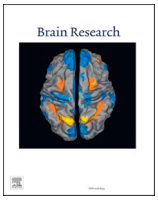
Brain Research 1832 (2024) 148827

Contents lists available at ScienceDirect

Brain Research

journal homepage: www.elsevier.com/locate/brainres



Review

The use of event-related potentials in the investigation of cognitive performance in people with Multiple Sclerosis: Systematic review

Robin Vlieger [a,*], Duncan Austin [b], Deborah Apthorp [c,d], Elena Daskalaki [d], Artem Lensky [e,h], Dianne Walton-Sonda [f], Hanna Suominen [a,d,g], Christian J. Lueck [a]

[a] School of Medicine and Psychology, College of Health and Medicine, Australian National University, Canberra, ACT, Australia
[b] Central Clinical School, Monash University, Melbourne, Victoria, Australia
[c] School of Psychology, Faculty of Medicine and Health, University of New England, Armidale, NSW, Australia
[d] School of Computing, College of Engineering, Computing & Cybernetics, Australian National University, Canberra, ACT, Australia
[e] School of Engineering and Technology, University of New South Wales, Canberra, ACT, Australia
[f] The ACT Health Library and Multimedia Service, The Canberra Hospital, Canberra, ACT, Australia
[g] Department of Computing, University of Turku, Turku, Finland
[h] School of Biomedical Engineering, Faculty of Engineering, University of Sydney, Sydney, NSW, Australia

ARTICLE INFO

Keywords:
Clinical markers
Cognitive function
Data science
Electroencephalography
Event-related potentials
Multiple sclerosis
Neuropsychological tests
Systematic review

ABSTRACT

A biomarker of cognition in Multiple Sclerosis (MS) that is independent from the response of people with MS (PwMS) to test questions would provide a more holistic assessment of cognitive decline. One suggested method involves event-related potentials (ERPs). This systematic review tried to answer five questions about the use of ERPs in distinguishing PwMS from controls: which stimulus modality, which experimental paradigm, which electrodes, and which ERP components are most discriminatory, and whether amplitude or latency is a better measure. Our results show larger pooled effect sizes for visual stimuli than auditory stimuli, and larger pooled effect sizes for latency measurements than amplitude measurements. We observed great heterogeneity in methods and suggest that future research would benefit from more uniformity in methods and that results should be reported for the individual subtypes of PwMS. With more standardised methods, ERPs have the potential to be developed into a clinical tool in MS.

1. Introduction

At present, disease severity in Multiple Sclerosis (MS) is most often measured clinically using the Expanded Disability Status Scale (EDSS; Kurtzke, 1983). However, this tool is recognised to lack precision and is dominated by assessment of motor activity, particularly at its more severe end (Meyer-Moock et al., 2014). The focus on motor activity means that the EDSS does not necessarily provide a holistic assessment of disability. For example, cognitive deficits are reported in 43 %-70 % of People with MS (PwMS) (Chiaravalloti and DeLuca, 2008). These cognitive deficits can often be detected in people with very mild disease, for example, a radiologically isolated syndrome (Lebrun et al., 2010) or a clinically isolated syndrome (Anhoque et al., 2012).

Several tests of cognitive function are already in clinical use, including the Symbol Digit Modalities Test (SDMT; Smith, 1973) and the Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977). These tests are not without criticism: for example, the SDMT, while being a sensitive test, measures cognition in a non-specific manner (Sandry et al., 2021; Berrigan et al., 2022), and the PASAT has been observed to suffer from learning effects (Tombaugh, 2006; Nagels et al., 2008). In an effort to provide a more complete assessment of cognition, test batteries consisting of tests assessing individual cognitive domains have been designed. Examples of such test batteries are the Minimal Assessment of Cognitive Function in MS (MACFIMS; Benedict et al., 2006), TRACK-MS (Taranu et al., 2022), and the Brief Repeatable Battery of Neuropsychological Tests (BRBNT; Boringa et al., 2001). While providing a more complete overview of cognition, the results of these tests are still dependent on responses made by the participants. ERPs, while often requiring a participant to generate a response, provide a direct measurement of brain activity itself, thereby providing an additional dimension to the quantification of cognitive functioning and offering a potential advantage over more conventional neuropsychological testing.

* Corresponding author at: The Australian National University, 145 Science Road, Acton, ACT 2601, Australia.
E-mail address: robin.vlieger@anu.edu.au (R. Vlieger).

https://doi.org/10.1016/j.brainres.2024.148827
Received 29 September 2023; Received in revised form 20 February 2024; Accepted 22 February 2024
Available online 24 February 2024

Accordingly, it has been suggested that the event-related potential (ERP), an electrophysiological brain response time-locked to a stimulus, might provide a more objective measure in this way (Newton et al., 1989; Vazquez-Marrufo, 2017). ERPs have been studied in several different neurodegenerative diseases, and many abnormalities have been described (Tanaka et al., 2000; Turner et al., 2015; Covey et al., 2017). The relative ease with which ERPs can be obtained (measurements can be made from just a small number of electrodes), together with research providing evidence that ERPs can detect compensatory mechanisms before cognitive impairment becomes apparent, indicate that ERPs might act as a clinically useful measure of disease severity and/or prognosis in PwMS. Specifically, combining ERPs with individual tests such as the SDMT and PASAT or the aforementioned test batteries could lead to a more holistic and objective assessment of cognition in MS and, very possibly, provide evidence of abnormality before it becomes detectable by the existing tests.

ERPs are measured using scalp electrodes that are most often embedded in an electrode cap. The stimuli used can be cognitive, visual, auditory, or tactile in nature, and the resulting ERPs are fairly stereotyped, with several components that are labelled according to their polarity (i.e., whether the component has a negative or positive peak) and their approximate latency. For example, P300 refers to a positive potential occurring at around 300 ms after a stimulus (Luck and Kappenman, 2011). This differs from the analysis of electroencephalographic data in the frequency domain (Cohen, 2014), which has been used both in clinical settings as well as to investigate cognition. In such scenarios characteristics of the signal such as power in particular frequency bands and connectivity are analysed (Leocani et al., 2010; Sjøgård et al., 2021; Jamoussi et al., 2023), rather than looking at specific responses in the time domain. Evoked potential (EP) studies in clinical use, on the other hand, do analyse data in the time domain but these tests, although sometimes used to provide information on broader cognitive functioning (Hansch et al., 1982; O'Donnell et al., 1987), generally focus on sensory pathway functioning rather than on specific cognitive processes. The components measured during ERP experiments can be related to specific processes, such as focusing attention on a stimulus or making a decision about the stimulus, enabling the study of different cognitive domains.

The current gold standard method to assess structural cortical changes in MS is magnetic resonance imaging (MRI), with functional MRI (fMRI) used as a method to measure cortical activity during tasks (Filippi et al., 2010; Filippi and Rocca, 2013; Benedict et al., 2020). The use of (f)MRI has two major drawbacks. First, the high cost of MRI equipment, often running in the millions of dollars, and the necessity for MRI scanners to be housed in purpose-built facilities, make that it is generally only available in areas with a sufficiently large population. EEG equipment, on the other hand, is available for a fraction of the cost of MRI scanners and can be set up in almost any room. Second, fMRI is based on the haemodynamic response to stimuli, which is slower than the underlying neural processes in response to the stimuli. This can be partially alleviated by the experimental designs, but does not reach the temporal specificity of ERP measurements (Glover, 2011).

Most research into degenerative diseases has looked at differences in the latency and amplitude of ERPs. A good example involves Huntington's Disease (HD) where differences in ERPs have been found between people with clinical HD, people who are asymptomatic HD gene carriers, and age-matched controls (Turner et al., 2015). ERPs have also been studied in PwMS. Collectively, these studies have used many different types of stimuli, required different types of responses, and measured both amplitudes and latencies of ERPs in many different ways. This means that it is difficult to determine from individual studies how ERPs might best be used to study MS, and to determine the potential role of ERPs in the diagnosis, assessment, and management of MS.

