; Cohen's $d = 0.725$) (Fig. 3).

After adjusting for disease severity, significant differences persisted between the two amyloid PET groups in terms of P200 target amplitude ($p = 0.014$) and P200 standard latency ($p = 0.012$).

As outlined in Table 3, in order to determine if the P200 ERP features were independent predictors of amyloid PET status, we completed a series of binary logistic regressions along with age, education, disease severity, and time between the amyloid PET scan and the ERP session. We found that both P200 standard latency ($p = 0.027$) (Model 1) and P200 target amplitude ($p = 0.042$) (Model 2) were significant predictors of A β status (Table 3).

ROC analyses

Regression scores for model 1 (age, education, disease severity, time difference between ERP and PET,

A. Standard Tone

■ = A β PET positive group
■ = A β PET negative group

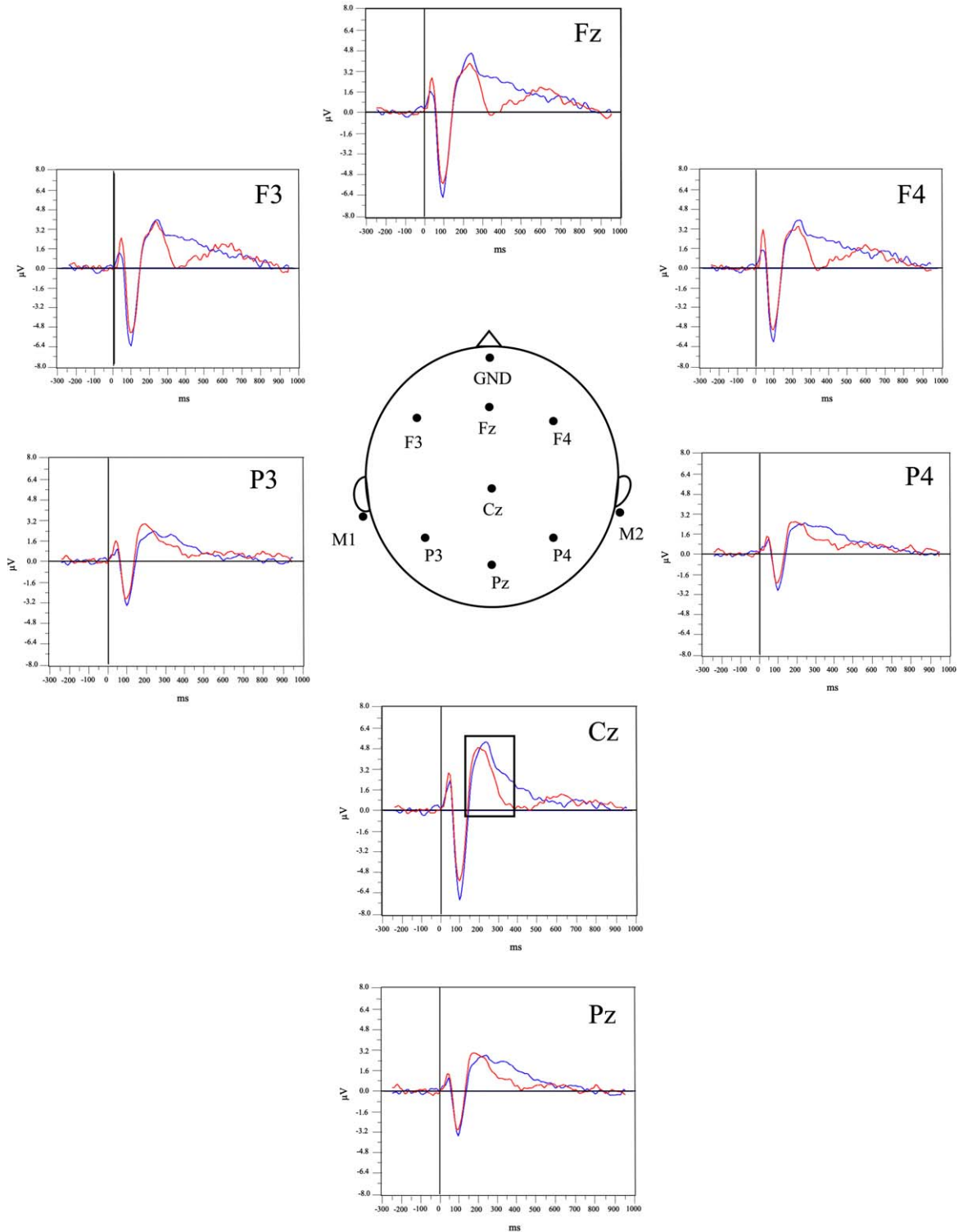


Fig. 2. (Continued)

