Neurofilament levels in combination with CCL22 and osteopontin in cerebrospinal fluid in clinically isolated syndrome and multiple sclerosis at baseline predict disease activity during two years of follow-up.

I Håkansson[a], A Tisell[b12], P Cassel[c], Blennow K[d], Zetterberg H[d], Lundberg P[b123], C Dahle [a], M Vrethem [a], J Ernerudh[c]

[a] Department of Neurology and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
[b1] Radiation Physics, Department of Medical and Health Sciences , Linköping University
[b2] Center for Medical Image Science and Visualization (CMIV), Linköping University
[b3] Radiology, Department of Medical and Health Sciences, Linköping University
[c] Department of Clinical Immunology and Transfusion Medicine and Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden
[d] C
istället för adressen ovan har vi två "nya":

d) Inst. of Neuroscience and Physiology, Dept. of Psychiatry and Neurochemistry, The Sahlgrenska Academy at University of Gothenburg, Sweden

e) Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden, Department of Molecular Neuroscience, UCL, London, UK.

Corresponding author
Irene Håkansson, Department of Neurology, University Hospital in Linköping, 581 85 Linköping, Sweden. Phone: +46 103 0000
Email: irene.hakansson@regionostergotland.se

Keywords: neurofilament light chain, CCL22, osteopontin, multiple sclerosis, clinically isolated syndrome, no evidence of disease activity, prognosis, cerebrospinal fluid, cytokines, chemokines

Abstract

Background: Improved biomarkers are needed to aid in clinical decision making and as surrogate endpoints in clinical trials in multiple sclerosis (MS).

Objective: To assess whether combinations of neurodegenerative and neuroinflammatory markers in cerebrospinal fluid (CSF) at diagnostic CSF sampling could predict disease activity during two years of follow-up in patients with clinically isolated syndrome (CIS) and MS.

Methods: Using multiplex bead array and enzyme-linked immunosorbent assay, CXCL1, CXCL8, CXCL10, CXCL13, CCL20, CCL22, neurofilament light chain (NFL), neurofilament heavy chain, glial fibrillary acidic protein, chitinase-3-like-1, matrix metalloproteinase-9 and osteopontin were analysed in CSF in 44 patients with CIS or MS and 23 healthy controls.

Results: In a logistic regression analysis model, NFL in CSF at baseline correctly classified 93% of patients that showed evidence of disease activity (relapses, magnetic resonance imaging activity or disability worsening) during two years of follow-up and 62% of patients that did not, with an overall percentage of 83% correctly classified. Combining NFL with CCL22 and osteopontin improved results to overall 91% correctly classified.

Conclusions: This study demonstrates the prognostic value of NFL in baseline-CSF in CIS and MS and indicates an added value of combining NFL with inflammatory markers.

1. Introduction

Multiple sclerosis (MS) is a chronic disease characterized by inflammation and degeneration of the central nervous system (CNS). The prognosis in MS is highly variable and magnetic resonance imaging (MRI), although being a very important diagnostic and follow-up tool, is not an optimal prognostic tool since conventional MRI data, like T2 lesion load, correlate poorly to the clinical disease manifestations and disease progression[1]. Hence, there is a need for non-MRI biomarkers in MS to aid individualized pharmacological treatment early in the disease course. The ample research in MS biomarkers has recently been reviewed[2-4]. Neurofilament light chain (NFL) in CSF has been reported to correlate to long term (8 – 20 years) prognosis[5] and treating MS with fingolimod and natalizumab has been shown to decrease NFL levels[6, 7]. Whether it is appropriate or not to measure NFL in CSF when monitoring response to treatment, to guide clinical decisions on therapy, is currently debated[8, 9]. There are complex relationships and overlaps between neuroinflammation and neurodegeneration in MS. Inflammatory biomarkers and markers of axonal damage in CSF have been shown to correlate, in different combinations in different types of MS[10]. However, results are inconsistent and there is no biomarker that is entirely specific for MS[3].

In this study, we evaluated a broad panel of neurodegenerative and inflammatory markers. NFL, neurofilament heavy chain (NFH) and glial fibrillary acidic protein (GFAP) were included as established markers of axonal degeneration[11] and astrogliosis[3], respectively. Chitinase-3-like 1 protein (CHI3L1) is as a marker of activated astrocytes and microglia[3] and matrix metallopeptidase 9 (MMP-9) is a matrix degrading enzyme that both have been suggested as surrogate markers of disease activity in multiple sclerosis[12, 13]. Osteopontin (OPN) is an early activation marker on T cells[14] that has been reported to be increased in

MS[15]. Chemokines are attractive markers because they recruit specific inflammatory cells and reflect ongoing type of inflammation. Also, chemokines are generally present at higher concentrations than cytokines and thereby easier to reliably measure in plasma and CSF. In MS, CXCL13 (a B cell chemoattractant) has been reported to be elevated in CSF[16] and associated with disease exacerbations and unfavorable prognosis[17]. Chemokines recruit different settings of Th subsets and they therefore represent and can be used as markers of Th1 (CXCL1, CXCL10), Th2 (CCL22) and Th17 (CXCL8, CCL20) immunity[18].

