Review Article

Predictive Power of Cognitive Biomarkers in Neurodegenerative Disease Drug Development: Utility of the P300 Event-Related Potential

John Olichney,1,2 Jiangyi Xia,1,2 Kevin J. Church,3 and Hans J. Moebius1,3

1Center for Mind and Brain, University of California, Davis, Davis, CA, USA
2Department of Neurology, University of California, Davis, Sacramento, CA, USA
3Athira Pharma, Inc., Bothell, WA, USA

Correspondence should be addressed to Hans J. Moebius; hans.moebius@athira.com

Received 10 June 2022; Accepted 8 October 2022; Published 8 November 2022

Copyright © 2022 John Olichney et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neurodegenerative diseases, such as Alzheimer’s disease (AD), and their associated deterioration of cognitive function are common causes of disability. The slowly developing pathology of neurodegenerative diseases necessitates early diagnosis and monitored long-term treatment. Lack of effective therapies coupled with an improved rate of early diagnosis in our aging population have created an urgent need for the development of novel drugs, as well as the need for reliable biomarkers for treatment response. These issues are especially relevant for AD, in which the rate of clinical trial drug failures has been very high. Frequently used biomarker evaluation procedures, such as positron emission tomography or cerebrospinal fluid measurements of phospho-tau and amyloid beta, are invasive and costly, and not universally available or accessible. This review considers the functionality of the event-related potential (ERP) P300 methodology as a surrogate biomarker for predicting the procognitive potential of drugs in clinical development for neurocognitive disorders. Through the application of standardized electroencephalography (EEG) described here, ERP P300 can be reliably measured. The P300 waveform objectively measures large-scale neuronal network functioning and working memory processes. Increased ERP P300 latency has been reported throughout the literature in disorders of cognition, supporting the potential utility of ERP P300 as a biomarker in many neurological and neuropsychiatric disorders, including AD. Specifically, evidence presented here supports ERP P300 latency as a quantitative, unbiased measure for detecting changes in cognition in patients with AD dementia through the progression from mild to moderate cognitive impairment and after drug treatment.

1. Introduction

Neurodegenerative diseases, such as Alzheimer’s disease (AD), Parkinson’s disease (PD), and Huntington’s disease (HD), can lead to dementia, which affects nearly 55 million people worldwide [1]. The neurodegenerative processes start decades before the appearance of clinical symptoms [2–6]. After diagnosis, the clinical course and life expectancy of patients with chronic neurodegenerative conditions often span many years, necessitating the long-term use of therapeutics. However, the progressive nature of neurodegenerative diseases poses a significant challenge for drug development. As the result of budgetary constraints and risk considerations, many early translational trials of drugs for these diseases had to be limited in size and/or duration [7]. Reliable cognitive assessment is particularly difficult because of placebo response and slow clinical decline. Longer trials (≥6 months) are usually necessary to allow for separation of trial drug effects from a placebo arm [2, 8]. Additionally, a leading cause of failure of central nervous system (CNS) drugs in clinical trials is use of the...
incorrect dose, which may only become apparent during phase 3 trials [9, 10]. Furthermore, the disease may be too progressed in patients with neurodegenerative diseases often selected for clinical trials [9]. Therefore, the exceptionally high failure rate of clinical trials for AD, coupled with the increasing rate of diagnosis has led to an urgent need for the development of novel therapeutics [7]. Despite this demand, the methods available for early prediction of clinical outcomes following pharmacological interventions are limited. Methods such as positron emission tomography are cost-prohibitive, and the cerebrospinal fluid measurement of phospho-tau and amyloid beta is invasive. Furthermore, these methods have limited predictive value [11]. Together, these factors underscore the need for objective, quantitative, reliable, noninvasive, repeatable, and cost-effective approaches for the assessment and monitoring of cognitive status. This review explores the ability of recording neurophysiological activity using the event-related potential (ERP) P300 to serve as a biomarker of cognition in clinical trials. The review focuses on the P300 component of the ERP waveform, including important technical aspects of the methodology, and summarizes the existing evidence of its translational utility in evaluating potential pro-cognitive properties of pharmaceutical interventions in AD, other neurodegenerative diseases, and neuropsychiatric disorders.

