Using biomarkers to predict clinical outcomes in multiple sclerosis

Daniel Castle, 1, 2 Ray Wynford-Thomas, 1, 2 Sam Loveless, 1, 2 Emily Bentley, 2 Owain W Howell, 3 Emma C Tallantyre 1, 2

1 Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK
2 Helen Durham Centre for Neuroinflammation, University Hospital of Wales, Cardiff, UK
3 Institute of Life Science (ILS), Swansea University Medical School, Swansea, UK

ABSTRACT

Long-term outcomes in multiple sclerosis (MS) are highly varied and treatment with disease-modifying therapies carries significant risks. Finding tissue biomarkers that can predict clinical outcomes would be valuable in individualising treatment decisions for people with MS. Several candidate biomarkers—reflecting inflammation, neurodegeneration and glial pathophysiology—show promise for predicting outcomes. However, many candidates still require validation in cohorts with long-term follow-up and evaluation for their independent contribution in predicting outcome when models are adjusted for known demographic, clinical and radiological predictors. Given the complexity of MS pathophysiology, heterogeneous panels comprising a combination of biomarkers that encompass the various aspects of neurodegenerative, glial and immune pathology seen in MS, may enhance future predictions of outcome.

INTRODUCTION

Multiple sclerosis (MS) is a common cause of neurological disability in young adults. The diagnosis can usually be made with a high degree of accuracy, but clinical outcomes in MS are very varied. Disease-modifying therapies (DMTs) hold promise for improving long-term outcomes, but also carry significant risks. Individualised MS treatment relies heavily on the availability of robust predictors of prognosis, an area where tissue biomarkers may add value.

Biomarkers are characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.2 The availability of a biomarker that reliably predicts outcomes in MS would allow clinicians to counsel patients more effectively, guide treatment decisions and also influence patient selection for clinical trials. Several markers of MS outcome have emerged from clinical practice, for example, T2-hyperintense brain/spinal cord MRI lesions predicting conversion from clinically isolated syndrome (CIS) to MS, or the use of anti-JC virus antibody index to predict risk of progressive multifocal leukoencephalopathy on natalizumab. However, a more recent targeted, screening approach is being widely applied to the discovery of tissue biomarkers for MS and other neurological diseases. This has generated hundreds of publications proposing candidate markers but it is often unclear how valid and robust these biomarkers are and how they should be applied in clinical practice.

In 2001, the National Institutes of Health Biomarkers Definition Working Group working group published consensus guidelines on the terminology and potential roles of biomarkers (table 1). However, although the Food and Drug Administration and European Medicines Agency have rigorous processes for licensing medications, their mechanisms for validating biomarkers are less standardised, and few biomarkers ever gain their formal approval.

Here, we report the current status of prognostic tissue biomarkers in MS. The role of imaging markers and some other biomarker types (table 2) are well summarised elsewhere.
CANDIDATE TISSUE BIOMARKERS MS is fundamentally considered to be an autoimmune disorder of the central nervous system (CNS). While T-cells have long been considered central to its pathophysiology, there is a growing appreciation of the important role of B-cells and the glial compartment in MS pathogenesis. Neuroaxonal pathology is now well recognised in MS and is thought to be largely responsible for long-term disability. Genome-wide association studies have identified numerous variants that confer a modest increase in the risk of MS, while observational studies have highlighted the relevance of environmental factors such as smoking, Epstein–Barr virus infection, vitamin D status and obesity. This knowledge has influenced the search for candidate biomarkers of outcome.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Glossary of terms formed by a Joint Leadership Council of the FDA and the NIH working group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker type</td>
<td>Definition</td>
</tr>
<tr>
<td>Susceptibility/risk biomarker</td>
<td>A biomarker that indicates the potential for developing a disease or medical condition in an individual who does not currently have clinically apparent disease or the medical condition.</td>
</tr>
<tr>
<td>Diagnostic biomarker</td>
<td>A biomarker used to detect or confirm presence of a disease or condition of interest or to identify individuals with a subtype of the disease.</td>
</tr>
<tr>
<td>Monitoring biomarker</td>
<td>A biomarker measured serially for assessing status of a disease or medical condition or for evidence of exposure to (or effect of) a medical product or an environmental agent.</td>
</tr>
<tr>
<td>Prognostic biomarker</td>
<td>A biomarker used to identify likelihood of a clinical event, disease recurrence or progression in patients who have the disease or medical condition of interest.</td>
</tr>
<tr>
<td>Predictive biomarker</td>
<td>A biomarker used to identify individuals who are more likely than similar individuals without the biomarker to experience a favourable or unfavourable effect from exposure to a medical product or an environmental agent.</td>
</tr>
<tr>
<td>Pharmacodynamic/response biomarker</td>
<td>A biomarker used to show that a biological response has occurred in an individual who has been exposed to a medical product or an environmental agent.</td>
</tr>
<tr>
<td>Safety biomarker</td>
<td>A biomarker measured before or after an exposure to a medical product or an environmental agent to indicate the likelihood, presence or extent of toxicity as an adverse effect.</td>
</tr>
<tr>
<td>Validated surrogate endpoint</td>
<td>An endpoint supported by a clear mechanistic rationale and clinical data providing strong evidence that an effect on the surrogate endpoint predicts a specific clinical benefit.</td>
</tr>
<tr>
<td>Validation of a biomarker</td>
<td>A process to establish that the performance of biomarker is acceptable for its intended purpose</td>
</tr>
</tbody>
</table>