B. Target Tone

■ = A β PET positive group
 ■ = A β PET negative group

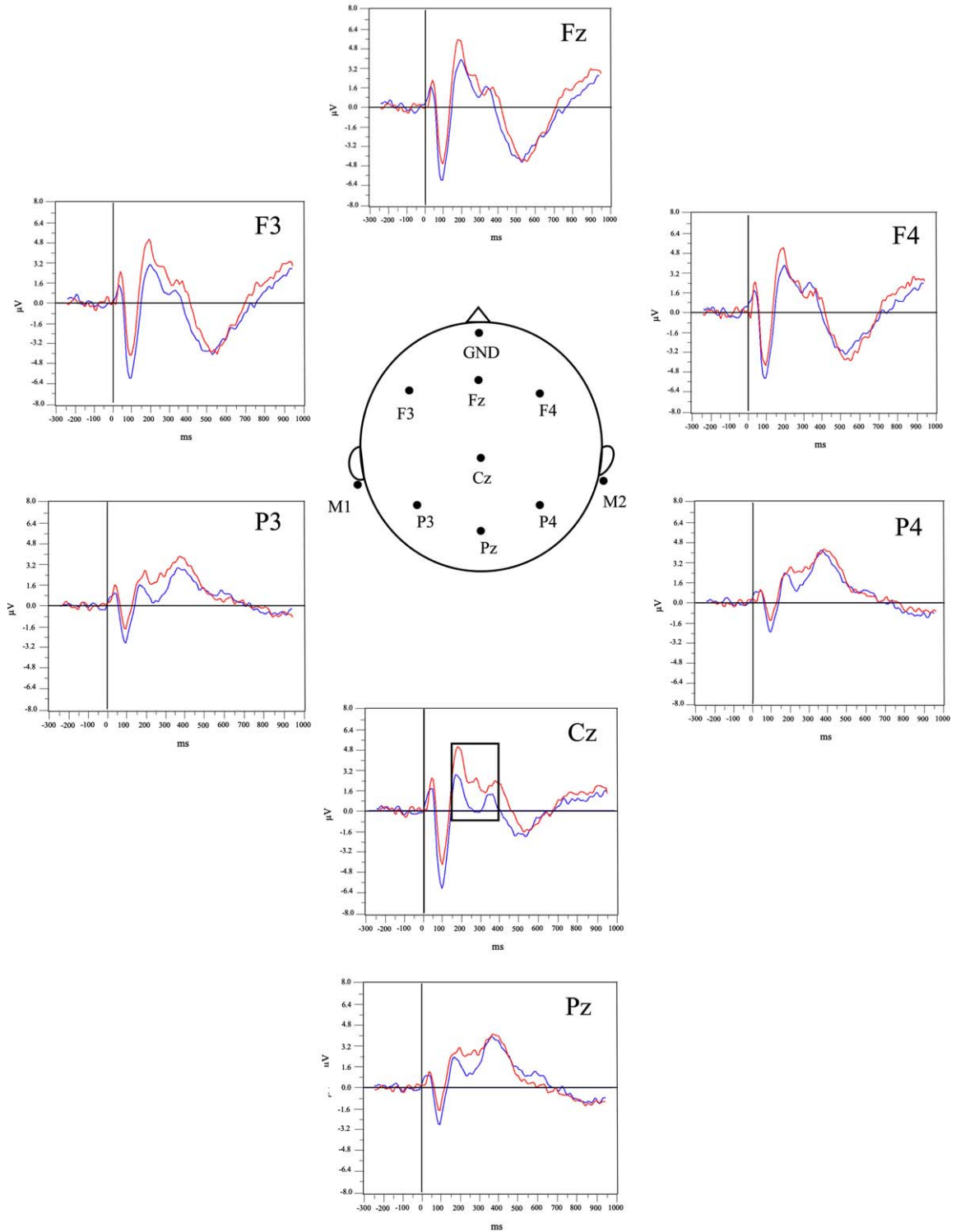


Fig. 2. Grand-average waveforms at the standard (A) and target (B) tones.

