The Growing Potential of Neurophysiology in Multiple Sclerosis

Canham LJW*, Western DG1,2, Walsh P3, Kane N4, Inglis K5 and Cottrell DA6

1Bristol and Avon Multiple Sclerosis Service (BrAMS), Department of Neurosciences, Bristol Brain Centre, North Bristol NHS Trust, Southmead Hospital, Southmead Road, Westbury-On-Trym, Bristol BS10 5NB, UK
2Department of Mathematical Engineering, University of Bristol, UK
3The Grey-Walter Neurophysiology Department, North Bristol NHS Trust, Southmead Hospital, Southmead Road, Westbury-On-Trym, Bristol, BS10 5NB, UK

Abstract

Neurophysiological techniques have a long history of clinical and research use in the setting of Multiple Sclerosis. With the output of evoked potential and electroencephalographic recordings being a direct consequence of the underlying functional integrity and activity of the nervous system it is unsurprising there is currently a resurgent interest in their possible use as biomarkers of disease related dysfunction.

Herein we discuss the promise and possible pitfalls associated with their use in translational research of MS therapeutics. The numerous advantages of these methods are touched upon as demonstrated by their application to date. A number of obstacles which must be overcome prior to their successful widespread implementation are also outlined and readily achievable solutions discussed.

Keywords: Evoked potential; Multiple sclerosis; Electroencephalography; Biomarkers


A Timely Return to Neurophysiology for Answers

Multiple Sclerosis remains the most common acquired neurodegenerative disease affecting young adults [1]. The cardinal feature of this condition is demyelinating inflammation of the central nervous system which is disseminated in time and space [2]. Prior to the advent of neuroimaging and its subsequent incorporation into modern diagnostic criteria [3], the ground-breaking work of Halliday et al. [4,5] in demonstrating increased latency of visual evoked potentials [VEP] in the setting of optic neuritis offered a valid technique for confirming second-site involvement in this condition.

The principle of demyelination causing slowed transmission and axonal loss leading to relatively reduced amplitudes of potentials evoked in a response time-locked to standardised stimuli [6] underpins the application of these techniques across a range of modalities.

This scientific rationale is directly supported by the established linear augmentation of saltatory axonal conduction velocity conferred by myelination observed in vitro and in vivo [7]. The large scale perturbation and slowing of peripheral nerve conduction in neuropathies similarly characterised by pathology of demyelination is also broadly familiar and frequently present with a diagnostic specificity [6].

Although human post-mortem studies [8-13] have yielded invaluable insight into the pathological cascades producing MS the direct nature of their relationship to electrophysiological disturbance in such tissue is less well established. However, application of near identical evoked potential paradigms to animal models of certain MS components, namely variants of Experimental Allergic Encephalitis in mice, has unequivocally demonstrated the tripartite interaction between morphologic, electrophysiology and function [14-16]. Furthermore such models have elegantly and directly demonstrated the consequences of typical MS pathology on these properties and most recently, not only the neuro-preservant effects of some interventions [17,18] but the directly reparative and remyelinating effects of others [15,19]. Putting acknowledged caveats regarding the generalizability of animal models to the singularly human condition [20] of MS temporarily aside (including the undeniable difference in morphological scale between human and murine neural pathways [21]) the tight causal association between neurophysiology and clinical function seen with both disease and its positive response to intervention in the animal setting at the very least suggests promise for human investigations.

Although axonal loss is considered both a driver and determinant of disability in the progressive phase [12,13,22-25] of Multiple Sclerosis comparatively less focus has fallen on evoked potential amplitudes which are effectively a function of synchronous induced neural activity and hence neuronal numbers [6,26]. The dominant interest in latency, as a function principally of myelination both for clinical and research purposes is not however inappropriate.

Myelination is a reasonable principle target of investigation given a) demyelination is the hallmark feature of the disease pathologically [2,12,13], b) the primacy of demyelination in leading to secondary axonal degeneration [24,27], c) the appreciation that myelin contributes not only to conduction but also to invariable trophic support of axons and their survival [28] and d) the amount, extent and pattern

*Corresponding author: Canham LJW, Bristol and Avon Multiple Sclerosis Service (BrAMS), Department of Neurosciences, Bristol Brain Centre, North Bristol NHS Trust, Southmead Hospital, Southmead Road, Westbury-On-Trym, Bristol BS10 5NB, UK. Tel: +441179505050; E-mail: luke.canham@nbt.nhs.uk

Received August 25, 2016; Accepted November 21, 2016; Published November 28, 2016


Copyright: © 2016 Canham LJW, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
of myelination is exquisitely 'tuned' to provide optimal function [21]. Far from being a binary, categorical or even ordinal property, where elicited EP latency offers a relatively continuous interval scaled system of quantification which offers classically accepted measurement attributes [29]. A discontinuity arises for both amplitude and latency when no response is provoked [30]. Such behaviour is not uncommonly encountered in neurophysiological practice in MS patients [31] and undeniably contributes to a ceiling effect but nonetheless does not preclude such findings from being properly informative.