Taken from the BEST produced in 2015.

BEST, Biomarkers, Endpoints, and other Tools; FDA, Food and Drug Administration; NIH, National Institutes of Health.
### Markers of Inflammation

#### Oligoclonal bands/Free light chains

Oligoclonal bands (OCBs) are bands of immunoglobulin (usually IgG), produced by clonal plasma cells. The presence of two or more bands in cerebrospinal fluid (CSF), but not serum, provides evidence of an intrathecal humoral immune response. The identification of CSF-specific OCBs has recently been adopted as a confirmatory criterion for the diagnosis of MS in people with a CIS plus MRI evidence of dissemination in space. However, OCBs also provide prognostic information. A meta-analysis of 12,253 people with MS reported the prevalence of CSF-specific OCBs at 87.7%. CSF-specific OCBs detected using isoelectric focusing (the gold standard) in 1918 people with MS were associated with an OR of 1.96 for reaching a disability outcome (Expanded Disability Status Scale [EDSS] 4 at 10 years, EDSS 6 during follow-up or increase of ≥1 EDSS point in 5 years). Likewise, Joseph et al found that OCB-negative status was associated with a HR of 0.65 (p=0.06) for reaching EDSS 4 and 0.52 (p=0.04) for reaching EDSS 6 (mean follow-up 72 months, adjusted for age, clinical presentation and baseline MRI findings). Tintore et al also reported that in 792 people with CIS, OCB-positive status was associated with a HR of 2.0 (95%CI 1.2 to 3.6) for reaching EDSS 3 (mean follow-up 6.8 years, adjusted for known clinicoradiological predictors of outcome).

Immunoglobulin-free light chains (FLC) are produced in excess during intrathecal immunoglobulin production and circulate in the CSF. The presence of CSF kappa and lambda free light chains (k-FLC and λ-FLC) is recognised in MS. CSF k-FLC may rival OCBs in diagnostic accuracy for MS, and may also offer modest prognostic information. In an early study, Rudick et al found CSF k-FLC in the upper quartile of the range to be a predictor of 36 month EDSS in 36 people with MS (HR 3.78, p=0.04). 11 Makshakov et al reported that in 98 people with CIS who subsequently converted to MS, k-FLC concentration at the time of first attack correlated with EDSS at 2 years (r=0.45, p=0.0016). 9 Rinker et al also found CSF k-FLC concentrations above 1.53 micg/mL predicted the need for ambulatory assistance in 57 people with MS when followed for a median of 15 years (specificity 87.5%, positive predictive value 88.9%). 12 While Voortman et al found no correlation between baseline CSF k-FLC and subsequent EDSS progression or relapse rate in 61 people with CIS or MS, by 4.8 years only 36 (59%) had converted to MS and most (61%) had received DMTs. 13 More recently, Rathbun et al suggested that the ratio of k:λ FLC may be more predictive than either in isolation. They found a CSF k:λ FLC ratio <9 at diagnostic lumbar puncture was moderately predictive of an EDSS >3 at 5 years in people with CIS/MS (75% sensitivity and 57.1% specificity). 14

#### Inflammatory cytokines

C-X-C motif chemokine 13 (CXCL13) is a potent B-cell activating chemokine, important in recruiting lymphocytes to sites of inflammation. Khademi et al found significantly higher CSF CXCL13 concentrations in people who had experienced a high cumulative number of MS relapses over their disease course (n=323), and that baseline CSF CXCL13 concentration was highest in those who went on to have over four relapses in the subsequent 5 years (n=52). They found no significant association between CSF CXCL13 levels and EDSS. 15 Alvarez et al reported a significant correlation between CSF CXCL13 and EDSS but this was in a small group of people with either neuromyelitis optica (n=9) or MS (n=9), and the effect may have been driven by high EDSS.