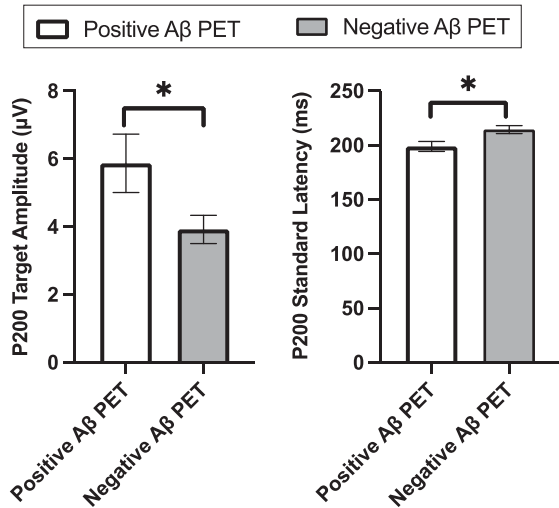


Fig. 3. P200 target amplitude and P200 standard latency.

and P200 standard latency) and model 2 (age, education, disease severity, time difference between ERP and PET, and P200 target amplitude) were submitted to ROC analysis to evaluate the sensitivity and specificity of these ERP measures alone (Fig. 4A, B). Model 1 led to a sensitivity of 86% and a specificity of 72%. Model 2 led to a sensitivity of 100% and a speci-

ficity of 45% (Table 4). We did not find any statistical difference between the two models (P200 standard latency versus P200 target amplitude) ($p = 0.925$).

Correlations between ERPs and neuropsychological data

P200 amplitude and latency at standard and target tones did not correlate with any neuropsychological measures after correcting for multiple comparisons.

DISCUSSION

We found increased P200 amplitude for the target tone and decreased P200 latency for the standard tone in the amyloid PET positive compared to amyloid PET negative patients. P200 target amplitude and P200 standard latency were both significant predictors of amyloid PET status, when controlling for age, education, disease severity, and time lag between the neuropsychological/ERP sessions and the PET scan. When submitting to an ROC, we found that P200 standard latency has the highest specificity and sensitivity in predicting amyloid PET status, and correctly classified 84% of memory disorders patients. The P200 ERP component could help in deciding

Table 3
Binary logistic regression model for Aβ PET positive versus Aβ PET negative

Variable	B (SE)	p	OR	95% CI
Model 1				
Age	0.065 (0.054)	0.230	1.067	0.960–1.187
Education	0.198 (0.145)	0.173	1.219	0.917–1.619
Disease severity	2.406 (0.886)	0.007	11.085	1.952–62.953
Days between PET and ERP	−0.003 (0.004)	0.477	0.997	0.990–1.005
P200 standard latency	−0.037 (0.016)	0.027	0.964	0.934–0.996
Nagelkerke (pseudo R^2) = 0.400				
Model 2				
Age	0.018 (0.062)	0.767	1.018	0.902–1.150
Years of Education	0.196 (0.141)	0.165	1.216	0.923–1.603
Disease severity	2.607 (0.891)	0.003	13.558	2.367–77.666
Days between PET and ERP	−0.001 (0.004)	0.722	0.999	0.991–1.006
P200 target amplitude	0.251 (0.123)	0.042	1.285	1.009–1.637
Nagelkerke (pseudo R^2) = 0.348				

Table 4
Results of ROC analyses

Aβ positive versus Aβ negative	P200 latency (Model 1)	P200 amplitude (Model 2)
Sensitivity (%)	86.4%	100.0%
Specificity (%)	71.9%	45.4%
AUC	0.841	0.798
CI 95%	0.716–0.926	0.667–0.895
SE	0.053	0.0606
Cut-off	>0.3066	>0.1873
P	<0.0001	<0.0001