This is supported by the use of a range of quantitative and semi-qualitative EP rating systems which have been applied to recordings from MS patients [32-38]. The fine granularity of EP measurement is a major advantage over traditional clinical rating systems wherein despite a steady rate of decline compatible with other neurodegenerative conditions [39-46] individuals spend unequal periods at disability steps [40,44] owing to very arbitrary placement of albeit clinically relevant milestones.

This latter factor not only confers a statistically under-powering effect of recruiting individuals at certain levels of disability [40,44] [with a resultant bias against their participation] it also creates difficulty in attempting to validate candidate biomarkers with possibly superior measurement properties against ill-constructed and misapplied ordinal scales whose persistence as the regulatory 'gold standard' owes as much to simple familiarity as it does accepted tradition [47]. Moving toward more useful systems of measurement using the EDSS as a foundation is therefore calibration through 'bootstrapping' and inherently less ideal in contrast to the more scientifically familiar definition of metrics and units with the highest attainable precision as their internationally standardised basis.

In judging the utility of neurophysiological techniques to biologically mark functionally relevant disease related decline in Multiple Sclerosis we should not therefore expect perfect correlation with accepted ordinal outcome ratings even though moderate correlation should at least be present.

We might also anticipate a similar pattern of dynamic changes over time between clinical and neurophysiological status; whether contemporaneously or with the latter antecedent to the former, particularly if there is a genuine causal relationship between the two.

A final reasonable expectation would be that interventions able to prevent, attenuate or even reverse neurophysiological decline should have a much greater than chance effect at ameliorating clinical decline or improving physical status also.

Several key issues could however affect the translation of a positive neurophysiological response at or before phase II into success at phase III using a clinical outcome accepted by regulatory bodies, which currently remains the EDSS [48].

A positive neurophysiological response simply may not be enough to produce a detectable clinical benefit; either because of the small magnitude of the former, the coarse measurement of the latter or lack of relation to the property measured on neurophysiology to physical disability itself. This pitfall is well highlighted by the recent failure of agents considered putatively neuroprotective on the basis of a manifest ability to generally attenuate MRI measured brain volume loss at phase II [49,50] to translate into any semblance of positive impact on hard clinical disability outcomes in progressive MS at phase III [51,52] and is readily understood in light of the recent volumetric work by Daams et al. [53].

Therein it is evident corticospinal tract integrity accounts for much variance in EDSS and global brain atrophy measures, in cross-section simply do not [53]. Therefore whilst both may deteriorate with disease, the former is causally associated with the clinical outcome in question whereas the latter is simply co-variant (albeit outwith the duration of a standard clinical trial nonetheless may have longer term predictive value [54]). Therefore any selection of EP metrics to act as surrogates of disability should focus on those modalities with a deterministic relationship. It is also notable that certain agents currently licensed for use in MS [Dalfampridine] exert their beneficial effects directly by improving conduction [55], which is evident on neurophysiology and yet without any attendant benefit on myelination or other proven disease modifying effect. Without clear pre-existing appreciation of such an agent's biological effects on the basis of neurophysiology alone it could be misconstrued as having reparative action. Conversely it will be necessary to identify any directly pro-conductive effects in putative reparative therapeutics to similarly avoid misinterpretation of positive biomarker responses. The lack of pseudoatrophic responses during the initiation phases of some conventional disease modifying therapies and not others. De Stefano and Arnold [56] and Zivadinov et al. [57] suggest the impact on volumetric surrogates is not solely mediated by a simple homogenous reduction of inflammation. This contemporary observation further highlights the need for caution when inferring meaning from biomarker outcomes and going forward underscores the need for great rigour in exploring the validity of neurophysiological surrogates and the effects of candidate therapeutics upon them.

The trophic support effect of myelin is also well recognised [28,58-61] and may dominate in relative importance over conduction benefits per se and persist, enabling tract survival long after evoked responses are no longer reliably clinically detectable. In our own experience [32], although higher neurophysiological burden is unquestionably associated with greater physical disability, rarely some patients may nonetheless entirely lose recordable sensorimotor long tract EPs yet still manifest reasonable function and very slow decline.