2. Event-Related Potentials (ERPs)

ERPs and evoked potentials (EPs) are particularly useful applications of electroencephalography (EEG), in that they capture averaged brain responses that are task-based and time-locked and thus, are associated with specific sensory and motor EPs or cognitive events. As a task-based methodology, and thus distinct from non-task-based quantitative electroencephalography (qEEG; Figure 1), the ERP is a highly suitable measure for correlating neuronal function with cognitive processes, such as working memory and executive function. The ERP waveform, which is composed mainly of summed inhibitory and excitatory postsynaptic potentials, measures the activity of networked neurons firing in synchrony, or integrated synaptic activity [3, 12]. Thus, ERPs can provide a neural correlate of working memory load and quantify other higher-level cognitive processes [3]. The ERP measurement is particularly appropriate for the evaluation of synaptic disorders, such as AD, because it reflects the spatial processing speed within large neuronal networks [12]. The superior temporal resolution of ERPs makes them well suited to detecting neural response to task manipulations (in milliseconds) and to distinguishing and quantifying both the early (generally ≤200 ms) and later stages of cognitive processing in individuals, including those with mild-to-moderate stage
dementia [12, 13]. Thus, the superior temporal resolution of ERPs enables pharmacodynamic changes to be detected within short time frames and allows the results to be compared with baseline data [14]. Capturing changes in cognitive function utilizing ERPs may serve as a promising noninvasive biomarker that may increase confidence in the detection of procognitive effects of drugs during later stages of development.

The deflections that occur within the ERP waveform depend on the task or stimuli presented. The early waves are most closely associated with sensory processing of a presented stimulus and...
usually occur within ~100 ms of stimulus onset [15]. The later stages of the ERP signal, i.e., those occurring after 200 ms, correspond to controlled attention, working memory access, and integrative processing (e.g., semantic and emotional) of the stimulus. The N component of the ERP waveform refers to the negative voltage deflection, whereas the P component refers to the positive voltage peak (Figure 2).

The number following N or P refers to the typical time from presentation of a stimulus to the peak or deflection (in milliseconds), or latency.

The P50 waveform is an early positive peak elicited by paired-click or steady-state paradigms [15] related to “sensory gating” that has been useful in studies of bipolar disorder and schizophrenia [16, 17]. The N100 and P200 waveforms represent the automatic sensory process responses elicited by auditory stimuli, whereas the N140 and N170 waveforms are elicited by visual stimuli. The P300 waveform appears in response to active engagement in the detection of task-relevant target stimuli. The N200 waveform, and associated “mismatch negativity,” is an additional late response to an infrequent or unexpected stimulus that is often studied in conjunction with the P300 waveform [3]. Additional late components of the ERP—the N400 and P600 waveforms—are elicited in language-processing contexts, such as semantic incongruity and syntax error processing. A late positive component that also has a positive peak at approximately 600 ms following stimulus onset has been shown to be one of the best predictors of human verbal memory ability [18–20].

The P300 component of the ERP waveform, which is considered a later-stage ERP signal, is the most widely analyzed ERP in cognitive research. The P300 component is characterized by the time it takes for the peak to occur (latency), normally ~300 ms after stimulus onset, and by its amplitude. The P300 peak latency reflects the timing of brain activity associated with accessing working memory and performing higher executive functions. The P300 component can be broken down further into subcomponents. The P3a subcomponent occurs in response to distracter (task-irrelevant) stimuli (e.g., a dog barking) and is associated with attentional processes involved in automatic novelty detection. The P3b subcomponent is evoked by target (task-relevant) stimuli [21]. In normal aging, age-related amplitude reduction, latency prolongation, and topographically more frontally oriented P300 have been consistently reported across studies [22]. Decreases in P300 amplitude and increases in latency more severe than those in normal aging have been reported in patients with dementia, psychiatric disorders, alcohol dependence, and traumatic brain injury (TBI), and in neurodevelopmental disorders [23–25]. The utility of the P300 latency as an objective measure in detecting changes in the cognitive performance for the application of testing therapeutics developed for neurodegenerative and neuropsychiatric disorders with associated cognitive deficits is discussed further in Section 5.