Accordingly, this systematic review set out to answer five questions related to the use of ERPs in distinguishing PwMS from controls, namely: (1) which stimulus modality (i.e., visual or auditory) is most discriminatory, (2) which experimental paradigm produces the most discriminatory results, (3) which individual electrodes generate the best results, (4) which electrophysiological component(s) of the ERP are most discriminatory and, (5) whether amplitude or latency is a better measure.

2. Methods

This review was carried out according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines for systematic reviews (Page et al., 2021), and was registered under the international prospective register of systematic reviews (PROSPERO) number CRD42020166633. Studies were retrieved from Medline, Embase, and Cumulative Index of Nursing and Allied Health Literature (CINAHL) databases in May 2023 using the search strategy "Multiple Sclerosis" AND ("contingent negative variation" OR "event-related potential" OR "readiness potential" OR "bereitschaftspotential" OR "bereitschafts potential" OR "event-related brain potential" OR "event-related auditory evoked potential"). Retrieved studies were screened for duplicates and uploaded to Covidence (Babineau, 2014). After title and abstract screening, a full-text review was carried out by two reviewers with relevant and complementary experience; any conflicts were resolved through mutual discussion or adjudication by a third reviewer. Validity of the studies was assessed using the Risk of Bias Assessment tool for Non-randomized Studies (RoBANS).

Studies were included if they met the following criteria:

Participants were 18 years of age or older.
A comparison was made between a control group and PwMS, or between different subgroups of PwMS.
ERPs were measured during a cognitive task.
Specific values for latencies and amplitudes were reported or could be reasonably accurately derived from published figures.

Studies into EPs, meaning those studies involving detection of an electroencephalographic response to a stimulus without requiring any cognitive processing of the stimulus, were excluded. However, because the terms ERPs and EPs have sometimes been used interchangeably, studies looking at 'cognitive EPs', meaning EPs that were measured during a cognitive task, were included. Other studies that were excluded were studies that were not original research, such as review articles, studies that did not use ERPs as the comparator, such as studies into the effects of drug interventions, studies in which ERP processing or the study design could not be extracted from the text, and studies that provided insufficient data for further analysis, such as conference abstracts. Studies that did not measure ERP component and did not report the temporal occurrence of responses were excluded. If a study into rehabilitation or drug intervention included baseline data, these baseline data were included. No limitations were placed on date of publication or language of publication.

For each study, the following information was extracted: demographic information on the PwMS and controls, the cognitive paradigm used, the neuropsychological tests that were administered, the number of electrodes, the reference, and the sample frequency used, and the latency and amplitude measures for each component of the ERP. In the PwMS group, the EDSS score and MS subtypes relapsing-remitting MS (RRMS), primary progressive MS (PPMS), and secondary progressive MS (SPMS) were recorded.

Pooled effect sizes and confidence intervals were calculated for each component and displayed as forest plots using RevMan 5.4 (The Cochrane Collaboration, 2020). Effect sizes quantify the magnitude of the difference between two groups. They have been used in medical research to describe treatment effects, for example, and it is argued they provide a better assessment of these effects than p-values alone (McGough and Faraone, 2009). Hedges' g was used to assess effect sizes: an effect size of 0.2 or smaller was considered small, an effect size of

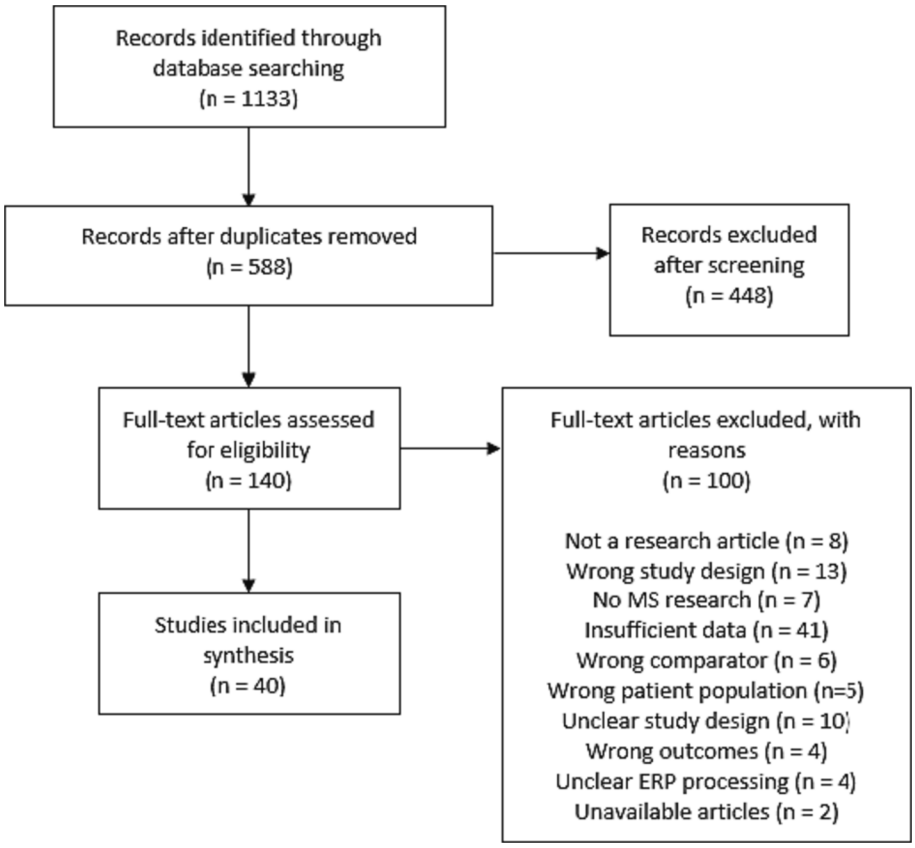


Fig. 1. PRISMA diagram from the 1133 identified to 40 included studies.

Table 1
Breakdown of participants in the 62 experiments. Ranges for all values across studies are given in brackets.

	CIS	RRMS	PPMS	SPMS	BMS	PwMS	controls
average number of participants	20.5 [7–44]	28.6 [11–72]	N/A	16 [-]	9.7 [9–10]	32.2 [9–101]	26.3 [7–89]
average number of female participants	16.3 [6–27]	18.2 [7–51]	N/A	11 [-]	6 [-]	18.6 [6–72]	17.9 [4–51]
average number of male participants	8.7 [4–17]	10.4 [0–32]	N/A	5 [-]	3.7 [3–4]	9.9 [0–32]	10.0 [0–38]
average age (years)	31.6 [27.4–40.4]	37.3 [32.5–45.6]	N/A	43.8 [-]	41.1 [38.7–42.3]	38.9 [27.4–51.0]	37.2 [277–46.5]
average disease duration (years)	3.3 [0.4–4.9]	6.8 [1.3–11.6]	N/A	13.9 [-]	12.2 [12.1–12.4]	8.2 [0.4–17.7]	–
average EDSS score	1.7 [1.4–3.0]	2.3 [0.9–3.1]	N/A	5.8 [-]	1.8 [1.6–1.9]	3.0 [0.87–5.78]	–

CIS: clinically isolated syndrome; EDSS: expanded disability status scale; PwMS: People with MS; RRMS: relapsing remitting MS; PPMS: primary progressive MS; SPMS: secondary progressive MS; BMS: benign MS. Note: average numbers for female and male participants may not correspond to averages for all participants because the number of studies reporting numbers for female and male participants differ.

0.2–0.5 medium, and an effect size of 0.8 or more large (Sullivan and Feinn, 2012). A list of reviewed studies can be found in Supplement 1.

3. Results

In total, 1133 studies were retrieved, of which 588 remained after removal of duplicates. 448 studies were removed based on title and/or abstract, and another 100 were excluded after a full-text review. Data were extracted from the remaining 40 studies (Fig. 1).