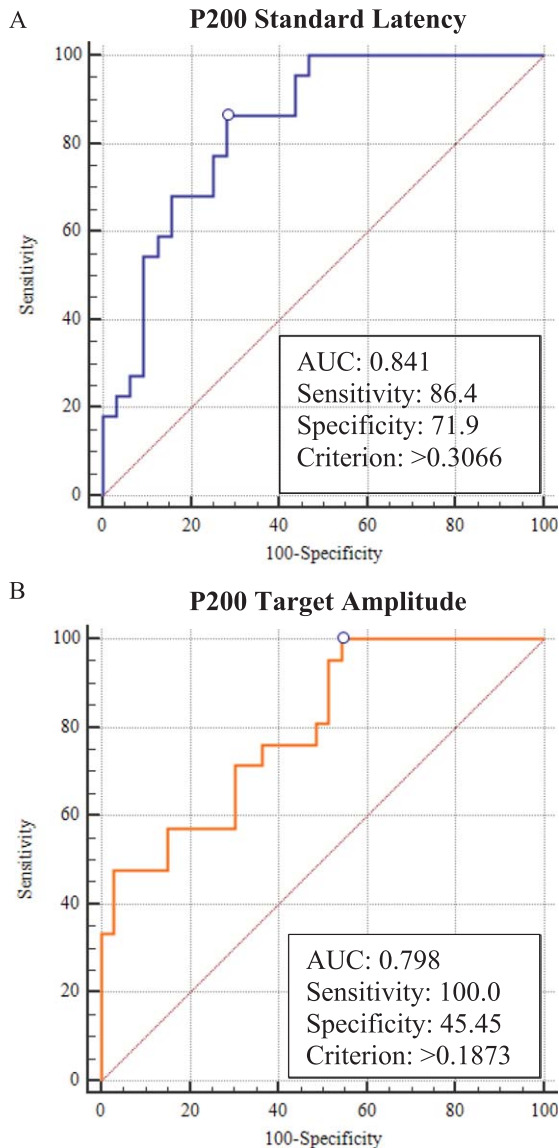


Fig. 4. ROC curves.

whether patients from an heterogeneous memory disorders clinic should undergo diagnostic screening with more invasive and expensive biomarker techniques.

P200 is often interpreted as a measure of perceptual processing involved in early attentional processes and is thought to reflect activation of frontal regions related to stimuli evaluation, as well as to perceptual representation sub-served by posterior regions.^{12,34} Although our findings do not parse out activation at different brain regions, the cumulative hyperactivation found at the P200 time point suggests that AD patients with amyloid beta deposition may recruit

attentional resources differently than non-AD memory patients, and thus may rely more upon bottom-up instead of top-down attentional networks, resulting in changes in the P200 component. This might be potentially caused by synaptic dysfunction and cholinergic deficits related to AD pathology.³⁵

Alterations in P200 amplitude and latency in amyloid PET positive patients may ultimately be related to two complementary mechanisms caused by the disruption of frontal-posterior networks in AD.³⁶ Firstly, as described above, an increase in perceptual processing (bottom-up) activation in AD may partially compensate for the disruption of the frontal-posterior attentional networks.^{10,37} Secondly, sensory gating deficits—especially attention allocation and inhibition—may also explain increases in P200 amplitude in AD.^{38–40} Sensory gating is a phenomenon thought to be associated with filtering out and inhibiting repetitive or unwanted stimuli.⁴¹ Sensory gating deficits in AD are associated with alterations in top-down cognitive function related to impaired inhibition and executive function, which are ultimately mediated by AD-related damage to the frontal cortex.^{42,43}

However, we found that P200 amplitude and latency at the standard and target tones do not correlate with any of the neuropsychological measures collected. This lack of correlation suggests that P200 is a measure of early pre-conscious cognitive processes which occurs too early on to be evaluated using a standard neuropsychological testing battery, but which is still affected by AD-related neurodegeneration.^{8,44} For this reason, the P200 ERP component may represent a potentially powerful measure when used in combination with standard neuropsychological measures as part of the clinical evaluation of memory patients, as it provides distinct information. However, future work should confirm this finding by further examining the correlation between P200 and neuropsychological deficits of attention using more tailored attentional measures which were not incorporated in our standard clinical neuropsychological battery.