3. ERP P300 Methodology

ERP P300 is a large positive waveform (normal amplitudes of up to 10–20 μV) that typically occurs with a latency, or time to peak, that ranges from 250 to 500 ms poststimulus onset [3]. Recommendations for standardizing the ERP P300 methodology have been made previously [23]. Electrodes for recording ERP P300

Figure 3: Schematic of the ERP P300 waveform in the progression of AD. Shown is a representative ERP waveform for neurotypical brain activity and normal cognition (b). A characteristic response is detected 300 ms after presentation of stimuli (P300). Early in disease onset, and usually prior to the diagnosis of AD, neural network communication is disrupted, leading to impaired brain connectivity and impaired cognition. This change in cognition can be detected as a change in the peak and/or latency (time to onset of peak) of the P300 response to stimuli (c).
<table>
<thead>
<tr>
<th>Study</th>
<th>Population (n)</th>
<th>P300 latency</th>
<th>P300 amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ally et al., [54]</td>
<td>HC (80) AD (80)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td></td>
<td>HC (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bennys et al., [47]</td>
<td>MCI (20) AD (30)</td>
<td>AD &gt; MCI &gt; HC</td>
<td>HC &gt; MCI = AD</td>
</tr>
<tr>
<td>Caravaglios et al., [63]</td>
<td>HC (16) AD (21)</td>
<td>AD &gt; HC</td>
<td>HC = AD</td>
</tr>
<tr>
<td>Cecchi et al., [105]</td>
<td>HC (101) Mild AD (103)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Cintra et al., [106]</td>
<td>HC (14) MCI (34) AD (17)</td>
<td>(ApoE4+) MCI + AD &gt; HC</td>
<td>HC = AD</td>
</tr>
<tr>
<td>Frodl et al., [48]</td>
<td>HC (26) MCI (26) AD (30)</td>
<td>AD &gt; MCI &gt; HC</td>
<td>HC &gt; MCI = AD</td>
</tr>
<tr>
<td>Fruehwirt et al., [107]</td>
<td>Mild AD (31) Severe AD (32)</td>
<td>Severe &gt; mild</td>
<td>NA</td>
</tr>
<tr>
<td>Golob and Starr, [32]</td>
<td>HC (12) AD (10)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Hirata et al., [40]</td>
<td>HC (12) AD (26)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Jervis et al., [28, 108–110]</td>
<td>HC (9) AD (9)</td>
<td>AD &gt; HC</td>
<td>HC = AD</td>
</tr>
<tr>
<td>Juckel et al., [111]</td>
<td>HC (18) AD (18)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Kraiwun et al., [112]</td>
<td>HC (100) AD (25)</td>
<td>AD &gt; HC (but not statistically significant)</td>
<td>NA</td>
</tr>
<tr>
<td>Lai et al., [53]</td>
<td>HC (16) AD (16)</td>
<td>AD &gt; HC</td>
<td>HC = AD</td>
</tr>
<tr>
<td>Lee et al., [64]</td>
<td>HC (31) AD (31)</td>
<td>AD = HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Marsh et al., [113]</td>
<td>HC (17) AD (18)</td>
<td>AD &gt; HC</td>
<td>NA</td>
</tr>
<tr>
<td>O’Mahony et al., [114]</td>
<td>HC (20) AD (18)</td>
<td>AD &gt; HC</td>
<td>NA</td>
</tr>
<tr>
<td>Papadaniil et al., [115]</td>
<td>HC (21) MCI (21) AD (21)</td>
<td>AD &gt; MCI &gt; HC</td>
<td>HC &gt; MCI &gt; AD (but not statistically significant)</td>
</tr>
<tr>
<td>Papaliagkas et al., [116]</td>
<td>HC (30) AD (5)</td>
<td>AD, MCI &gt; HC</td>
<td>HC = MCI, AD</td>
</tr>
<tr>
<td>Papaliagkas, et al., [117]</td>
<td>MCI (15) AD (5)</td>
<td>AD &gt; MCI</td>
<td>MCI &gt; AD</td>
</tr>
<tr>
<td>Pedroso et al., [118]</td>
<td>HC (30) Mild AD (24)</td>
<td>AD &gt; HC</td>
<td>NA</td>
</tr>
<tr>
<td>Pedroso et al., [119]</td>
<td>HC (37) AD (48)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Pokryszko-Dragan et al., [120]</td>
<td>HC (13) AD (13)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Polich et al., [121]</td>
<td>HC (16) AD (16)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Polich and Corey-Bloom, [30]</td>
<td>HC (NA) AD (NA)</td>
<td>AD &gt; HC</td>
<td>HC &gt; AD</td>
</tr>
<tr>
<td>Tachibana et al., [122]</td>
<td>AD (NA)</td>
<td>AD &gt; HC</td>
<td>NA</td>
</tr>
</tbody>
</table>
often include three midline electrode sites—frontal, central, and parietal—with the remaining electrodes used as a reference, monitoring of the electrooculogram for eye movement and blinks, or for better definition of scalp voltage and current maps. An earlobe is often used for a reference electrode, and linked mastoids or noncephalic references can also be used effectively. ERP P300 can measure brain responses to any type of stimulus (i.e., auditory, visual, and somatosensory). P300 latency increases with the difficulty in distinguishing between stimuli. The oddball task (described further in Section 4) is the most common paradigm used for measuring the P300; however, other tasks have also been used, including the continuous performance task, the Eriksen flanker task, the Stroop task, [23], and the sustained attention to response task, or go/no-go task [26].