In all, the 40 studies reported on a total of 62 experiments. Of these experiments, 33, 8, and 6 reported on people with RRMS, SPMS, and PPMS, respectively, and 4 experiments included people with Clinically Isolated Syndrome (CIS). Eleven studies included PwMS with different subtypes of MS but did not distinguish between the subtypes when analysing the results. Unfortunately, demographic data on people with PPMS were not provided by the specific studies so this information was not available for inclusion in Table 1. Study population sizes ranged from 19 to 172, with an average of 32 PwMS and 26 controls per study. In the studies that reported on the sex of participants, 62.5 % of PwMS were female compared to 63 % of controls. Unfortunately, while most, but not all, studies reported characteristics such as sex, age, disease duration, and EDSS score, characteristics such as ethnicity were absent. Participant demographics are provided in Table 1 and details per experiment can be found in Supplement 2.

Five experimental paradigms, namely the oddball paradigm, Posner paradigm, n-back tasks, choice reaction experiments, and sensorimotor integration paradigm (see Supplement 3 for general examples of these paradigms), were used in more than one experiment. Table 2 provides

Table 2
Overview of the experimental paradigms that were used in more than one experiment.

Experimental paradigm	N experiments		N visual and auditory experiments
Oddball experiments	42		
		visual	8
		auditory	34
Posner paradigms	4		
		visual	4
Choice reaction experiments	3		
		visual	2
		auditory	1
N-back tasks	2		
		visual	2
Sensorimotor integration paradigm	2		
		multimodal	2

an overview of these paradigms, but a full overview of paradigms and experimental settings used per study can be found in Supplement 4. Most studies reported on oddball experiments, in which participants are presented with a string of stimuli of which some differ. The participants have to either count these 'oddballs' or respond to them through, for example, a button press. While the aim of the paradigm is the same, there were large differences in individual paradigms. There was considerable variation in the number of stimuli per experiment, ranging from 100 to 900 for auditory stimuli and 100 to 400 for visual stimuli. In addition, the frequency of the tone used in auditory experiments was also variable, with the oddball tone ranging from 1000 Hz to 2000 Hz and the standard tone from 500 Hz to 1000 Hz. Even though differing in their specifics, most oddball experiments used auditory stimuli and asked participants to count the number of oddballs.

In the Posner paradigm experiments (Posner, 1980), the paradigms were more uniform. A participant was first shown a cue in the form of an arrow, followed by a target or standard stimulus. Participants had to withhold responses to the standard stimulus and press the left or right button depending on the side on which a target stimulus appeared. (Gonzalez-Rosa et al., 2006; Vázquez-Marrufo et al., 2009; Gonzalez-Rosa et al., 2011) For the choice reaction experiments, (Sundgren et al., 2015; Cooray et al., 2020), participants were presented with two different auditory or visual stimuli and asked to press a button that corresponded to the presented stimulus with their left or right index finger. In the n-back tasks, (Covey et al., 2017) participants had to compare the current letter that was being presented to a letter presented n stimuli earlier. They pressed two buttons with their thumbs if they detected a match, and two different buttons if there was no match. The study also included a 0-back experiment where participants had to press a button the moment a stimulus appeared, effectively making it a choice-reaction experiment. In the sensorimotor integration experiments (Uysal et al., 2014), participants were presented with a warning tone, followed by stimulus 2 s later to which they had to respond as quickly as possible by pressing a button with their dominant hand. Supplement 4 provides an overview of experimental parameters per experiment.

The many differences in experimental paradigms required analysis from several different angles. In order to address the five questions listed in the introduction, the results of the relevant studies had to be collapsed across different paradigms, electrodes, modalities, etcetera, even though the data were very heterogeneous.

3.1. Comparison of auditory and visual stimuli

Overall, studies using visual stimuli generated greater effect sizes (Fig. 2) and there was a larger difference for latency than for amplitude. Pooled effect sizes for latency measurements were 0.3 CI: 0.15, 0.46] and 0.68 CI: 0.55, 0.81] for auditory and visual stimuli, respectively, while the same measurements for amplitude were 0.15 CI: 0.05, 0.26] and 0.52 CI: 0.42, 0.62].

3.2. Comparison of experimental paradigms

Oddball experiments made up approximately 63 % of experiments but the Posner paradigm produced the largest pooled effect size for latency (Fig. 3a) and the Attention Network Test (ANT) produced the largest effect size for amplitude (Fig. 3b). Latencies were longer for PwMS than controls except for the working memory tasks and modified Iowa Gambling Tasks (marked with an asterisk).

3.3. Comparison of electrodes

Almost all experiments reported measurements from electrodes Fz, Cz, and Pz. (Fig. 4). Latency measurements were largest for electrode Cz (ES = 0.84 0.62, 1.05]), while the pooled effect size for Fz was noticeably smaller. Pooled effect sizes for all three electrodes were smaller for amplitude measurements than for latency measurements.

3.4. Comparison of ERP components

The precise component of the ERP that was measured in each experiment depended on the paradigm used. As expected from the large number of oddball experiments, the component that was most often measured was P300, followed by N100 and N200 (Fig. 5). The N200-P300 complex was frequently measured as a single entity: this measure refers to the peak-to-peak difference in amplitude between the N200 to P300 components, as these two potentials immediately follow each other. For latency measurements, P300 and N200 produced the largest pooled effect sizes, followed by P150, but this was based on only 1 experiment, resulting in a large confidence interval. In auditory