That ultimate axonal loss perhaps bears a tighter cross-sectional association with disability than burden of demyelination [62,63] per se is perhaps congruent with such observations and an important consideration. However the absence of significant axonal loss without pre-existing myelin injury [13,27] as the index event reinforces the standing of EPs as both indicators of disease presence and ensuing downstream severity.

A further reflection is that to date exploration of combining different EP types into Multi-modal composite batteries has through various means of scoring and summation generally considered their contributions equally [33-36,38,64-67]. Given the variable eloquence and prognostic impact of identical MS pathology arising in different sites [41,68] such composites will perhaps not offer their optimal real world ecological value unless their components are tuned, for instance by regression, against desired clinical outcomes.

Finally, there is the challenge conferred by the millimetre or smaller scale of pathological change [8,10,11,69] and impressive sensitivity of neurophysiological techniques to it [70], which frequently exceeds the resolution of even contemporary imaging techniques and certainly clinical detection. As MS moves possibly into an era of simultaneously applied therapeutics used for the synergistic benefit of multiple actions--in a manner familiar to many other chronic diseases where real difference has been achieved, only the availability of such sensitive measures will enable accurate estimation of individual effect sizes. In isolation the magnitude of effect may be less than the standard and
large-trial demanding 0.3 but nonetheless substantial when applied synergistically (cf. combination therapy in cerebrovascular secondary prevention [71]). It is therefore a real concern that agents provoking neurophysiological improvements may be disregarded entirely from further consideration if they do not lead to accompanying clinical and radiological benefits as monotherapy, despite objective findings which point to promise as part of polytherapy.

The phenomenon of clinical deterioration in the absence of contemporaneous radiological alteration and vice versa is already familiar to the MS clinician [62,63]. Whilst neurophysiology may ultimately offer a route to understand such apparent structure-function dissociations [72-75], its cross-validation against such modalities is likely to generate similar challenges for contemplation.

The purpose of performing clinical disability ratings or biomarker quantification in the context of clinical investigation is to ascertain the impact of disease upon individuals, the effect of any candidate intervention on such a condition and the relationship of both disease and therapy to ultimate outcome [76]. Discerning such truth with reasonable statistical confidence demands precision and both the recognition and minimization of error.

It is broadly appreciated that a range of concrete and abstract factors hugely modify the translation of disease related damage into functional impairment and this subsequently into disability. In the setting of Multiple Sclerosis body weight, age and physical fitness are examples of the former, with mood, motivation and placebo effects typifying the latter.

Various subjective biases from clinician and patient alike also couple with the natural fluctuation [77-79] of clinical status to increase error and reduce precision. Such issues coupled with ordinality and not-insignificant phenotypic heterogeneity demand a prohibitive scale of investigation when using the EDSS outcome system [40,44,80].

Although it is unlikely to be imminently supplanted as the instrument of investigation in pivotal phase III studies supporting regulatory approval, the use of evoked potential analysis at phase II could offer metrics relatively invulnerable to many if not all of these confounding factors and facilitate more feasible enquiry [32].

Most EP paradigms involve the extraction of a summed average waveform from background physiological noise following the presentation of several hundred stimuli [31,70]. Such an approach therefore offers an objective, minimally variant and highly confident estimation of mean latency and amplitude of the underlying neurophysiological process. The rapidity of these processes which occur on the millisecond scale makes their acquisition either in isolation or as part of a wider battery logistically very feasible, non-labour intensive and ultimately affordable. In our experience they are also non-invasive, easy to perform and very well tolerated [32,70,81].

The standard MS clinical trial duration, whether interventional or observational is typically in the order of years and rarely more than 24 months [56]. Whilst evoked potential studies have demonstrated a relationship at baseline with the clinical change that ensues over such a period [33,34,37,65] and a natural deterioration over such an interval themselves, if they are to be a useful instrument in short [12 month] and small scale [n<50] ‘signal studies’ of putative reparative therapies it will be essential to quantify the normal near and medium term variability of such metrics. This is particularly pertinent in a condition characterised for many by almost daily fluctuation [77,78] in their real-world functional capacity.

Application of Evoked Potentials to MS

With respect to those EP techniques which have already gained widespread clinical application Visual, Brainstem, Somatosensory and Motor are the most established, with international consensus based standards of acquisition [26] and comparative normative data from healthy controls available to enable standardisation.