When comparing results of the ERP P300 across the literature, certain methodological factors should be considered, including test-retest variability. It is possible to reduce test-retest variability by increasing the signal-to-noise ratio and by recording a sufficient prestimulus baseline. To increase the signal-to-noise ratio, ERP P300 recordings are generally repeated over several (35–60) experimental trials, and an averaged representative waveform with several components is produced [27]. Recommendations provided in Duncan et al. [23] include utilizing a baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

Methods to reduce recording noise in the analysis phase include subtracting the baseline from the poststimulus signal. Another method used to reduce noise is to average the experimental trials for one subject to produce a representative waveform for that subject, excluding trials with incorrect responses. A further step often employed entails averaging the waveform for all subjects within a group to produce a grand average that can be used for comparison. Most studies use grand averaged waveforms for data presentation. However, the grand average does not capture signal-to-noise ratio and by recording a sufficient prestimulus baseline of 100–150 ms prior to a stimulus and 800–1000 ms after a stimulus in each epoch [23]. Recording of a sufficient baseline allows distinguishing between normal fluctuations within a subject’s response that can be later subtracted from the poststimulus response during analysis [27]. Using varying interstimulus intervals helps reduce stimulus predictability and some anticipatory brain potentials (e.g., contingent negative variation and lateralized readiness potential). Additionally, using the proper band-pass can filter out frequencies to reduce noise without affecting signal; however, it is recommended that filtering take place during analysis rather than during recording [23, 27].

### 4. ERP P300 Auditory Oddball Paradigm

The oddball paradigm is a discrimination task that requires working memory to decipher between a standard stimulus and an infrequent stimulus. This task is the most used technique to evaluate changes in cognition using the ERP P300 measurement, and it can be performed with visual or auditory stimuli. Several considerations must be taken in the design of the oddball task (i.e., stimulus relevance, probability, distractibility, and focused attention) as they can affect cognitive response. For example, an individual’s arousal state can be impacted by the stimulus presented and the subsequent response [29]. The difficulty of the task can also affect the results: for example, an easier task may not discriminate between groups, while a more difficult task may be able to detect differences between the same subject groups [30].

An auditory oddball task paradigm consists of the presentation of audio tones, including frequent standard tones (e.g., 85% low-pitch [500-Hz]) and rare oddball (e.g., 15% high-pitch [2000-Hz]) tones with relatively short (e.g., 1.2 s to 1.9 s) interstimulus intervals. Over the course of the task, participants are required to count the number of oddball tones they hear; therefore, it is important to conduct prior testing to demonstrate adequate hearing capabilities for the frequencies utilized. The P300 deflection is evoked within a typical range of 250–500 ms after the presentation of the oddball tone, with a positive voltage peak occurring normally at approximately 300 ms after the end of the odd tone [3, 29]. A similar peak is not seen when the standard tone is presented as this tone is not being actively stored in working memory; in contrast, both types of tones elicit N1, P2, and N2 activity, which are associated with sensory and automatic attentional processes.