---

**Table 2: Applications of biomarkers with examples relevant to MS**

<table>
<thead>
<tr>
<th>Potential applications of biomarkers</th>
<th>Example in MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Providing proof of concept in early clinical trials</td>
<td>Brain volume as a surrogate endpoint in phase II trials of disease modifying therapy</td>
</tr>
<tr>
<td>Monitoring disease activity</td>
<td>T2 hyperintense and gadolinium enhancing lesions as subclinical markers of disease activity</td>
</tr>
<tr>
<td>Expediting disease detection</td>
<td>Oligoclonal bands confirming diagnosis of MS in a patient with typical clinical and imaging features</td>
</tr>
<tr>
<td>Predicting clinical outcome</td>
<td>Number of T2 lesions on baseline MRI brain scan predicting time to disability milestones</td>
</tr>
<tr>
<td>Predicting treatment response</td>
<td>Neutalising antibody status predicting likelihood of response to beta-interferon</td>
</tr>
</tbody>
</table>

---

MS, multiple sclerosis.
grades in neuromyelitis optica. Novakova et al found no relationship between CSF CXCL13 and either relapse rate or EDSS in 59 people with MS.17

CSF CXCL13 also shows a treatment response in MS. One group has shown that the magnitude of reduction in CSF CXCL13 is greater in those who receive immune-modulating treatment than those who receive placebo.18 Alvarez et al found that a higher CSF CXCL13 concentration at baseline was associated with response to rituximab (defined as no evidence of disease activity, NEDA) in 30 people with MS, and observed reductions in CSF CXCL13 after treatment.19 However, some of these studies also highlight the lack of specificity offered by CSF CXCL13, which can be significantly elevated in people with other neuroinflammatory disorders and CNS infections.15 16

MARKERS OF NEURODEGENERATION

Neuronal proteins are promising biomarkers in MS. The most studied of this subgroup are the neurofilaments, comprising neurofilament heavy, medium and light chains and ainternexin. These proteins contribute to the neuronal cytoskeleton and allow radial cell growth. Increased concentration in CSF and serum provides evidence of CNS axonal damage. The relatively short half-life of neurofilament light (NFL) compared with the more extensively phosphorylated heavy chains suggests that levels are more likely to reflect recent axonal damage.

NFL has become one of the most widely studied prognostic biomarkers in MS. Most early work was performed on CSF-NFL; the low sensitivity of ELISA techniques was previously a major barrier to using serum NFL as a prognostic biomarker. However, a novel single-molecule array technique (SiMoA) is now available to measure NFL in the serum.20 SiMoA appears sensitive and shows strong correlations between CSF-NFL and serum-NFL (sNFL) concentrations suggesting promise for a less invasive method of prognostic testing.21–23 Numerous studies have shown that CSF-NFL rises at the time of clinical relapse,17 24–26 a finding that has also been reproduced using sNFL.21–23 CSF-NFL appears to remain elevated for 2–3 months after relapse before returning to baseline.27 There have been no consistent correlations between either serum or CSF-NFL and concurrent disability. Correlations tend to be imperceptible in small cohorts,17 23 26 28 lost in multivariate analyses by the addition of clinicodemographic covariates,29 or associated with relatively modest effect sizes in larger cohorts.21 22 27

Starting a DMT is often associated with a fall in CSF-NFL,17 25 26 28 30 and sNFL.21–23 However, it is difficult to exclude regression to the mean in many of these studies given that patients often start DMT in response to a recent relapse.