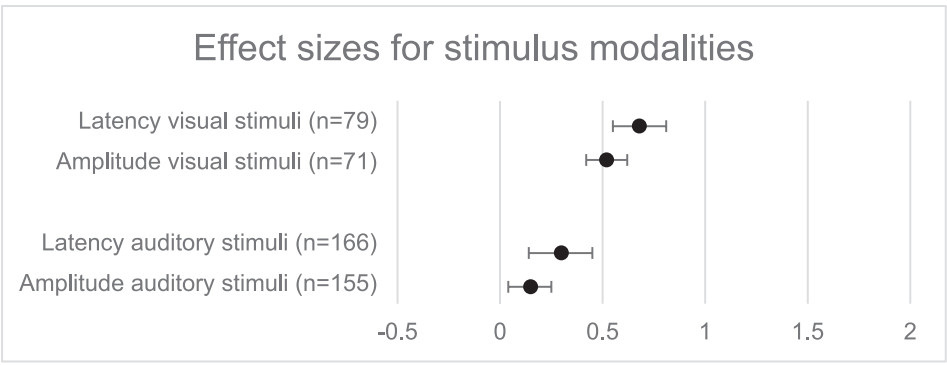
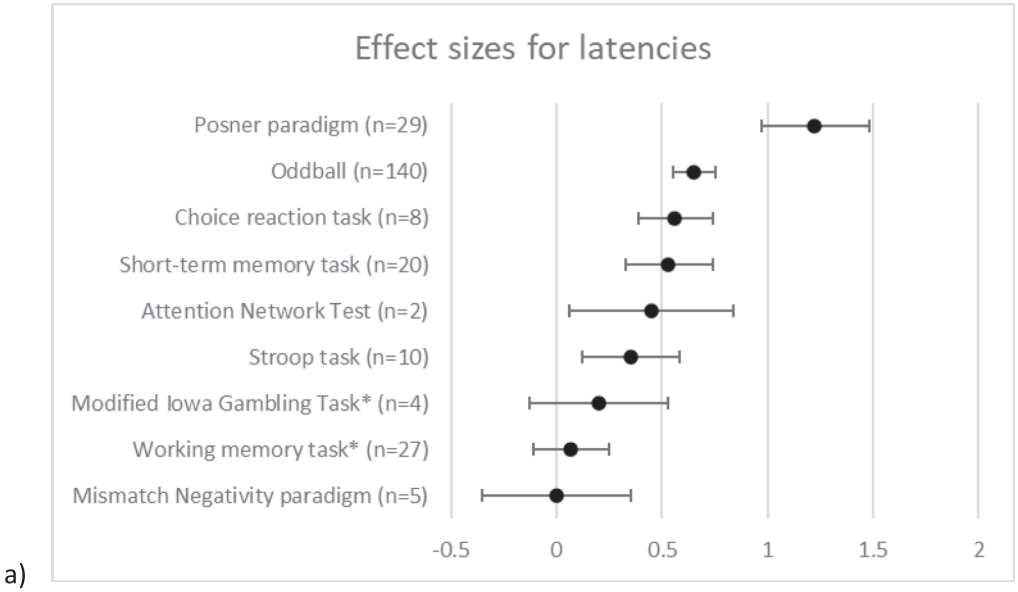
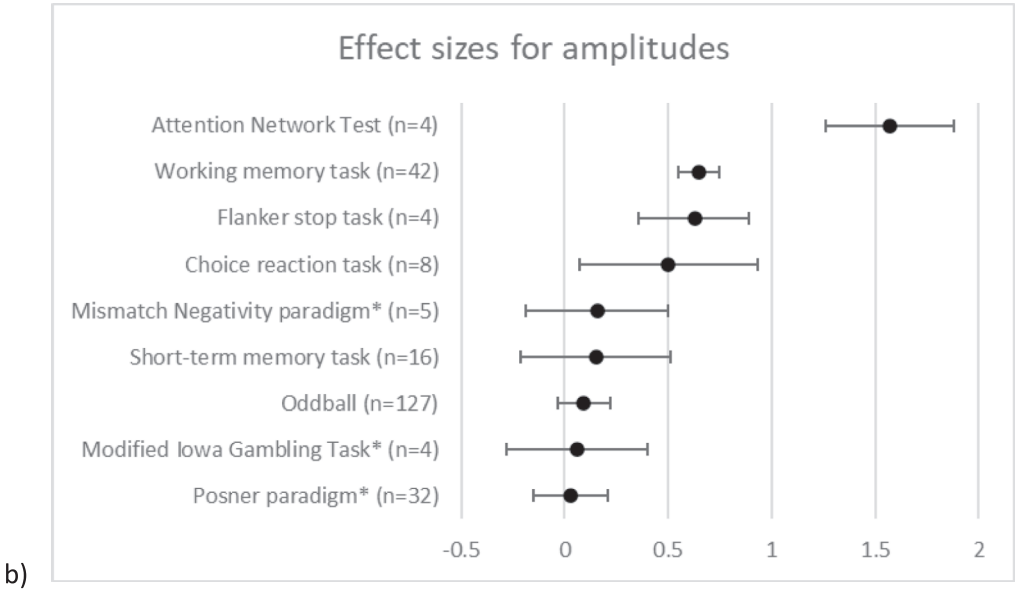


Fig. 2. Pooled effect sizes and 95% confidence intervals comparing visual and auditory stimuli for latency and amplitude. 'n' refers to the total number of measurements made across all experiments, because multiple measurements from different electrodes were often made in any given experiment.



a)



b)

Fig. 3. a and b. Pooled effect sizes and 95% confidence intervals for the experimental paradigms looking at latency (a) and amplitude (b). 'n' refers to the total number of measurements made across all experiments, because multiple measurements from different electrodes were often made in any given experiment. In measurements with an asterisk, measurements for PwMS were larger than those of controls, while all other measurements were larger for the control group.

experiments, the largest pooled effect sizes for amplitude were observed at the contingent negative variation (CNV, g = 0.9 0.46, 1.35]). The Error-Related Negativity (ERN, g = 0.68 0.3, 1.06]), which is the potential that occurs when a participant commits an error in a cognitive task, and P600 (g = 0.63 0.51, 0.75]), generated the second and third largest pooled effect sizes.

3.5. Comparison of latency and amplitude

Effect sizes for individual latency measurements ranged considerably, with a value of 1.51 for shorter latencies in PwMS at one extreme to 5.02 for shorter latencies in controls at the other. In total, there were 34 measurements in which latencies in PwMS were shorter than those in controls compared to 211 measurements in which latencies were shorter in controls. Regarding amplitude measurements, effect sizes ranged from 1.47 higher in PwMS to 22.98 higher amplitudes in controls. In total, there were 83 measurements in which amplitudes in PwMS were higher than those in controls compared to 143 measurements in which amplitudes were higher in controls.

Overall, latency produced larger pooled effect sizes than amplitude when comparing experiments with auditory and visual stimuli. When pooling based on electrode location, latency produced larger pooled effect sizes for all three locations. For the modified Iowa Gambling Tasks (mIGT), the short-term memory tasks, the oddball experiments, and the Posner paradigm experiments, latency produced larger pooled effect sizes, while the pooled effect sizes produced by measuring amplitude were larger for the n-back tasks, working memory tasks, and the ANTs. Interestingly, the largest individual pooled effect size occurred for the amplitude measurements in the ANTs. When calculating pooled effect sizes for components, the latency pooled effect sizes were larger for components N100, N200, P200, and P300, while amplitude pooled effect sizes were larger for P100 and P600.

3.6. Direction of effects

As mentioned above, not all effect sizes were in the same direction.

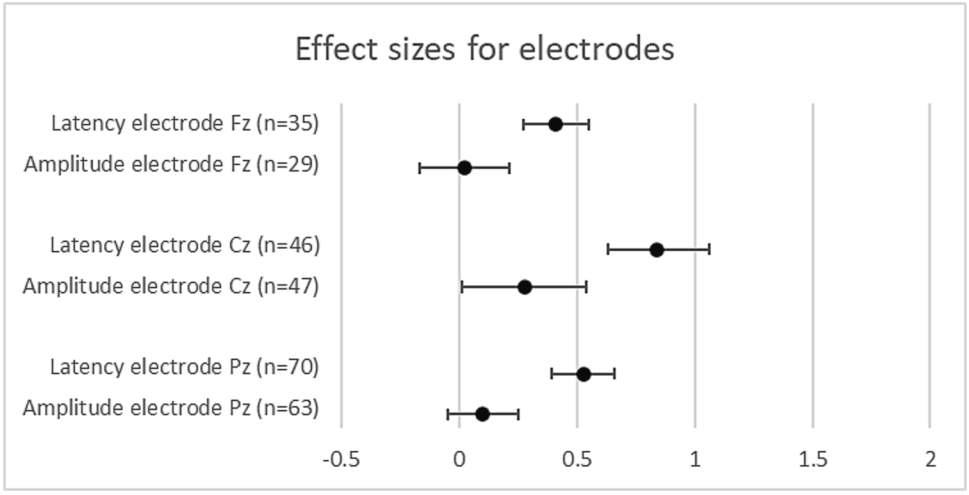


Fig. 4. Pooled effect sizes and 95% confidence intervals for measurements derived from electrodes Fz, Cz, and Fz for latency and amplitude. 'n' refers to the total number of measurements made across all experiments, because multiple measurements from different electrodes were often made in any given experiment.