Counting of the oddball tones may be recorded mentally, or participants may be asked to respond to each target tone by pressing a button. While the use of the button-press allows for the recording of both accuracy and reaction time to correlate with P300 latency and may help maintain engagement with the task, evidence suggests that its use may introduce movement-related artifacts (and lateralized readiness potentials) and may reduce the amplitude of the P300 [31]. Mental recording of the response allows participants to focus on the accuracy of their response rather than on responding as quickly as possible. However, while reaction times have been shown to increase in patients with AD as compared to healthy age-matched controls, accuracy

<table>
<thead>
<tr>
<th>Study</th>
<th>Population (n)</th>
<th>P300 latency</th>
<th>P300 amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AD (15)</td>
<td>HC (15)</td>
</tr>
<tr>
<td>Yamaguchi et al., [123]</td>
<td>HC (16)</td>
<td>AD (16)</td>
<td>AD &gt; HC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AD (16)</td>
<td>HC (16)</td>
</tr>
</tbody>
</table>

Abbreviations: AD: Alzheimer’s disease; ApoE4: apolipoprotein E4; HC: healthy control; MCI: mild cognitive impairment; and NA: not available.
did not differ between the groups [32]. Additional evidence suggests that mental count may require additional working memory capacity; therefore, altering the P300 waveform [33] and the frontal P300 produced may also be an indicator of frontal-executive abilities [34].

5. Neuroanatomical Substrates of the P300

Human lesion studies and intracranial studies demonstrate temporoparietal, frontal, limbic, and paralimbic P300 generators [35]. Convergent neuroimaging evidence using functional magnetic resonance imaging (fMRI) indicates that the neural activation in brain regions surrounding the temporal-parietal junction and lateral prefrontal cortex acts as core generators of P300 [35, 36]. In addition, modality-specific activations in regions such as the superior temporal gyrus (auditory) and the occipital regions (visual) are also involved in a task-specific manner [35]. Neuroanatomical substrates of P300, as shown previously within, comprise prominent nodes of major functional brain networks such as the frontoparietal attentional control network, saliency network, and default mode network [37]. Disruption of these networks impairs cognitive function, including attention, working memory, and episodic memory, and has important impact on cognition in neurological and psychiatric disorders such as AD, depression, and schizophrenia [37]. Specific to age-related neurodegeneration, brain areas of convergent age- and AD-related atrophy are present in the parietal angular gyrus and dorsolateral prefrontal cortex [38], which are core neural generators of P300.

6. ERP in Dementia

Numerous studies have used the ERP to study cognition in various neurodegenerative, neuropsychiatric, and neurodevelopmental disorders. Several studies have detected changes in the early ERP waveform components associated with sensory processes in patients with AD. Longer latencies of several components of the ERP have been reported in patients with AD, as well as family members carrying genetic mutations related to AD [39]. Hirata et al. [40] reported a decreased global field power of the N100 in an oddball paradigm, while Tarkka et al. [41] showed decreased N100 peak amplitude and latency in patients with familial AD compared with healthy controls when performing a habituation task [40, 41]. Longer latencies of both P200 and N200 in the two-back task in patients with AD have also been reported [42]. Additional studies have evaluated late ERP waveforms, such as the LPC/P600 and N400 measurements to evaluate cognition in patients with AD. For example, studies of patients with mild cognitive impairment (MCI) that converted to AD showed reduced or absent LPC/P600 word repetition or N400 effects prior to conversion [18, 43].

6.1. ERP P300 in AD. While several ERP components are altered in patients with AD, the ERP P300 latency is a particularly useful tool for measuring synaptic function [6, 44–46]; the usefulness of ERP P300 in AD stems from the fact that it is only elicited when working memory is active, the simplicity of the task employed in order to evoke the response, and the size of the response in comparison to other waveforms of the ERP. Furthermore, the ERP P300 has been used for several decades to detect cognitive changes in dementia [44]. Therefore, a large body of literature is available for standardization and comparison. In diseases in which there are deficits in working memory, P300 amplitude would be expected to decrease and P300 latency would be expected to increase with disease progression (Figure 3).

ERP P300 can be used for the early assessment of cognitive decline in patients with AD. Indeed, numerous studies have shown abnormalities/differences in P300 amplitude and latency in patients with both MCI and AD [47–49]. Furthermore, P300 studies have been able to sensitively track progression of MCI and AD dementia over time [50–53]. Several studies have also shown that ERP P300 can detect differences in P300 latency in people with a family history of AD as compared to age-matched controls [39, 54–56]. Additionally, results of several meta-analyses of ERP P300 latency studies in patients with AD and MCI support the use of ERP P300 as a biological marker for prodromal AD [57–61]. While most studies have reported a decreased P300 amplitude in patients with AD [62], there have been some exceptions, in which amplitude in patients with AD was reported as being equivalent to that in healthy controls [53, 63]. In contrast, increased P300 latency has been consistently supported in the AD literature (Table 1). Although at least one study found no difference in P300 latency between patients with AD and healthy controls [64], notably, no studies have reported decreased latency in patients with AD compared with healthy controls.