Our ERP results are suggestive of sensory gating and potential early attentional deficits in AD versus other memory disorders, likely due to neuronal loss and cholinergic deficits commonly associated with AD,⁴⁵ that are not yet apparent in standard neuropsychological measures of cognitive function. Recent ERP work in AD also shows increased P50 and qualitative differences at the P200 time frame in patients with confirmed CSF A β pathology.⁸ Related work

by Quiroz and colleagues in preclinical carriers of high-risk AD genes (PSEN1 mutation carriers) found an increase in P200 amplitude in occipital regions¹⁰ using a visual recognition memory paradigm.¹⁰ This paradigm tests visual perceptual processing, stimulus evaluation, and episodic memory, including familiarity and recollection processing. Using an auditory oddball paradigm, our study examines overlapping perceptual and attentional cognitive functions while employing an auditory instead of visual paradigm.

Work is still limited on the relationship between P200 latency and sensory gating; however, some initial work has found prolonged P200 latency correlates with better attentional performance, in line with our finding that patients with amyloid positivity had a shorter P200 latency.³⁹ More research is needed to better understand how attentional deficits in AD affect P200 amplitude and latency.

Other studies that have used ERPs to study cortical dysfunction in AD have identified additional ERP features, such as P3b and P50, that aid diagnostic accuracy. For example, Bennys and colleagues assessed cognitive function using ERPs to predict amyloid PET status, with a sample of MCI patients in a non-clinical setting. The authors found decreased amplitude and increased latency in the late ERP feature P3b, commonly associated with declarative memory and decision-making processes, in AD patients with amyloid positive PET scans.¹³ They used P3b as an adjunctive tool with neuropsychological scores, such as semantic fluency, to predict amyloid PET positivity in MCI patients. In their study, they examined individuals aged 70 years and older with spontaneous memory complaints and limitations in one instrumental activity of daily living while excluding participants in the dementia stage of the disease. In our study, we further examined the utility of ERPs as an adjunctive tool to neuropsychological scores in a heterogeneous sample of memory disorders patients who spanned from subjective cognitive decline to dementia. The P200-ERP feature might represent a more generalizable indicator of amyloid PET positivity in a wider range of AD disease stages.

Others have examined the utility of ERPs to detect AD pathology in MCI patients.⁸ They compared MCI patients with and without confirmed AD pathology, based on tau and A β ₄₂ levels in the cerebrospinal fluid (CSF). Using an auditory oddball paradigm, they found P50 was greater in amplitude among MCI patients with positive CSF AD biomarkers. The

authors argue that increased P50 amplitude may be linked to a decreased contribution of the prefrontal cortex which can be attributed to a functional disconnection between the frontal-parietal networks caused by AD-mediated cortical neurodegeneration.

The current study provides new evidence as to the utility of an additional early ERP feature, the P200 component, to help determine the presence AD pathology in a large sample of memory disorders patients recruited from a clinical setting, at different stages of cognitive impairment. We did not find significant differences in P50 and P3b, likely because we investigated a heterogeneous sample that did not solely include MCI patients and that had a more diverse presentation of non-AD etiologies from the clinic.

Previous work has investigated the utility of neuropsychological measures in predicting A β positivity, such as tests of Word Total Recall and Wechsler Memory Scale Logical Memory test immediate and delayed recall,⁴⁶ and the 30-min delayed recall of the Rey Auditory Verbal Learning test.⁴⁷ We found that the CERAD word list memory test, delayed recall measure, and the MMSE are significant indicators of A β positivity in a heterogeneous memory disorders clinic. However, disease severity acted as a confounding factor in our analysis of neuropsychological differences between groups.

The use of ERPs could help improve the current diagnostic practices in a memory clinic, which at the time of the initial visit relies heavily upon the collection of neuropsychological testing measures. ERPs may be a useful additional clinical tool to help determine whether patients from a memory disorders clinic should undergo more invasive biomarker techniques to confirm A β positivity. Future work should investigate how to incorporate easy-to-use and cost-efficient ERP paradigms in the clinical practice to enhance diagnostic accuracy and circumvent the current challenges in providing amyloid PET scans to all patients. ERPs could theoretically be a useful clinical tool in triaging which patients need further confirmatory biomarker testing such as Amyloid PET or lumbar puncture. However, it should be noted that by measuring synaptic dysfunction, ERPs may correlate with tau measures including tau PET to an even greater degree than amyloid PET and that future studies should investigate to what degree P200 and other ERP measures reflect tau burden in AD.