The original visual stimulus and recording paradigm of Halliday et al. [4] has evolved considerably with the development of multi-focal recording techniques [82,83] allowing finer localisation of abnormalities within the optic pathway and addition of electro-retinography has offered further insights and precision [84]. The development of fine structural imaging with MRI and optical coherence tomography [OCT] in particular has given support to the localisation and interpretation of VEP findings and yet not significantly surpassed their relationship with objective ratings of ocular function [85] whilst also appearing less sensitive to MS damage than VEP measurement [86]. A strong positive correlation between lesion size on MRI of the optic nerves and VEP latency has been observed in neuritis acutely [87,88], with follow up over several years [88] and in the progressive phase. Congruent with its manifestation of axonal integrity VEP amplitudes are proportional to both cross-sectional optic nerve area on MRI and also Retinal Nerve Fibre Layer [RNFL] thickness on OCT [89].

Local experience in the Grey Walter Neurophysiology Unit found that of 273 patients referred on clinical suspicion of MS, 92.5% demonstrated characteristic abnormality in those with eventual clinically definite diagnoses [70]. Typically 80% of patients will have an abnormality even without a history of optic neuritis [70], with this rising to over 90% in those who do [90]. A particular pitfall relating to use of newer digital rather than older and more precise optomechanical checkerboard stimuli presentation is thought to underlie variability in estimates of abnormality prevalence in the wider literature [70] and is therefore relevant during consideration of trial design.

Although the broad transverse course of the optic radiation fibres suggests VEP may offer a useful index of intracranial demyelination burden, extrapolation to serving as a cognitive biomarker would not be wholly valid as any relationship would be inherently non-deterministic.

Interrogation of afferent responses to auditory stimuli has offered indicators of demyelination of the respective pathways at the level of the brainstem [26,91,92]. Natural history imaging studies have demonstrated the adverse long term prognosis of brain stem damage which is indeed common [93], suggesting utility of incorporating Brainstem Auditory Evoked Potentials [BSEP] into larger multimodality batteries. However, the correlation between BSEP and Brainstem FSS score is only modest \( r=0.36 \) \( p=0.0008 \) [94] likely consequent of exquisite pathway specificity and FSS ordinality. Nonetheless sensitivity to subclinical damage is substantial with BSEP abnormality being evident in 40% of MS patients without clinically evident deficit [70]. Examination of the efferent component of the auditory pathway by measurement of Transient Evoked Otoacoustic Emissions from medial olivo-cochlear bundle mediated hair cell tuning is a newer technique which although simple and sensitive to MS damage [95,96] has enjoyed limited exploration to date [95-97].

Alternate brainstem modalities of facial sensation and vestibular function have been explored [again to a much lesser degree than BSEP] with elicitation of Brainstem Trigeminal EP [BTEP] [91] and Vestibulo-Ocular, Vestibulo-Masseteric and Auditory-Masseteric responses respectively [73,98]. Their performance in detection of subclinical abnormalities and relationship to Brainstem FSS is similar to conventional BSEP; however the combination of BSAEP and BTEP appears to offer a synergistic gain in sensitivity to brainstem functional
abnormality and lesion burden [91].

Exploration of long tract functional integrity with evoked potentials has a much longer history dating back to the original demonstration of ulnar nerve stimulation somatosensory EP by time-locked cathode ray tube photography in 1947 by Dawson [6] and the temporal intervals and morphology of peripherally elicited waveforms on route to the primary sensory cortex from all limbs have subsequently been well characterised [26]. Importantly, meaningful linear relationships between SSEP abnormality and quantitative sensory thresholds to vibratory and thermal stimuli in all limbs have been demonstrated [99-101]. Such abnormalities are also evident in 80% of MS patients without referable signs or symptoms suggesting impressive subclinical sensitivity [70,102]. The established paradigm relies on induction and subsequent conduction within the predominantly large-fibre dorsal column pathway [26]; although peripheral laser stimulation has yielded some specific insights as to spinthalamic integrity in MS patients [103], the residual tissue effects of the technique suggest further exploration at this time is likely to be limited.

The extensive anatomical course of the long tracts, which are typically over a metre in length, is considered a principle determinant of their eloquence from a neurophysiological perspective [30,104,105]. Their marked sensitivity, superior to both clinical and radiological detection [105], stands in contrast to the poor spatial resolution of the technique. Nonetheless, and unsurprisingly, in keeping with the physical disability of MS being principally driven by myopathy [32,53,106-111] at least as captured by the EDSS, the SSEP from lower limb stimulation bear the closest direct association with Global EDSS [101,104].