ERP P300 latency is sensitive to drug effects on cognitive performance, highlighting its utility as a biomarker in early clinical trials. As early as 1987, ERP P300 was used to assess the efficacy of treatment in patients with AD; in a double-blind crossover study, an oral muscarinic agonist improved ERP P300 amplitude compared with placebo [65]. Nicergoline (an ergoline derivative) administration resulted in reduced P300 latency in patients with dementia [66, 67]. Treatment with donepezil and rivastigmine (acetylcholinesterase inhibitors) also resulted in reduced P300 latency and improved cognitive scores in patients with AD [68–72]. A recent randomized trial showed an increase in amplitude of the ERP P300 in patients with cognitive impairment after treatment with HTL0009936 (a selective muscarinic M1-acetylcholine receptor agonist) compared to treatment with placebo [73]. Most recently, a double-blind, placebo-controlled phase 1 clinical trial showed a reduction in P300 latency in patients with AD treated with fosgonimeton (a hepatocyte growth factor (HGF)/MET-positive modulator) compared with placebo [74]. Additionally, an ongoing randomized, double-blind, placebo-controlled phase 1 trial (ACT-AD, NCT04491006) in patients with mild to moderate AD is using the change in P300 latency as the primary endpoint in assessing the effects of fosgonimeton. Another ongoing randomized, double-blind, placebo-controlled phase 1 trial (NCT04759365) is measuring the change in P300 latency to assess ASNS1 (an oral, small molecule inhibitor of (protein) 3-O-(N-acetyl-D-glucosaminyll)-L-serine/threonine N-
acetylglucosaminyl hydrolase) in healthy subjects and patients with AD.

Beyond drugs that specifically target AD, the antihypertensive valsartan improved cognitive scores and reduced the P300 latency [75]. Administration of modafinil (a wakefulness agent) in subjects with narcolepsy or idiopathic hypersomnia improved P300 latency and cognition compared with placebo [76, 77]. In addition to the evidence for restoration of P300 latency by therapeutics, the contrary has also been demonstrated with the muscarinic receptor antagonist scopolamine. Scopolamine has been used to model cognitive deficits and has been shown to increase visual P300 latency while reducing cognitive performance [78, 79].

6.2. ERP P300 in Other Neurological Disorders and Neuropsychiatric Disorders. Additional utility of ERP P300 has been noted in the studies of neurodegenerative disorders other than AD. For example, Parkinson’s disease dementia (PDD) and dementia with Lewy bodies (DLB) are associated with increased P300 latency. In one study, patients with DLB showed more severe delayed P300 latency on an auditory oddball paradigm than did patients with AD [80]. Both auditory [81–83] and visual [84] paradigms have been used to evaluate cognition in patients with PDD and have demonstrated increased P300 latency [81–83] and reduced amplitudes [83, 84] as compared to age-matched healthy controls. Furthermore, ERP P300 may also be a useful tool in diagnosing MCI in patients with PD [85, 86]. While several studies have utilized ERP P300 to study PDD and DLB, we are unaware of any clinical trials that have used ERP P300 to detect drug-induced changes in cognition for PDD or DLB to date. An ongoing randomized, double-blind, placebo-controlled, parallel-group, 26-week phase 2 trial (SHAPE, NCT04831281) that will evaluate changes in cognition and ERP P300 latency in patients with mild to moderate PDD or DLB treated with fosgonimeton or placebo is the first controlled clinical trial that includes this outcome measure.

A wide range of studies have utilized the P300 in patients with multiple sclerosis. These include studies on cognitive impairment and fatigue [87], deep gray matter atrophy association with cognition [88], cognition during relapse [89], and the ability of ERP P300 to predict long-term disability [90]. Also, the use of ERP P300 has been suggested as a complementary tool to correlate with cognitive scores in patients with epilepsy [91]. A recent meta-analysis showed that patients with epilepsy have longer P300 latency and lower P300 amplitude than controls [92]. Meador and colleagues’ study had a large influence on clinical practice when it demonstrated prolonging of P300 latency with antiepileptic drugs [93]. Indeed, the prolonged P300 latency has also been demonstrated in other indications with working memory deficits, including HD [26], transient ischemic stroke [94], intellectual disability [25], attention deficit hyperactivity disorder [95], and sleep deprivation [96, 97]. Additionally, evaluation of ERP P300 in patients with TBI demonstrated increased P300 latency and reduced amplitude compared with healthy controls [24, 98]. Furthermore, evidence from a study of military service members supported P300 latency change as a quantitative biomarker for monitoring mental health changes following trauma, including posttraumatic stress disorder, depression, and psychosocial functioning [99].