Overall, the present study has a number of strengths. Firstly, study participants represented a

realistic sample of clinical patients seen in a memory disorders clinic. All patients completed the neuropsychological and ERP portions of the study prior to their second visit to the clinic (first follow-up visit), replicating the real-life clinical work up that most patients go through before completing *in vivo* biomarker tests to examine the presence of AD pathology. Secondly, participants came from the memory disorders clinic at the VA Boston. The VA Boston Healthcare System allows for amyloid PET ordering as part of routine clinical care in contrast to other clinical settings, avoiding the sampling bias commonly seen in studies that analyze amyloid PET data retrospectively.^{1,48} Thirdly, our ERP data were collected using a seven-active electrode EEG rig with built-in oddball paradigm, making our results more clinically applicable as clinicians and other clinical staff without prior EEG/ERP expertise can collect and interpret the data in less than an hour.

The present study also has a number of limitations. One important limitation is that, since we tested a memory disorders population from the VA healthcare system, some of the participants also had psychiatric comorbidities such as depression and post-traumatic stress disorder. Future work should replicate these findings in a non-veteran population to confirm that both the ERP and the neuropsychological findings also align in a heterogenous memory disorders clinic outside the VA healthcare system.

Another limitation to consider in this retrospective clinical study is the significant variability in the time gap between the neuropsychological/ERP session and the PET scan. This introduces a confounding variable that may affect the predictive value of the measures on amyloid PET status. We have accounted for this variability by including this variable as a covariate in our models and found our ERP results were still significant predictors of amyloid PET positivity even after accounting for variable time lag between ERP and PET scans. However, future studies should aim to conduct all scans within a 6-month window of the neuropsychological and ERP session to minimize the extent of disease progression over time.

Moreover, this study explored the utility of ERP-P200 in predicting the amyloid PET status of patients who were already showing symptoms related to MCI or dementia. The predictive utility of P200 should also be tested in individuals at the pre-clinical stages of the disease. Given these limitations, future studies should validate our findings in a larger sample of older adults from a memory disorders clinic.

Conclusion

The P200 ERP measure is a promising diagnostic tool to reliably predict the presence of A β plaques among patients presenting to a memory disorders clinic. Changes in P200 amplitude and latency between the PET positive and negative groups may indicate compensatory mechanisms and sensory gating deficits related to underlying neurodegeneration. Further work is needed to evaluate differences specifically related to increased P200 amplitude and decreased latency in the PET positive group. Our findings show that P200 latency differences could be especially useful to evaluate patients with cognitive complaints, aiding disease prognosis and treatment plan. Overall, the P200 ERP feature is a useful marker to help predict A β pathology in patients from a memory disorders clinic. Future work is needed to explore how P200 may correlate with other AD biomarkers including tau and the generalizability of the current findings to non-veteran populations.

AUTHOR CONTRIBUTIONS

Anna Marin (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Validation; Visualization; Writing – original draft; Writing – review & editing); Katherine Turk (Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing – original draft; Writing – review & editing); Kylie Schiloski (Data curation; Formal analysis); Ana Vives-Rodriguez (Data curation; Formal analysis); Cheongmin Suh (Data curation; Formal analysis); Prayerna Uppal (Data curation); Brigid Dwyer (Data curation; Formal analysis); Rocco Palumbo (Formal analysis; Methodology); Andrew Budson (Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing – original draft; Writing – review & editing).

ACKNOWLEDGMENTS

The authors have no acknowledgements to report.

FUNDING

This study was supported by the Alzheimer's Association Clinical Fellowship Research Award

(AACF-16-443347 to KT). VA Career Development Award-2 (IK2 CX002065 to KT). NIH/NIA P30 (AG072978). VA Merit Award (CX001698 to AEB).

CONFLICT OF INTEREST

Katherine Turk has received grant funding from VoxNeuro.

Andrew Budson has been a consultant for Eli Lilly, Genentec and Roche, and a clinical trial investigator for Biogen, Eli Lilly, Genentec, and Cognito and has received funding from VoxNeuro.