Several studies have validated the prediction of future disability using CSF-NFL but the effect sizes are modest. In one of the earliest studies, Lycke et al found that CSF-NFL at baseline predicted annualised relapse rate at 2 years in 60 people with treatment-naïve relapsing-MS (r=0.38, p<0.01).27 Selleberg et al found that twofold higher baseline CSF-NFL was associated with a HR of 1.17 for relapse at 2 years (95% CI 1.01 to 1.36), but with no association between base-line CSF-NFL and conversion to secondary progressive (SP) MS or change in EDSS at 5 years.31 Bhan et al showed that for 35 people with relapsing-remit-ting MS at baseline, conversion to SPMS at 5 years was associated with a significantly higher median baseline CSF-NFL but with wide variation (2122 ng/L (IQR 1921 ng/L) vs 246 ng/L (IQR 885 ng/L), p=0.01).32 In a small cohort of people with progressive MS, Trentini et al found CSF-NFL significantly predicted EDSS progression over 9 years of follow-up, (0.44, R2=0.195, p=0.025).33 Salzer et al found that CSF-NFL at diagnostic lumbar puncture in 89 people with MS correlated with MS severity scale (MSSS) after median 14 years of follow-up (r=0.3, p=0.005), and that baseline CSF-NFL >386 ng/L was associated with a higher risk of severe MS (MSSS above median) at 14 years (OR 5.2, 95% CI 1.8 to 15).34

Serum NFL has also been validated in some studies to exert modest predictions of MS outcomes. Disanto et al found baseline sNFL predicted an increased risk of relapse in the subsequent 12 months in an unadjusted analysis in 388 people with CIS/MS (HR 1.94 (95% CI 1.21 to 3.10), p=0.006). Risk of EDSS worsening in the 12 months after blood sampling was higher in people with sNFL>97.5th percentile (appropriate for age) versus those with sNFL<80th percentile (15% vs 6.7%; OR 2.41, (95% CI 1.07 to 5.42) p=0.03).34 Barro et al found that sNFL>90th percentile predicted EDSS worsening in the next 12 months, in a multivariate model adjusting for other predictors of outcome including baseline T2 lesion load in 259 people with MS (OR 2.786 (95% CI 1.609 to 4.826) p<0.001).35 sNFL does not predict medium-term or long-term outcomes in MS. Piehl et al. (n=243) and Kuhle et al. (n=42) found no associations between baseline sNFL and 2 year outcomes.23 36 Chitnis et al also found no relationship between sNFL and 10-year disability outcomes (n=122).37

There are no universally agreed clinically relevant cut-off values for CSF and serum NFL. Some authors have used the median value in their cohort, whereas others have expressed NFL concentrations relative to age-matched controls. One study found that CSF-NFL above 450 ng/L was associated with a sensitivity of 81% and
specificity of 57% for disease activity at 4 years, and CSF NfL above the median expected value in an age-matched control was associated with 67% sensitivity and 75% specificity for disease activity at 12 months. Likewise, in the same two studies, sNfL above 14.2 ng/L provided a sensitivity of 72% and specificity of 57% for disease activity (modified NEDA) at 4 years, and sNfL above 18.2 ng/L was associated with 45% sensitivity and 80% specificity for disease activity at 12 months.

In summary, CSF-NfL appears to be of some help in predicting short-term and long-term outcomes in MS, while sNfL is more useful for predicting short-term outcomes. However, the effect sizes are modest and confidence intervals are wide.

Neurofilament heavy (NfH) in the CSF has been shown in several studies to be a surrogate marker of current ability but with no established utility as a predictor of outcome. Neurofilament medium and α-internexin have been studied in less detail in MS, in part related to difficulties in their analysis.

GLIAL MARKERS

Astroglial activation is associated with activation of the immune cascade and is thought to play a role in the demyelination and neuroaxonal injury observed in MS. Glial fibrillar acidic protein (GFAP) is the major constituent of gliotic scarring. CSF-GFAP correlates modestly with EDSS in cross-sectional analyses, and baseline CSF-GFAP appears modestly predictive of long-term clinical outcomes. Axelsson et al found baseline CSF-GFAP predicted EDSS score at 8–10 years (r=0.45, p<0.05). Martinez et al found that baseline CSF-GFAP predicted time to EDSS 3 (HR 1.83 [95% CI 1.01 to 3.35], p=0.04). Encouragingly, Abdelhak et al recently reported using SiMoA technology to measure serum GFAP in MS and found that after correction for age, serum GFAP concentrations modestly correlated with EDSS in progressive MS (r=0.4, p=0.008).