Examination of the efferent long tract integrity by the alternate means of non-invasive cortical stimulation with Transcranial Magnetic Stimulation [TMS] to provoke downstream electromyographic responses [MEP–Motor Evoked Potential] [26] has generated parallel findings [112]. Central Motor Conduction Latency correlates with voluntary phasic motor strength [113], limb motor function [114,115] and general walking ability [116]. As anticipated the burden of Central Conduction abnormality is also markedly greater in those patients with progressive disease compared those at the earlier relapse-remitting phase [117].

Application of TMS-MEP is not without two important caveats; firstly although broadly safe they are relatively contra-indicated in subjects with a liability to seizure disorder [118] itself more common amongst MS patients [119]. Secondly, the actual navigation of stimulus delivery by clinical means, although standard and logistically simple has generated concern about the differential elicitation of direct and indirect stimulation of the corticospinal fibres which in turn produces a variation in resultant MEP latency and ultimate interpretation [120]. Stereotactic navigation systems of TMS delivery, co-registered with subject neuroimaging have overcome this challenge and offered insight into the aberrant motor unit recruitment patterns within the motor cortices of MS patients [121]. However the current cost and limited availability of these systems is likely to prohibit their candidacy of becoming a widely employed translational biomarker. Even so, clinically navigated TMS MEP paradigms have offered not only meaningful quantification of corticospinal damage but repetitive stimulation with variation of interstimulus interval to elicit phenomena such as Intra-Cortical Inhibition and Intra-Cortical Facilitation, Cortical Silent Periods and particularly premotor facilitation mediated by cortico-cortical connectivity in the wider motor hierarchy have recently illustrated the relationship of damage therein to the previously poorly understood pathophysiology of fatigue [122-124] which is ubiquitous amongst MS patients.

**Application of Multi-Modality Evoked Potentials**

In respect of the multi-domain impairments encountered with MS the collection of quantified output from various modalities of EP into composite MMEP [Multi-Modal Evoked Potential] batteries is intuitive.

Numerous studies in both relapsing and progressive phenotypes have demonstrated a significant correlation between MMEP burden and physical disability rated by EDSS and the Multiple Sclerosis Functional Composite [MSFC] [32-34,64,65].

Several methods of abnormality quantification have been described [32], initially offering a binary categorisation of normal vs. abnormal waveforms and yielding a quotient of the number of abnormalities from the number actually acquired [The PATH-Q] [38,67].

A lower-resolution qualitative ordinal rating influenced by not only latency but also morphological characteristics including amplitude and asymmetry between contralateral recordings, termed the GEPS [Global Evoked Potential Score] [33,125] was subsequently deployed offering 0-3 points for every EP undertaken. A similar but higher resolution 0-5 point scale termed the MEP [Multimodal Evoked Potential Score] [34,64] was subsequently developed and included semi-quantitative rating of the EP based on its pathological prolongation against the established normal range [26] for the modality in question. Both techniques handled the discontinuity arising from absent response simply by awarding maximal points.

Fuhr’s group [36,37,65-67] have deployed a method of quantification focussed principally on the recorded latency of individual EP responses by offering summed Z scores from latencies in reference to published normative data. Alterations in morphology and amplitude are not included; this averts difficulties associated with inter-rater variability in making qualitative judgements of morphology and also the risk of erroneous interpretation of amplitude disparity between individuals due to non-disease related factors. Ultimately it offers a precise, unbiased and purely quantitative system, of fine granularity and standardised units. The problem of evaluating absent responses in this case however is met by the imperfect solution of using either the maximally recorded latency as a surrogate prior to normalisation or as we did by taking this value and simply adding 1 to offer Zmax+1 [32]. Nonetheless all such techniques have offered meaningful cross sectional and longitudinal findings and our own recent work demonstrates that the strength of relationship is perhaps largely independent of quantification scale selected [32]. This feature of relative scale-independence likely reflects genuine measurement of functional properties deterministically related to disability outcomes.

The positive association between EDSS and global EP burden on all scales is consistent amongst all MS phenotypes [33,34,67] but notably increases in strength with advancing disease severity [64].

A reasonable interpretation is that EP are sensitive to early damage, namely the index event of demyelination which is initially often subclinical [70] and masked by various forms of functional adaptation and a limited degree of repair. Over time such events have translated into downstream cascades of axonal loss and a growing disease burden has exhausted limits of adaptation and with such decompensation progressive accrual of disability arises [126].

Not only are recent imaging findings increasingly supportive of the view that initially non-eloquent lesions have profound longer term consequences [127,128], but the relatively consistent finding of higher
baseline EP abnormality burden predicting worse disability outcomes over short [34,129] and longer term [65,130] intervals is also congruent with such a model of ‘delayed effect’.