All other authors have no conflict of interest to report.

DATA AVAILABILITY

The data supporting the findings of this study is not publicly available due to privacy or ethical restrictions.

REFERENCES

1. Turk KW, Vives-Rodriguez A, Schiloski KA, et al. Amyloid PET ordering practices in a memory disorders clinic. *Alzheimers Dement (N Y)* 2022; 8: e12333.
2. Cecchi M, Moore DK, Sadowsky CH, et al. A clinical trial to validate event-related potential markers of Alzheimer's disease in outpatient settings. *Alzheimers Dement (Amst)* 2015; 1: 387–394.
3. Fruehwirt W, Dorffner G, Roberts S, et al. Associations of event-related brain potentials and Alzheimer's disease severity: A longitudinal study. *Prog Neuropsychopharmacol Biol Psychiatry* 2019; 92: 31–38.
4. Friedman D. Cognition and aging: a highly selective overview of event-related potential (ERP) data. *J Clin Exp Neuropsychol* 2003; 25: 702–720.
5. Kropotov J, Ponomarev V, Tereshchenko EP, et al. Effect of aging on ERP components of cognitive control. *Front Aging Neurosci* 2016; 8: 69.
6. Hampel H, Hardy J, Blennow K, et al. The amyloid- β pathway in Alzheimer's disease. *Mol Psychiatry* 2021; 26: 5481–5503.
7. Olichney JM, Yang J-C, Taylor J, et al. Cognitive event-related potentials: biomarkers of synaptic dysfunction across the stages of Alzheimer's disease. *J Alzheimers Dis* 2011; 26: 215–228.
8. Green DL, Payne L, Polikar R, et al. P50: A candidate ERP biomarker of prodromal Alzheimer's disease. *Brain Res* 2015; 1624: 390–397.
9. Golob EJ, Irimajiri R, Starr A. Auditory cortical activity in amnesic mild cognitive impairment: relationship to subtype and conversion to dementia. *Brain* 2007; 130: 740–752.
10. Quiroz Y, Ally B, Celone K, et al. Event-Related potential markers of brain changes in preclinical familial Alzheimer disease. *Neurology* 2011; 77: 469–75.
11. Parra MA, Ascencio LL, Urquina HF, et al. P300 and neuropsychological assessment in mild cognitive impairment and Alzheimer dementia. *Front Neurol* 2012; 3: 172.
12. Katada E, Sato K, Ojika K, et al. Cognitive event-related potentials: useful clinical information in Alzheimer's disease. *Curr Alzheimer Res* 2004; 1: 63–69.
13. Bennys K, Gabelle A, Berr C, et al. Cognitive Event-related potential, an early diagnosis biomarker in frail elderly subjects: The ERP-MAPT-PLUS Ancillary Study. *J Alzheimers Dis* 2017; 58: 87–97.
14. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; 7: 270–279.
15. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; 7: 263–269.
16. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005; 53: 695–699.
17. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989; 39: 1159–1165.
18. Mueller KD, Kosciak RL, LaRue A, et al. Verbal fluency and early memory decline: results from the Wisconsin Registry for Alzheimer's Prevention. *Arch Clin Neuropsychol* 2015; 30: 448–457.
19. Calero MD, Arnedo ML, Navarro E, et al. Usefulness of a 15-item version of the Boston Naming Test in neuropsychological assessment of low-educational elders with dementia. *J Gerontol B Psychol Sci Soc Sci* 2002; 57: P187–P191.
20. Reitan RM. The relation of the trail making test to organic brain damage. *J Consult Psychol* 1955; 19: 393–394.
21. Folstein MF, 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: 189–198.
22. Sheikh JJ, Yesavage JA. Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clin Gerontol* 1986; 5: 165–173.
23. Johnco C, Knight A, Tadic D, et al. Psychometric properties of the Geriatric Anxiety Inventory (GAI) and its short-form (GAI-SF) in a clinical and non-clinical sample of older adults. *Int Psychogeriatr* 2015; 27: 1089–1097.
24. Work SS, Colamonico JA, Bradley WG, et al. Pseudobulbar affect: an under-recognized and under-treated neurological disorder. *Adv Ther* 2011; 28: 586–601.
25. Klem GH, Lüders HO, Jasper HH, et al. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl* 1999; 52: 3–6.
26. Cecchi M, Adachi M, Basile A, et al. Validation of a suite of ERP and QEEG biomarkers in a pre-competitive, industry-led study in subjects with schizophrenia and healthy volunteers. *Schizophr Res* 2023; 254: 178–189.
27. Kappenman ES, Luck SJ. The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology* 2010; 47: 888–904.
28. Tsivilis D, Otten LJ, Rugg MD. Context effects on the neural correlates of recognition memory: an electrophysiological study. *Neuron* 2001; 31: 497–505.