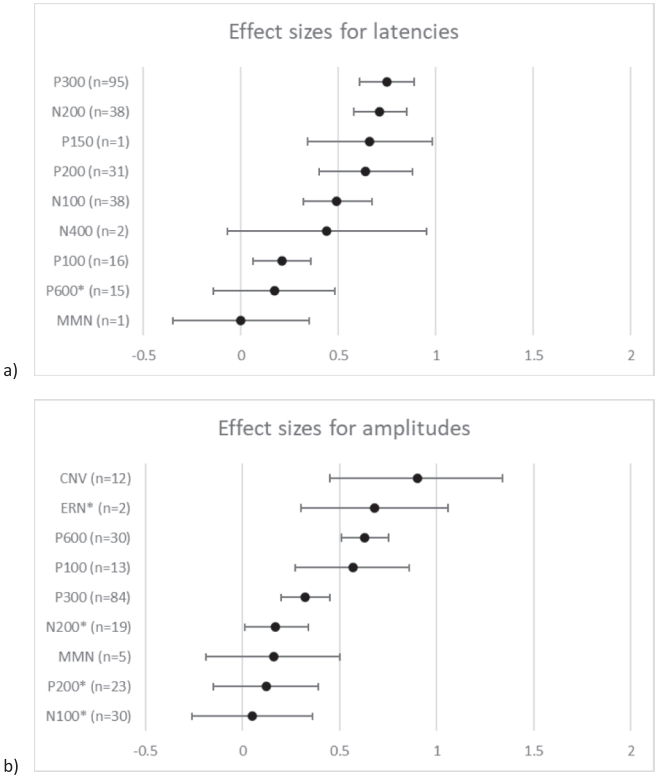


Fig. 5. a and b. Pooled effect sizes and 95% confidence intervals for the different ERP components looking at latency (a) and amplitude (b). 'n' refers to the total number of measurements made across all experiments, because multiple measurements from different electrodes were often made in any given experiment. In measurements with an asterisk, measurements for PwMS were larger than those of controls, while all other measurements were larger for the control group.

For most paradigms, latencies were longer and amplitudes lower for PwMS compared to controls, but this was not the case in experiments using the mIGT and working memory tasks for latencies, and Posner paradigms, the mIGT, and mismatch negativity tasks for amplitudes. Components, locations and stimulus modality leading to these results vary. P300 was measured from electrode Pz at four different moments during a visual mIGT, with latencies being longer for controls than

Table 3
An overview of the versions of the SDMT and PASAT as described by the authors in case such details were provided.

SDMT Reference	Version	PASAT Reference	Version
Artemiadis et al., 2018	Oral SDMT	Covey et al., 2017	2-second PASAT 3-second PASAT
Bissonnette et al., 2023	Oral SDMT	Gerschlager et al., 2000	Easy PASAT Hard PASAT
Chinnadurai et al., 2016	Oral 60 s modified SDMT Oral 180 s modified SDMT	Kocer et al., 2008	
Covey et al., 2017	Rao adaptations	López-Góngora et al., 2015	2-second PASAT 3-second PASAT
López-Góngora et al., 2015	Oral SDMT	Magnié et al., 2007	3-second PASAT
Paolicelli et al., 2021		Paolicelli et al., 2021	3-second PASAT 5-second PASAT
Pokryszko-Dragan et al., 2016a		Pokryszko-Dragan et al., 2016a	3-second PASAT
Pokryszko-Dragan et al., 2016b		Pokryszko-Dragan et al., 2016b	
Vazquez-Marrufo et al., 2008		Uysal et al., 2014	3-second PASAT
Vázquez-Marrufo et al., 2014		Vázquez-Marrufo et al., 2014	3-second PASAT
Vázquez-Marrufo et al., 2019		Vázquez-Marrufo et al., 2019	3-second PASAT
Waliszewska-Prosół et al., 2018			

PwMS for all four, but amplitude only being larger for PwMS than controls in one measurement, and P600 showed longer latencies in an auditory working memory task at 9 of 15 measured locations, mainly over frontal and central electrodes. 3 out of 5 measurements in electrode Fz during a mismatch negativity experiment showed larger amplitudes for PwMS than controls, while the majority of measurements during the Posner paradigm for components N100, N200, and P200 were larger for PwMS than controls in unspecified electrodes. Lastly, two measurements of the ERN during a flanker test also showed larger amplitudes for PwMS

than controls.

3.7. ERP components and neuropsychological tests

Of the forty studies included in this systematic review, 26 also reported on the results of neuropsychological tests. Of the 26 studies, 12 employed the SDMT and 11 employed the PASAT (see Table 3 for details on the versions used) but not all studies reported on the relation between the results of the tests of cognitive function and the ERPs. Supplement 5 provides an overview of the individual tests used in the various studies.

For the SMDT, two studies (Chinnadurai et al., 2016; Artemiadis et al., 2018) found a difference between controls and PwMS, four studies (López-Góngora et al., 2015; Covey et al., 2017; Waliszewska-Prosół et al., 2018; Paolicelli et al., 2021; Bissonnette et al., 2023) found no difference, and five (Vazquez-Marrufo et al., 2008; Vázquez-Marrufo et al., 2014; Pokryszko-Dragan et al., 2016a; Pokryszko-Dragan et al., 2016b; Vázquez-Marrufo et al., 2019) did not report on the comparison between controls and PwMS. Nine of these studies (Vazquez-Marrufo et al., 2008; Vázquez-Marrufo et al., 2014; López-Góngora et al., 2015; Chinnadurai et al., 2016; Pokryszko-Dragan et al., 2016b; Covey et al., 2017; Waliszewska-Prosół et al., 2018; Artemiadis et al., 2018) found differences in their ERP measurements between controls and PwMS. Of these, two (Vázquez-Marrufo et al., 2014; Artemiadis et al., 2018) reported a correlation between ERP measurements and the SDMT results with correlations between worse SDMT performance and CNV amplitude, longer P300 latency, and reduced P300 amplitude. One study (Covey et al., 2017) reported no significant group differences in the SDMT between controls and PwMS, but did find more variability in PwMS, as well as different relationships between ERPs and SDMT score, with P300 latency on a 2-back task being predictive of SDMT performance in controls. Two studies (López-Góngora et al., 2015; Waliszewska-Prosół et al., 2018) found differences in ERP but no corresponding difference in SDMT.

For the PASAT, three studies (Gerschlager et al., 2000; Uysal et al., 2014; Paolicelli et al., 2021) found a difference between controls and PwMS, three studies (Kocer et al., 2008; López-Góngora et al., 2015; Covey et al., 2017) found no difference, and five (Magnié et al., 2007; Vázquez-Marrufo et al., 2014; Pokryszko-Dragan et al., 2016a; Pokryszko-Dragan et al., 2016b; Vázquez-Marrufo et al., 2019) did not report comparisons between controls and PwMS. Eight studies (Magnié et al., 2007; Kocer et al., 2008; Uysal et al., 2014; Vázquez-Marrufo et al., 2014; López-Góngora et al., 2015; Pokryszko-Dragan et al., 2016a; Pokryszko-Dragan et al., 2016b; Covey et al., 2017) reported differences in ERPs between controls and PwMS. Two studies (Magnié et al., 2007; Uysal et al., 2014) found a correlation between PASAT and ERP: one study reported a correlation between PASAT and the P300 amplitude in a visual oddball experiment while the other reported that an increase CNV amplitude correlated with a better PASAT performance. One study (Covey et al., 2017) reported P300 amplitude on the 2-back task in electrode Pz was predictive of performance of PwMS in the 2- and 3-second PASAT. Additionally, reduced P100 amplitude in electrode Pz on the 2-back task was associated with better performance by PwMS on the 2-second PASAT.

4. Discussion

This study reviewed the existing literature to determine which stimulus modality, experimental paradigm, scalp electrode, ERP component, and parameter were most discriminatory between PwMS and controls in cognitive ERP experiments, with a view to trying to standardise future experiments using this technique. Overall, the largest pooled effect sizes were seen for visual stimuli, the Posner and ANT paradigms, and electrode Cz. The ERN and P300 were the best-performing ERP components, and latency was more discriminatory than amplitude.

Collapsing across the various studies, visual stimuli produced larger pooled effect sizes than auditory stimuli, regardless of whether they were expressed in terms of latency or amplitude. This difference was especially noticeable for amplitude measurements where visual stimuli produced an overall pooled effect size of 0.52 while auditory stimuli resulted in a pooled effect size of only 0.15. Interestingly, there were more measurements from experiments using auditory stimuli than from experiments using visual stimuli, mostly because of the large number of auditory oddball experiments. Given the more discriminatory results produced by the experiments using visual stimuli, this seems to be the better modality to use in future experiments.