Chitinase-like proteins (YKL-40), expressed by astrocytes in the CNS, is thought to play a role in chronic inflammation in MS via regulation of the innate immune system. Modvig et al found baseline CSF-CHI3L1 predicted cognitive impairment at 13.6 years (mean) in 34 people with MS (r=0.41, p=0.015). Martinez et al. found CHI3L1 was the best predictor of disability progression in a multivariate analysis of several CSF biomarkers. CHI3L1 predicted time to EDSS 3 (HR 2.78 [95% CI 1.48 to 5.23], p=0.001) and time to EDSS 6 (HR 5.7 [95% CI 1.01 to 20.83], p=0.05). Canto et al. found CSF-CHI3L1 to be the only significant independent risk factor for reaching EDSS 3 (HR=3.82; p=5.3×10−8) in 813 people with CIS or MS, in a multi-variate analysis including baseline MRI and OCB status. CSF-CHI3L1 levels>170 ng/mL were associated with 74% sensitivity and 60% specificity of reaching EDSS 3 at (mean) 5.4 years.

CLINICAL USE OF CANDIDATE BIOMARKERS IN MS

There is a long list of candidate biomarkers that differ in expression between people with MS and healthy controls but require validation for prognostic utility (figure 1). There are several features to consider regarding possible prognostic biomarkers in the clinic (table 3). First, the in vitro test of a biomarker requires laboratory validation for its sensitivity, reproducibility, repeatability and standardisation in relation to reference materials. An inappropriate assay could lead to false negative results as is likely in studies using older versions of sNfL assay. Second, analytical validation requires a disease cohort to measure the sensitivity and specificity of the marker for relevant outcomes. It is important to distinguish markers that correlate with and markers that predict outcome. There is no universal agreement on the optimum clinical outcome measure for prognostic biomarkers in MS. Many groups have explored the prediction of disease activity (e.g., relapses) in the subsequent 12–24 months, which would provide a rationale for influencing short-term treatment decisions. Others argue that an ideal MS biomarker would predict disability status after 10–20 years. However, this approach has some limitations, such as the known inadequacies of MS disability rating scales, which may undermine the ability to demonstrate the predictive performance of a biomarker, for example, baseline sNfL correlates with 10-year brain volume but not 10-year disability. Further, the widespread and often heterogeneous use of DMTs in any MS cohort prohibits the study of the prediction of long-term outcomes. Almost all of the studies aimed at validating the utility of NfL as a biomarker used cohorts with DMT exposure rates exceeding 40%.
After analytical validation, candidate biomarkers require clinical validation in independent large clinical cohorts from different centres. Clinical validation should also evaluate a biomarker’s independent contribution in predicting outcome when models are adjusted for known demographic, clinical and radiological prognostic factors. A minority of studies have incorporated candidate biomarkers into multivariate models adjusting for known demographic, clinical and radiological predictors of outcome. Furthermore, several studies have highlighted the lack of specificity of several candidate biomarkers, which may give confounding effects of, for example, coexistent neurological disease, concurrent systemic illness, acute trauma or age. Before a candidate could be adopted into routine clinical use, a reference range is required, including age-adjustment where relevant for example, in the case of NFL.

THE FUTURE OF MS PREDICTIVE BIOMARKERS

None of the existing tissue biomarkers of MS prognosis appear ready for widespread clinical use. Given the gaps in the current data, identifying markers that will most likely help with clinical prognosis remains speculative. Given the complexity of MS pathophysiology, we probably need panels comprising a combination of markers. Other advances include the ability of screening technologies to profile a person’s gene expression, cytokine transcription, T-cell receptors, autoantibody repertoire, microRNA expression or the microvesicle contents, providing a unique signature that may better predict outcome or treatment response.

There are clear advantages in the convenience of a serum biomarker, avoiding the need for lumbar puncture. However, the utility of serum biomarkers will inevitably depend on the extent to which the CNS pathology in MS is reflected by changes in the blood. There is likely to be a trade-off between the convenience of obtaining samples and their disease relevance; the most valuable biomarkers may be derived from the disease-relevant cell populations in the CNS.
CONCLUSION

Clinical outcomes in MS are highly varied and existing prognostic variables are of limited value in predicting outcome for an individual. Tissue biomarkers show promise for improving the prediction of outcome in MS and may also offer additional value in diagnosis and monitoring. While some candidate markers show considerable promise, most are not disease-specific and may require adjustment for confounding effects of age and/or comorbidity. In the context of a disease with complex pathophysiology, it may be best in the future to develop heterogeneous panels of biomarkers that encompass the various aspects of neurodegenerative, glial and immune pathology seen in MS.

REFERENCES

29 KhalilM, EnzingerC, LangkammerC, et al. CSF neurofilament and N-acetylaspartate related brain changes in clinically