29. 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: 1438–1449.
30. Turk KW, Marin A, Schiloski KA, et al. Head injury exposure in Veterans presenting to memory disorders clinic: an observational study of clinical characteristics and relationship of event-related potentials and imaging markers. *Frontiers in Neurology* 2021; 12: 626767.
31. Doucet C, Stelmack RM. The effect of response execution on P3 latency, reaction time, and movement time. *Psychophysiology* 1999; 36: 351–363.
32. Faul F, Erdfelder E, Lang A-G, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007; 39: 175–191.
33. Golob EJ, Ringman JM, Irimajiri R, et al. Cortical event-related potentials in preclinical familial Alzheimer disease. *Neurology* 2009; 73: 1649–1655.
34. Fogarty JS, Barry RJ, Steiner GZ. The first 250 ms of auditory processing: no evidence of early processing negativity in the Go/NoGo Task. *Sci Rep* 2020; 10: 4041.
35. Bourisly AK, Shuaib A. Neurophysiological effects of aging: a P200 ERP study. *Transl Neurosci* 2018; 9: 61–66.
36. Oh H, Mormino EC, Madison C, et al. β -amyloid affects frontal and posterior brain networks in normal aging. *Neuroimage* 2011; 54: 1887–1895.
37. Bobes MA, García YF, Lopera F, et al. ERP generator anomalies in presymptomatic carriers of the Alzheimer's disease E280A PS-1 mutation. *Hum Brain Mapp* 2009; 31: 247–265.
38. Boutros NN, Korzyukov O, Jansen B, et al. Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. *Psychiatry Res* 2004; 126: 203–215.
39. Lijffijt M, Lane SD, Meier SL, et al. P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory. *Psychophysiology* 2009; 46: 1059.
40. Shen C-L, Chou T-L, Lai W-S, et al. P50, N100, and P200 auditory sensory gating deficits in schizophrenia patients. *Front Psychiatry* 2020; 11: 868.
41. Jones LA, Hills PJ, Dick KM, et al. Cognitive mechanisms associated with auditory sensory gating. *Brain Cogn* 2016; 102: 33–45.
42. Cancelli I, Cadore IP, Merlino G, et al. Sensory gating deficit assessed by P50/Pb middle latency event related potential in Alzheimer's disease. *J Clin Neurophysiol* 2006; 23: 421–425.
43. Thomas C, vom Berg I, Rupp A, et al. P50 gating deficit in Alzheimer dementia correlates to frontal neuropsychological function. *Neurobiol Aging* 2010; 31: 416–424.
44. Friedman G, Turk KW, Budson AE. The current of consciousness: neural correlates and clinical aspects. *Curr Neurol Neurosci Rep* 2023; 23: 345–352.
45. Ally BA, Jones GE, Cole JA, et al. Sensory gating in patients with Alzheimer's disease and their biological children. *Am J Alzheimers Dis Other Demen* 2006; 21: 439–447.
46. Alves L, Cardoso S, Silva D, et al. Neuropsychological profile of amyloid-positive versus amyloid-negative amnesic mild cognitive impairment. *J Neuropsychol* 2021; 15: e12218.
47. Kandel BM, Avants BB, Gee JC, et al. Neuropsychological testing predicts cerebrospinal fluid amyloid- β in mild cognitive impairment. *J Alzheimers Dis* 2015; 46: 901–912.
48. Vives-Rodriguez AL, Schiloski KA, Marin A, et al. Impact of amyloid PET in the clinical care of veterans in a tertiary memory disorders clinic. *Alzheimers Dement (N Y)* 2022; 8: e12320.