An important consideration is that some of the difference found between visual and auditory stimuli may relate to fact that visual disturbance, particularly in the form of optic neuritis (Chan, 2002), is common in PwMS. Involvement of the auditory pathways is relatively rare. Indeed, in some of the studies reviewed here, it was observed that the latency of N100, a component associated with attention, was prolonged with visual stimuli (Vázquez-Marrufo et al., 2014). Research elsewhere (Covey et al., 2021) has shown that there may be an impairment of integration of visual information. To clarify this issue, future research will need to look into the correlation between visual function and the ERP.

Analysis of experimental paradigms revealed that oddball paradigms were by far the most frequently used paradigm, particularly as an auditory version. However, this paradigm did not produce the largest pooled effect sizes, either for latency or amplitude. Other paradigms, like the Posner paradigm (measuring latency) or the ANTs (measuring amplitude), produced much larger effect sizes than the oddball paradigm. While this finding could relate to the issue of visual involvement in PwMS discussed above, the lack of available information about participants' visual function made it impossible to comment on this issue here. The ANT results of $g = 1.57$ ($CI = 1.26, 1.88$]) would need further investigation, as these are generally only observed for very noticeable differences between the groups that are being compared (Hilgard 2021). For example, the difference in height between men and women born in Spain in 1980 produces a Hedges' g of 1.87, based on data reported by (Garcia and Quintana-Domeque, 2007).

The vast majority of experiments reported results from the three midline electrodes Fz, Cz, and Pz, with only few reporting results in other electrodes. This observation held true across all experiments regardless of the paradigm used. In this regard, the positioning of a midline electrode does not relate well to specific underlying brain areas and, therefore, it is difficult to draw any 'anatomical' functional conclusions from the results provided by these electrodes. However, placement of midline electrodes is relatively easy, making them potentially more convenient for future clinical use.

Overall, latency measurements produced larger effect sizes for all electrodes than amplitude measurements, Cz produced the largest pooled effect sizes for both latency and amplitude measurements, and Pz produced larger pooled effect sizes than Fz. Based on these results it would be worth exploring central and posterior electrodes, possibly using regions of interest (ROIs) instead of single electrodes or independent components, if high-density electrode setups with 128 or more electrodes are available.

The exact component that was reported by each of the studies appeared to depend on which experimental paradigm was employed. For example, P300 and N200 were most frequently measured during oddball experiments. In fact, the overall effect sizes for latency at both P150 and N400 were larger than those at P300, and P600 performed better when measuring amplitude. These larger effect sizes were, however, based on a relatively small number of measurements, often from different sections of the same experiment. It was therefore not possible to determine conclusively which EEG component was best at distinguishing PwMS from controls.

In general, latency measurements produced the largest effect sizes, though the largest effect size overall was seen for amplitude in the ANTs. However, this latter result was derived from only four individual

Table 4
An overview of recommendations for future studies.

1) Visual stimuli are preferable to auditory stimuli
2) Paradigms other than the oddball paradigm are likely to be more discriminatory
3) Sufficiently large cohorts are required to generate meaningful results
4) Reporting of measured latencies and amplitude in ms and μV, respectively, as such measurements permit easier comparison
5) The precise subtype of MS being studied should be clarified

experiments, compared with 85 for oddball experiments, making comparison difficult. The same was true for many of the other measurements, meaning that further investigation is required before drawing a definitive conclusion.

The direction of effects was not the same in each study and for each component. Unfortunately, the results that produced these differences come from very different studies, hindering comparison. For example, the two studies producing shorter latencies for PwMS than controls measured P300 and P600, respectively, from different electrodes and using visual stimuli in one experiment and auditory stimuli in the other. For the amplitude measurements that were larger for PwMS than controls, the study providing most of the results did not specify which electrodes were used. This heterogeneity in methods and the fact information for comparison is missing means the differences can be noted but not further analysed.

This study has the following two main limitations. First, the literature consisted of reports of many different experimental arrangements in each individual study, meaning that the numbers pertinent to any one group were often very small. The way this was dealt with here was to collapse across paradigms, electrodes, etc. to try to generate an overall 'summary' picture. This means that the heterogeneity for any comparison was very large. It is likely that more meaningful (and clinically useful) results would be obtained if researchers adopted a more standardised methodology in the future. To assist with this, we have provided recommendations for future research in Table 4.

A second limitation of our research is that the results presented here treat PwMS as a single entity. 'PwMS' actually comprises several different clinical subtypes such as RRMS, PPMS and SPMS, and clinically or radiologically isolated syndromes. While some information about subtype was available, the numbers were too small to allow meaningful comparison. With this in mind, however, it is important for future investigators to clarify the precise subgroup of MS being studied.

Another important consideration is that there may have been an effect arising from technical issues such as the type of electrode used, the method of electrode application, or the location of the reference electrode. Granted the large variation in the information provided by the various studies (where this was available) this issue could not be commented on here. However, it is worth pointing out that it would be important to assess this in future studies.

The review presented here synthesised results from research over a period of more than 30 years, from 1992 (Filippi, 1992; Giesser, 1992; Honig, 1992) until April 2023. (Bissonnette et al., 2023) In doing so, this review provides an overview of the most prevalent paradigms and methods, similar to reviews that synthesized results in other neurodegenerative diseases (Seer et al., 2016), and, by doing so and in combination with other reviews, can help inform decision making about experimental setups not only in MS but also other diseases.

More broadly, research into cognitive performance of PwMS, would benefit from a more uniform and shared approach to data collection, analysis, and reporting methods, and results should be reported for the different subtypes of PwMS. In particular, equipping research into ERPs in the investigation of cognitive performance of people with MS with more standardised methods, would facilitate the potential of ERPs to be developed into a clinical tool in MS. Such more uniform and shared approaches have been successful in driving transformational research in other medical informatics contexts (Chapman et al., 2011; Huang and Lu, 2016) and are hence likely to benefit studying clinical markers of MS as well. If this can be accomplished, ERPs may become useful in the management of MS, by providing a more objective measure of cognitive function. This would potentially assist PwMS, who are often uncertain about what the future holds for them (Hossain et al., 2022), by providing greater clarity on their individual rate of disease progression, as well as by providing a useful measure of the effectiveness (or not) of disease-modifying medication (Capone et al., 2020; Gerschlager et al., 2000). At this point, however, considerable further work is required before the ERP can be considered clinically useful.

5. Conclusion

This study has synthesised a substantial body of existing work with a view to assisting future studies to use the most discriminatory experimental setups, i.e. those that distinguish best between PwMS and controls. Our results show visual stimuli are preferable over auditory stimuli, and that paradigms other than the oddball paradigm (which has been most frequently used in the past) are likely to provide better results. To derive more conclusive outcomes, larger cohorts are needed, and any results should be divided into specific MS subtypes and reported in terms of raw numbers. In summary, future work looking at the role of ERPs in PwMS would benefit from a more systematic and uniform approach to collection, analysis, and reporting methods. However, the results show that, with more standardised methods, ERPs have significant potential to contribute to the clinical assessment of PwMS, possibly as an objective biomarker of cognitive decline in the disease. This would provide additional information to that provided by the EDSS, generate a more holistic assessment of patients and, therefore, offer an improved health experience to PwMS.