While other components of the ERP are often used to assess various aspects of neuropsychiatric disorders, reviewed by Sur and Sinha [15], the ability of the P300 component to predict drug effects on cognition is also relevant to drug development for mood disorders regularly associated with cognitive impairment, such as major depressive disorder (MDD) and schizophrenia. Prolonged P300 latency and cognitive deficits have been demonstrated in patients with MDD [100], and P300 latency is directly proportional to MDD severity [101, 102]; these findings support the use of P300 latency as a prognostic indicator for and potential measure of therapeutic response to antidepressant treatment. A recent study of the N-methyl-D-aspartate (NMDA) receptor antagonist ketamine used healthy subjects and ERP P300 to evaluate its potential utility for schizophrenia research [103]. A systematic review recently reported that the increased latency and reduced amplitude of the P300 are common findings in the early stages of schizophrenia, which lends support for the concept that synaptic dysfunction precedes the onset of severe symptoms, akin to the neuropathology of neurodegenerative diseases [104].

7. Conclusion

ERP P300 assessments can directly measure large neuronal network functioning in distinct settings (e.g., in response to auditory stimuli) and can answer a spectrum of important questions in early procognitive drug development. ERP P300 assessment offers several advantages over the current practice of measuring certain brain protein concentrations as a potentially predictive biomarker for cognitive outcomes [11]. The ability to extrapolate results of ERP P300 assessment early on in drug development lends support for rational target dose range decisions and increases confidence in staging larger and longer controlled clinical trials. The utility of ERP P300 latency in studying the progression of cognitive decline in patients with AD has been consistently supported in the literature. Additionally, ERP P300 latency has potential uses in many other neurological and psychiatric disorders. Importantly, P300 latency is an ideal measure that is sensitive to change in cognitive processing occurring in both, disease progression in AD, and in response to drug treatments over short durations. Overall, the evidence reviewed here supports the use of ERP P300 latency as an objective and noninvasive surrogate biomarker for predicting the therapeutic potential of drugs in clinical development for neurocognitive disorders.

Abbreviations

AD: Alzheimer’s disease
ApoE4: Apolipoprotein E4
CNS: Central nervous system
DLB: Dementia with Lewy bodies
EEG: Electroencephalography
EPs: Evoked potentials
ERPs: Event-related potentials
fMRI: Functional magnetic resonance imaging
HD: Huntington’s disease
HGF: Hepatocyte growth factor
MCI: Mild cognitive impairment
MDD: Major depressive disorder
NMDA: N-methyl-D-aspartate
PD: Parkinson’s disease
PDD: Parkinson’s disease dementia
qEEG: Quantitative electroencephalography
REM: Rapid eye movement
TBI: Traumatic brain injury.

Conflicts of Interest
K.C. and H.M. are employees and stockholders of Athira Pharma, Inc. J.O. has worked as a paid consultant for Athira Pharma, Inc., Biogen, and H. Lundbeck A/S. He also serves on the Scientific Advisory Board of Athira Pharma, Inc. and previously on the Scientific Advisory Board of Neurone-trix Solutions, LLC. J.X. has no conflicts of interest to disclose.

Authors’ Contributions
All authors were involved in the conceptualization, reviewing, and editing of all drafts and approval of the final version of this manuscript. John Olichney and Jiangyi Xia contributed equally to this manuscript (co-first authors).

Acknowledgments
Medical writing and editorial support, conducted in accordance with the Good Publication Practice 3 (GPP3) and the International Committee of Medical Journal Editors (ICMJE) guidelines, were provided by Oxford PharmaGenesis Inc., Newtown, PA, USA, and funded by Athira Pharma, Inc., Bothell, WA, USA. J.O. and J.X. were supported by NIH, grant 5R01AG048252.

References
[20] J. Xia, A. Mazaikeri, K. Seguert et al., ”Event-related potential and EEG oscillatory predictors of verbal memory in mild
cognitive impairment,” *Brain Communications*, vol. 2, no. 2, article fca213, 2020.