Longitudinal follow up of 245 patients with initial diagnoses of CIS demonstrated a substantially increased risk of moderate disability developing in those with a significant burden of abnormality on VEP, SSEP and MEP compared to those without [72]. Although conversion to CDMS risk was not related to EP burden in that investigation, this was found in a smaller cohort of 27 patients with solely SSEP and MEP considered, and to a degree greater than that predicted by MRI lesion burden in the same individuals at baseline [131].

Retrospective analysis of 94 subjects [130] with MMEP at baseline and followed up at 5 and 10 years subsequently demonstrated unequivocally increased risk and severity of disability progression in those with abnormalities on MEP and SSEP particularly. Giffroy et al. [132] have also recently published a 6 year retrospective analysis of a further 100 mixed-phenotype patients with impressively congruent findings demonstrating the independent adverse prognostic effect of a higher MMEP abnormality burden.

A smaller investigation of MEP in 15 RRMS patients has also highlighted the increased likelihood of worsening disability even over an interval as short as 6 months with a greater MEP abnormality burden [129]. Increased risk of disability progression with greater MMEP burden has been observed in both RR and PMS phenotypes over the typical 1-2 year time periods of clinical trials in several investigations [35,36,133].

In the context of RRMS this was demonstrated by Schlaeger et al. in 50 patients prospectively evaluated at 6 monthly intervals over 3 years. The baseline MMEP burden on VEP, SSEP and MEP correlated strongly [r>0.7, p<0.001] with final disability outcome [65]. The strength of association observed therein with the quantitative system of rating was greater than that observed in a study of 37 RRMS patients over a slightly shorter 2 year interval when graded by the qualitative ordinal scaling systems [r=0.39, p=0.02] [34]. This may highlight a superiority of purely quantitative MMEP evaluation. This said, in our own direct comparison of such methods [32] they all performed well with only a trend to superiority in association with disability measures from the higher-resolution qualitative semi-quantitative MEPS system.

Although clinical and paradigm heterogeneity may also account for the discrepancy in the above longitudinal studies, that no significant association was evident on cross-sectional evaluation at baseline but nonetheless emerged over the course of the investigation is in further support of the ‘delayed effect’ model. Indeed in a large scale retrospective analysis of 143 patients with Clinically Definite MS with a relapse-remitting phenotype and less than moderate disability there was observed to be no association with MMEP below an EDSS of 1.5 and between 1.5 and 3.5 only a weak rho of 0.39 [p=0.0114] [64]. Within the same cohort a similar trend to increasing strength and significance of positive association was also observed between MMEP and temporal interval between MS onset and time of baseline evaluation [64]. Furthermore regression based modelling using such data has begun to offer impressively accurate probability estimation for risk of disability progression over several years from the early stages of clinical dysfunction and may even offer a way to prospectively identify those with so termed ‘Benign MS’ who would presumably benefit from avoidance of the risks associated with current disease modifying therapies [64].

The longest interval study of baseline combined EP abnormalities [VEP and MEP in this case] and ultimate disability outcome is from 20 year follow up data from 28 initially relapsing patients published by Schlaeger et al. [65]. Association between MMEP and EDSS at 20 years was strong [r=-0.72, p<0.0001] and following inclusion into a regression model provided predictive ability unsurpassed by either consideration of the baseline clinical status or inflammatory indices on baseline MR imaging [65]. Such modelling conducted on prospectively collected data in a cohort of 22 patients with the purely progressive phenotype of PPMS over a 3 year period suggested an ability to predict final EDSS from baseline VEP, SSEP and MEP with 92% accuracy [67] and most intriguingly the dynamics of EP change preceded EDSS declines by an average of 6 months, again congruent with the model outlined above.

The large scale natural history studies of disability progression in MS [39,41,44,134-136] with anticipated variation are nonetheless remarkably consistent in the estimation of median time to salient clinical milestones. Importantly such averages are accompanied by an exceptionally broad range of clinical trajectories unfamiliar to most other neurodegenerative contexts. With both natural history and MMEP studies also supporting the adverse prognostic outcome particularly of myelopathic damage [36,134,136,137], with the former evidently being foreshadowed by the latter [138] and with the challenge to date of determining disease duration with any reasonable accuracy, it is perhaps not unreasonable to consider using MMEP to provide some form of disease staging, as so effectively applied in oncological and other medical disciplines featuring disease processes manifesting disseminated biological attack. The particular utility in translational research may in fact be to enrich recruitment for only those individuals likely to manifest confirmed disability progression within the forthcoming trial period—or alternately select for those of early chrono-biological stage to assess for the genuinely prophylactic effects of agents intended to prevent the index event of demyelination. Such an approach may offer a route to ultimately reduce trial size by favourably altering anticipated progression probabilities and partially disentangle the overlapping phenomena of inflammation and neurodegeneration in patient groups, which are not sufficiently segregated by the purely clinical criteria currently employed [80].