CRediT authorship contribution statement

Robin Vlieger: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Duncan Austin:** Writing – original draft, Validation, Investigation, Data curation. **Deborah Apthorp:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Conceptualization. **Elena Daskalaki:** Writing – original draft, Supervision, Methodology, Conceptualization. **Artem Lensky:** Writing – original draft, Methodology, Data curation. **Dianne Walton-Sonda:** Resources, Methodology, Conceptualization. **Hanna Suominen:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Formal analysis, Conceptualization. **Christian J. Lueck:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data used in this review was extracted from studies by other researchers. A list of these studies is provided

Acknowledgements

This research was funded by and has been delivered in partnership with Our Health in Our Hands (OHIOH), a strategic initiative of the Australian National University, which aims to transform health care by developing new personalised health technologies and solutions in collaboration with patients, clinicians, and health-care providers. We gratefully acknowledge the funding from the ANU School of Computing for the first author's PhD studies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brainres.2024.148827.

References

Anhoque, C.F., Biccas Neto, L., Domingues, S.C.A., Teixeira, A.L., Domingues, R.B., 2012. Cognitive impairment in patients with clinically isolated syndrome. Dement. Neuropsychol. 6, 266–269.
Artemiadis, A., Anagnostouli, M., Zalonis, I., Chairopoulos, K., Triantafyllou, N., 2018. Structural MRI correlates of cognitive function in multiple sclerosis. Mult. Scler. Relat. Disord. 21, 1–8.
Babineau, J., 2014. Product review: covidence (systematic review software). J. Can. Health Libraries Assoc./Journal De L'association Des Bibliothèques De La Santé Du Canada 35 (2), 68–71.
Benedict, R.H., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., Weinstock-Guttman, B., 2006. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). J. Int. Neuropsychol. Soc. 12 (4), 549–558.
Benedict, R.H., Amato, M.P., DeLuca, J., Geurts, J.J., 2020. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. The Lancet Neurology 19 (10), 860–871.
Berrigan, L.I., LeFevre, J.A., Rees, L.M., Berard, J.A., Francis, A., Freedman, M.S., Walker, L.A., 2022. The symbol digit modalities test and the paced auditory serial addition test involve more than processing speed. Mult. Scler. Relat. Disord. 68, 104229.
Bissonnette, J.N., Pimer, L., Francis, A.M., Hull, K.M., Leckey, J., MacGillivray, M., Fisher, D.J., 2023. An investigation of auditory processing in relapsing–remitting multiple sclerosis. Exp. Brain Res. 241 (5), 1319–1327.
Boringa, J.B., Lazeron, R.H., Reuling, I.E., Ader, H.J., Pfennings, L.E., Lindeboom, J., Polman, C.H., 2001. The brief repeatable battery of neuropsychological tests: normative values allow application in multiple sclerosis clinical practice. Mult. Scler. J. 7 (4), 263–267.
Capone, F., Motolese, F., Falato, E., Rossi, M., Di Lazzaro, V., 2020. The potential role of neurophysiology in the management of multiple sclerosis-related fatigue. Front. Neurol. 11, 251.
Chan, J.W., 2002. Optic neuritis in multiple sclerosis. Ocul. Immunol. Inflamm. 10 (3), 161–186.
Chapman, W.W., et al., 2011. Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. J. Am. Med. Inform. Assoc. 18 (5), 540–543.
Chiaravalloti, N.D., DeLuca, J., 2008. Cognitive impairment in multiple sclerosis. The Lancet Neurology 7 (12), 1139–1151.
Chinnadurai, S.A., Venkatesan, S.A., Shankar, G., Samivel, B., Ranganathan, L.N., 2016. A study of cognitive fatigue in multiple sclerosis with novel clinical and electrophysiological parameters utilizing the event related potential P300. Mult. Scler. Relat. Disord. 10, 1–6.
Cohen, M.X., 2014. Analyzing neural time series data: theory and practice. MIT press.
Cooray, G.K., et al., 2020. Mechanism of visual network dysfunction in relapsing-remitting multiple sclerosis and its relation to cognition. Clin. Neurophysiol. 131 (2), 361–367.
Covey, T.J., Shucard, J.L., Shucard, D.W., 2017. Event-related brain potential indices of cognitive function and brain resource reallocation during working memory in patients with multiple sclerosis. Clin. Neurophysiol. 128 (4), 604–621.
Covey, T.J., Golan, D., Doniger, G.M., Sergott, R., Zarif, M., Srinivasan, J., Gudesblatt, M., 2021. Visual evoked potential latency predicts cognitive function in people with multiple sclerosis. J. Neurol. 268, 4311–4320.
Filippi, M., Rocca, M.A., Benedict, R.H.B., DeLuca, J., Geurts, J.J.G., Rombouts, S.A.R.B., Comi, G., 2010. The contribution of MRI in assessing cognitive impairment in multiple sclerosis. Neurology 75 (23), 2121–2128.
Filippi, M., et al., 1992. Correlations between auditory event-related potentials (ERPs), brain MRI and neuropsychological tests in multiple sclerosis. Rivista di Neurobiologia 38 (5), 369–375.
Filippi, M., Rocca, M.A., 2013. Present and future of fMRI in multiple sclerosis. Expert Rev. Neurother. 13 (sup2), 27–31.
Garcia, J., Quintana-Domeque, C., 2007. The evolution of adult height in Europe: a brief note. Econ. Hum. Biol. 5 (2), 340–349.
Gerschlager, W., Beisteiner, R., Deecke, L., Dirnberger, G., Endl, W., Kollegger, H., Lang, W., 2000. Electrophysiological, neuropsychological and clinical findings in multiple sclerosis patients receiving interferon β-1b: a 1-year follow-up. Eur. Neurol. 44 (4), 205–209.
Giesser, B.S., et al., 1992. Endogenous event-related potentials as indices of dementia in multiple sclerosis patients. Electroencephalography and clinical neurophysiology 82 (5), 320–329.
Glover, G.H., 2011. Overview of functional magnetic resonance imaging. Neurosurgery Clinics 22 (2), 133–139.
Gonzalez-Rosa, J.J., et al., 2006. Differential cognitive impairment for diverse forms of multiple sclerosis. BMC Neurosci. 7, 1–11.
Gonzalez-Rosa, J.J., et al., 2011. Cluster analysis of behavioural and event-related potentials during a contingent negative variation paradigm in remitting-relapsing and benign forms of multiple sclerosis. BMC Neurol. 11, 1–19.
Gronwall, D.M.A., 1977. Paced auditory serial-addition task: a measure of recovery from concussion. Percept. Mot. Skills 44 (2), 367–373.
Hansch, E.C., Syndulko, K., Cohen, S.N., Goldberg, Z.I., Potvin, A.R., Tourtellotte, W.W., 1982. Cognition in parkinson disease: an event-related potential perspective. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society 11 (6), 599–607.
Hilgard, J., 2021. Maximal positive controls: a method for estimating the largest plausible effect size. J. Exp. Soc. Psychol. 93, 104082 https://doi.org/10.1016/j.jesp.2020.104082.
Honig, L.S., et al., 1992. Event-related potential P300 in multiple sclerosis: relation to magnetic resonance imaging and cognitive impairment. Archives of neurology 49 (1), 44–50.
Hossain, M.Z., Daskalaki, E., Brüstle, A., Desborough, J., Lueck, C.J., Suominen, H., 2022. The role of machine learning in developing non-magnetic resonance imaging based biomarkers for multiple sclerosis: a systematic review. BMC Medical Informatics and Decision Making 22 (1), 242.
Huang, C.C., Lu, Z., 2016. Community challenges in biomedical text mining over 10 years: success, failure and the future. Brief Bioinform 17 (1), 132–144.
Jamoussi, H., Ali, N.B., Missaoui, Y., Cherif, A., Oudia, N., Anane, N., Fredj, M., 2023. Cognitive impairment in multiple sclerosis: utility of electroencephalography. Mult. Scler. Relat. Disord. 70, 104502.
Kocer, B., Unal, T., Nazliel, B., Biyikli, Z., Yesilbudak, Z., Karakas, S., Irkec, C., 2008. Evaluating sub-clinical cognitive dysfunction and event-related potentials (P300) in clinically isolated syndrome. Neurol. Sci. 29, 435–444.
Kurtzke, J.F., 1983. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 33 (11), 1444.
Lebrun, C., Blanc, F., Brassat, D., Zephir, H., de Seze, J., & CFSEP, b. o. (2010). Cognitive function in radiologically isolated syndrome. *Multiple Sclerosis Journal, 16*(8), 919-925.
Leocani, L., Gonzalez-Rosa, J.J., Comi, G., 2010. Neurophysiological correlates of cognitive disturbances in multiple sclerosis. Neurol. Sci. 31 (Suppl 2), 249–253.
López-Góngora, M., Escartín, A., Martinez-Horta, S., Fernandez-Bobadilla, R., Querol, L., Romero, S., Riba, J., 2015. Neurophysiological evidence of compensatory brain mechanisms in early-stage multiple sclerosis. PLoS One 10 (8), e0136786.
Luck, S.J., Kappenman, E.S., 2011. The Oxford handbook of event-related potential components. Oxford University Press.
Magnié, M.N., Bensa, C., Laloux, L., Bertogliati, C., Faure, S., Lebrun, C., 2007. Contribution of cognitive evoked potentials for detecting early cognitive disorders in multiple sclerosis. Rev. Neurol. 163 (11), 1065–1074.
McGough, J.J., Faraone, S.V., 2009. Estimating the size of treatment effects: moving beyond p values. Psychiatry (Edgmont) 6 (10), 21.
Meyer-Moock, S., Feng, Y.-S., Maeurer, M., Dippel, F.-W., Kohlmann, T., 2014. Systematic literature review and validity evaluation of the expanded disability status scale (EDSS) and the multiple sclerosis functional composite (MSFC) in patients with multiple sclerosis. BMC Neurol. 14 (1), 1–10.
Nagels, G., D'hooghe, M. B., Kos, D., Engelborghs, S., & De Deyn, P. P. (2008). Within-session practice effect on paced auditory serial addition test in multiple sclerosis. Multiple Sclerosis Journal, 14(1), 106-111.
Newton, M.R., Barrett, G., Callanan, M.M., Towell, A.D., 1989. Cognitive event-related potentials in multiple sclerosis. Brain 112 (6), 1637–1660.
O'Donnell, B.F., Squires, N.K., Martz, M.J., Chen, J.R., Phay, A.J., 1987. Evoked potential changes and neuropsychological performance in Parkinson's disease. Biol. Psychol. 24 (1), 23–37.
Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Moher, D., 2021. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J. Clin. Epidemiol. 134, 103–112.
Paolicelli, D., Manni, A., Iaffaldano, A., Tancredi, G., Ricci, K., Gentile, E., Trojano, M., 2021. Magnetoencephalography and high-density electroencephalography study of acoustic event related potentials in early stage of multiple sclerosis: a pilot study on cognitive impairment and fatigue. Brain Sci. 11 (4), 481.
Pokryszko-Dragan, A., Dziadkowiak, E., Zagrajek, M., Slotwinski, K., Gruszka, E., Bilinska, M., Podemski, R., 2016a. Cognitive performance, fatigue and event-related potentials in patients with clinically isolated syndrome. Clin. Neurol. Neurosurg. 149, 68–74.
Pokryszko-Dragan, A., Zagrajek, M., Slotwinski, K., Bilinska, M., Gruszka, E., Podemski, R., 2016b. Event-related potentials and cognitive performance in multiple sclerosis patients with fatigue. Neurol. Sci. 37, 1545–1556.
Posner, M.I., 1980. Orienting of attention. Q. J. Exp. Psychol. 32 (1), 3–25.
Sandry, J., Simonet, D.V., Brandstadter, R., Krieger, S., Sand, I.K., Graney, R.A., Sumowski, J.F., 2021. The symbol digit modalities test (SDMT) is sensitive but non-specific in MS: lexical access speed, memory, and information processing speed independently contribute to SDMT performance. Mult. Scler. Relat. Disord. 51, 102950.

Seer, C., Lange, F., Georgiev, D., Jahanshahi, M., Kopp, B., 2016. Event-related potentials and cognition in parkinson's disease: an integrative review. Neurosci. Biobehav. Rev. 71, 691–714.
Sjøgård, M., Wens, V., Van Schependom, J., Costers, L., D'hooghe, M., D'haeseleer, M., ... & De Tiège, X. (2021). Brain dysconnectivity relates to disability and cognitive impairment in multiple sclerosis. Human brain mapping, 42(3), 626-643.
Smith, A. (1973). *Symbol digit modalities test* (p. 22). Los Angeles: Western psychological services.
Sullivan, G.M., Feinn, R., 2012. Using effect size—or why the P value is not enough. J. Grad. Med. Educ. 4 (3), 279–282.
Sundgren, M., et al., 2015. P300 amplitude and response speed relate to preserved cognitive function in relapsing–remitting multiple sclerosis. Clin. Neurophysiol. 126 (4), 689–697.
Tanaka, H., Koenig, T., Pascual-Marqui, R.D., Hirata, K., Kochi, K., Lehmann, D., 2000. Event-related potential and EEG measures in Parkinson's disease without and with dementia. Dement. Geriatr. Cogn. Disord. 11 (1), 39–45.
Taranu, D., Tumani, H., Holbrook, J., Tumani, V., Uttner, I., Fissler, P., 2022. The TRACK-MS test battery: a very brief tool to track multiple sclerosis-related cognitive impairment. Biomedicines 10 (11), 2975.
The Cochrane Collaboration. (2020). Review Manager (RevMan) (Version 5.4).
Tombaugh, T.N., 2006. A comprehensive review of the paced auditory serial addition test (PASAT). Arch. Clin. Neuropsychol. 21 (1), 53–76.
Turner, L.M., Croft, R.J., Churchyard, A., Looi, J.C., Apthorp, D., Georgiou-Karistianis, N., 2015. Abnormal electrophysiological motor responses in Huntington's disease: evidence of premanifest compensation. PLoS One 10 (9), e0138563.
Uysal, U., et al., 2014. Contingent negative variation is associated with cognitive dysfunction and secondary progressive disease course in multiple sclerosis. J. Clin. Neurol. 10 (4), 296–303.
Vázquez-Marrufo, M., et al., 2009. Cognitive evoked potentials in remitting-relapsing and benign forms of multiple sclerosis. Rev. Neurol. 48 (9), 453–458.
Vázquez-Marrufo, M., Galvao-Carmona, A., González-Rosa, J.J., Hidalgo-Muñoz, A.R., Borges, M., Ruiz-Peña, J.L., Izquierdo, G., 2014. Neural correlates of alerting and orienting impairment in multiple sclerosis patients. PLoS One 9 (5), e97226.
Vázquez-Marrufo, M., Galvao-Carmona, A., Caballero-Díaz, R., Borges, M., Paramo, M. D., Benítez-Lugo, M.L., Izquierdo, G., 2019. Altered individual behavioral and EEG parameters are related to the EDSS score in relapsing-remitting multiple sclerosis patients. PLoS One 14 (7), e0219594.
Vazquez-Marrufo, M. (2017). Event-related potentials for the study of cognition. *Event-Related Potentials and Evoked Potentials*, *1*.
Vazquez-Marrufo, M., Gonzalez-Rosa, J.J., Vaquero, E., Duque, P., Escera, C., Borges, M., Izquierdo, G., Gomez, C.M., 2008. Abnormal ERPs and high frequency bands power in multiple sclerosis [Research Support, Non-U.S. Gov't]. International Journal of Neuroscience 118 (1), 27–38.
Waliszewska-Prosół, M., Nowakowska-Kotas, M., Kotas, R., Bańkowski, T., Pokryszko-Dragan, A., Podemski, R., 2018. The relationship between event-related potentials, stress perception and personality type in patients with multiple sclerosis without cognitive impairment: a pilot study. Adv. Clin. Exp. Med. 27 (6), 787–794.