To date there have already been several historical and contemporary translational investigations using MMEP as outcome measures of intervention effect in the setting of MS, in addition to their aforementioned and increasingly accepted use in pre-clinical testing in animal models of the condition.

It is now 3 decades since the utilisation of MMEP [featuring VEP, BSAEP and upper limb SSEP] in the double blind placebo-controlled Azathioprine and Methylprednisolone Study involving panicipating 101 patients over 3 years [139]. In the investigation deterioration in VEP and SSEP were reported to parallel and precede clinical decline by an average of 1 year in the Chronic Progressive cohort [139]. The EP outcomes in even earlier studies of the effects of plasmapheresis were less conclusive but nonetheless were again seen to mirror clinical trajectory [140,141]. In subsequent studies of methylprednisolone treatment for acute relapse improvements in clinical status have been accompanied by positive changes in global EP scores and such findings extend to benefit within individual components, with MEP CMCT showing responses which match clinical motor resolution [120,142].

Contemporary disease modifying therapies including Interferon [143] and Natalizumab [144] have been associated with beneficial effects on EP outcomes and positive effects have also been observed in phase I investigation of mesenchymal stem cells in the setting of PMS [81]. Such a finding has prompted the selection of MMEP as the primary outcome measure in the subsequent phase II investigation of the technique in that
context [145]. Although such pioneering use remains unaccompanied amongst the increasingly numerous PMS studies currently underway [146], it is likely to represent the start of a growing trend particularly if successful and a relationship to improvements in physical disability is evident alongside a degree of neurophysiological rescue.

For the reasons discussed above, this is clearly not guaranteed. As the first successful pro-remyelination clinical trial [147] demonstrated— it is possible to successfully induce seemingly beneficial change in VEP over placebo but this may not be accompanied by contemporaneous benefits to structural metrics on OCT or more importantly objective ocular function. Similarly longitudinal observational studies of VEP over 3 and 5 years [148-151] following neuritis have shown spontaneous improvement to be not-uncommon and yet after the familiar clinical interval of 3-6 months such changes are not accompanied by ocular functional improvements. Electrophysiological improvements are considered a consequence of natural partial remyelination [152] and ion channel reorganisation in demyelinated regions.

Therefore, on one hand EP clearly represent a technique enabling detection of improvement and by inference some degree of often very subtle repair, and thereby stand in contrast to numerous other metrics and outcomes in which focus on retarding decline or tissue loss. On the other hand it remains uncertain how much EP benefit [for a given modality] is required to translate into a meaningful clinical outcome and given the discussion of delayed effects how long it may subsequently take to become apparent. It is not unreasonable to presume that any anti-progressive benefits of partial remyelination or similar may only be readily evident clinically several years post-intervention, in much the same way that the benefit of DMT on Disability Progression in the UK NHS Risk Sharing Scheme only became grossly evident after 3-6 years [153], having been not overtly apparent at earlier intervals [154].

Therefore again interpreting EP benefits, as with any finding, should be done with caution particularly if serving as a basis to inform subsequent phase III studies with clinical outcomes.

Robust power calculations to inform study design using EP outcomes are outstanding and will clearly vary according to quantification technique, composite battery, cohort demographics and clinical phenotype. Given the growing pace of investigation underway and planned in all groups featuring accepted clinical and radiological outcomes it would be both extremely useful and logistically readily achievable to incorporate EP batteries alongside such instruments. This would offer cross-validation with other modalities and enable rapid acquisition of normative data for EP behaviour within this disease.

Larger more extensive cohorts would also further illuminate the dynamics of EP change with increasing pathological burden. Is accrual of neurophysiological decline generally linear over time? or ‘front-loaded’ being greatest after index insults as is the case seen with structural metrics such as spinal cord atrophy [108,155] and Retinal Nerve Fibre Layer thinning following acute myelitis and neuritis respectively [156].

Closing the ‘Cerebral Gap’

Whilst MMEP have much in their favour as candidate biomarkers of physical disability, our own analysis demonstrated no meaningful relationship with cognitive performance [32]. This is unsurprising and highlights the need for an additional approach to capture this important domain which becomes impaired in the majority of patients [157]. The growing appreciation of the contribution of MS related cognitive impairment to real-world disability [158], occupational loss [159,160], health economic burden and most importantly the quality of life of patients [160-163], has not been met by a parallel surge of translational studies for cognitive interventions, with methodological constraints likely representing a greater barrier than any lack of putative candidates worthy of test [164].

The tightest relationship between cognitive performance and conventional imaging metrics are those of atrophy, in the cortical and deep grey structures particularly [57,165-177]. Newer modalities such as DTI and MRS which disclose pathology in ’Normal Appearing White Matter’ and ‘Normal Appearing Grey Matter’ also bear some relationship to cognitive performance particularly in frontal and limbic regions [166,168,178-182]. However, such volumetric metrics typically represent substantial and to date irreversible tissue loss with the result being an inability to manifest dynamic response to intervention, especially over short term intervals.

Cerebral tissue is the structural medium from which cognitive processing arises, however it is both the quantity and quality of functional coupling within and between specialised regions that provides the actual substrate of human thought [183,184]. Measurement of such coupling should offer superior relationships with objective cognitive performance compared to structural metrics and this has recently been demonstrated through the application of fMRI to key processing regions [185]. However temporal resolution of this modality is fundamentally constrained by the dynamics of neurovascular coupling which generate its output [186] and which are also demonstrably perturbed in Multiple Sclerosis [187].

In contrast, neurophysiological output is time-locked directly to the cortical neuronal activity generating cognitive processes [6]. Any inferiority in spatial resolution is more than compensated for by superior temporal [millisecond, ms] resolution which is of heightened relevance in the investigation of a condition wherein the dominant feature of its ‘Cognitive Footprint’ is reduced Information Processing Speed [IPS] [188].

It has long been appreciated that Cognitive Evoked Potentials can be elicited in an almost identical manner to those in the primary afferent pathways considered above [31]. In this instance the stimulus is typically a modality-independent discrepancy or ‘oddball’ embedded within a stream of regular presentations.

The attention-based decision in cognitive recognition of difference elicits a characteristic time-locked positive waveform ~300 ms later [189]. Several studies have demonstrated a prolongation of such P300 latencies in the setting of MS in a manner associated with IPS [190-194]. Recent methodological consensus [189] and normative data [195] from large cohorts have also emerged. It has been observed to dynamically improve over short intervals in response to the use of Methylprednisolone for MS relapse [190] and also Modafinil for MS Fatigue [196]. The impact of formal immunomodulatory therapies on P300 latency has been explored in small cohorts with varying outcomes [197,198] from which it is difficult to make conclusive inference.

Understanding of the physiological mechanism underlying the P300 waveform itself remains incomplete [199,200] which is perhaps a limitation to making deduction about effects and its wider implementation, nonetheless it does at least at the group level, offer an index of attentional decision making speed [201].

A lower-level sub-awareness response to detection of novelty or change with otherwise identical auditory paradigms is Mismatch Negativity [MMN], seen as a negative deflection typically 200 ms post stimulus [189]. Although the neuroanatomical basis of this passive response is better delineated and less dependent on any active engagement its exploration...
Identification of a reliable cognitive surrogate, through EEG analysis would enable application of a widely available and inexpensive technique with unbiased outcome production. It would also enable inclusion of the not insignificant number of patients currently excluded from clinical trials on purely mobility grounds [40,44].

**Conclusion**

MS is arguably the most complicated disease process affecting the most complex organ system known to humans. It is most unlikely any single modality in isolation will or even could provide the insights and metrics needed to translate therapeutic concepts into much needed validated treatments.

Neurophysiological techniques, of evoked potential recording and EEG analysis offer quantitative outputs which are causally related to cardinal disease processes, are of fine granularity and meet the proper scientific criteria for true systems of measurement.

International consensus for systems of acquisition is already widely established and the real world meaning of electrophysiological abnormality increasingly understood, in both the near and longer term.

The relationship of dysfunction to disability is not straightforward, with the former appearing to antedate the latter by a significant interval. This has implications for trial design and suggests the possible utility of introducing a staging model.

The tighter relationship of disability to functional compared to structural integrity is a common theme. But this advantage of neurophysiology over imaging could be coupled with the exquisite spatial resolution and pathological specificity of scanning to yield the most informative handle on MS. This likely applies equally to both the cognitive and physical aspects of the disease.

As required of a valid biomarker, Evoked Potential techniques have a solid conceptual basis and display meaningful cross-sectional and longitudinal relationships to the outcome ratings of importance in translational enquiry. With a need for methodological refinement, larger scale acquisition and clarification of norms readily acknowledged, the authors here feel further prompt exploration is surely warranted.

**References**


The Growing Potential of Neurophysiology in Multiple Sclerosis.