There is no universally accepted definition of freedom of disease activity in MS. "No evidence of disease activity" (NEDA) is increasingly used as a comprehensive measure of treatment response in MS and the combination of no relapses, no brain MRI activity (new or enlarging T2 lesions or Gadolinium-enhancing lesions) and no disability worsening is often called NEDA-3. Expanding the NEDA concept to NEDA-4, where brain volume loss is added, or NEDA-5, where a biomarker in CSF or serum is also added, is currently discussed[19].

This study aimed to assess whether concentrations of neurodegenerative and neuroinflammatory markers in CSF and plasma at baseline were able to predict disease activity during two years of follow-up in a longitudinal cohort of patients with CIS and MS, and to identify the best prognostic markers in our material. In this study we report that NFL in CSF at baseline classified 83% of patients correctly with regard to NEDA-3 or not NEDA-3 for two years of follow-up. Combining NFL with CCL22 and OPN improved results to 91% correctly classified patients.

Material and methods

Patients and controls

44 patients with CIS or MS were consecutively enrolled in a prospective longitudinal cohort study of CIS and newly diagnosed MS at the Department of Neurology at the University Hospital in Linköping, Sweden. All patients fulfilled the revised McDonald criteria from 2010[20] for CIS or MS. Patients underwent clinical neurological examination including expanded disability status scale (EDSS), peripheral blood and CSF sampling and MRI at baseline and at one and two years of follow-up. Patients received immunomodulatory treatment according to Swedish clinical practice. Patient characteristics are presented in Table 1 and Table 2.

For peripheral blood and CSF, 23 age- and sex-matched healthy controls (HC) were recruited from healthy blood donors (Table 1). Healthy controls were free from past and current neurological and autoimmune disease and their clinical neurological examinations were normal as were routine findings in CSF (Table 1). No medication, except oral contraceptive pills, was allowed in healthy controls.

Table 1. Patient and healthy control characteristics at baseline.

Clinical and laboratory data	Patients n=44	Healthy controls n=23	p-value
Women/men (%women)	35/9 (80%)	18/5 (78%)	0.9
Age median, 25:th-75:th percentile	31 (25-42)	32 (26-46)	0.4
Diagnosis (CIS/RRMS/PPMS)	19/22/3	NA	
Relapse within last 3 months	25	NA	
Relapse within last 2 months	23	NA	
Relapse within last 1 months	16	NA	
Mean disease duration (months)	13	NA	
Median disease duration (months)	4.8	NA	
Disease duration (number of subjects)		NA	
0-1 months	*10*		
1.25-2 months	*7*		
2.25-3 months	*3*		
3.25-6 months	*6*		
6.5-12 months	*7*		

13-24 months	*5*		
25-36 months	*2*		
37-48 months	*1*		
49-60 months	*1*		
61-72 months	*1*		
120 months	*1*		
Median EDSS	2.0	NA	
EDSS (number of subjects)			
0	*6*	*23*	
1.0	*12*		
1.5	*2*		
2.0	*14*		
2.5	*4*		
3.0	*1*		
3.5	*2*		
4.0	*2*		
4.5	*1*		
Treatment (number of subjects on DMT)	0	NA	
CSF mononuclear cell count*	4.5 (1.9-8.9)	2.1 (0.9-2.6)	0.001
Albumin ratio*	4.8 (3.4-5.8)	4.9 (3.6-5.5)	0.7
IgG index*	0.7 (0.5-1.2)	0.5 (0.5-0.5)	<0.001
IgG synthesis index*	1.3 (1.0-2.2)	0.9 (0.9-1.0)	<0.001
Oligoclonal CSF IgG bands (pos/neg)	35/9	0/23	

P-values from Chi-square test for sex distribution and Mann-Whitney U test for independent samples for age and CSF data.
**Median and within brackets interquartile range.*
NA, not applicable
DMT, disease-modifying treatment

Table 2. Patient diagnoses, relapse status and treatment status at baseline and at one and two years of follow-up.

Clinical and laboratory data	Baseline	1 year	2 years
Number of subjects	44	44	43
Diagnosis (CIS/RRMS/PPMS)	19/22/3	12/29/3	9/31/3
Relapse within last 2 months (yes/no)	23/21	4/40	2/41
Treatment (number of subjects)			
No DMT	*44*	*20*	*21*
Interferon-β 1b	*0*	*18*	*12*
Interferon-β 1a	*0*	*1*	*1*
Fingolimod	*0*	*1*	*2*
Natalizumab	*0*	*4*	*7*

Ethics Statement

The study was approved by The Regional Ethics Committee in Linköping and written informed consent was obtained from all participants.

Cerebrospinal fluid and plasma analyses

All CSF sampling was carried out by the same neurologist (IH) and CSF was always collected 8 – 12 a.m. Plasma samples were collected directly after CSF collection. One aliquot of the CSF sample was used for cell counting, CSF/serum albumin ratio, IgG index, IgG synthesis index and isoelectric focusing for detection of oligoclonal IgG bands, all according to clinical routines performed at the Department of Clinical Chemistry. Within one hour, the remaining CSF was centrifuged (300 x g for 10 min.) and the supernatant was aliquoted and immediately frozen and stored at -70°C until use.

Plasma and CSF samples were analyzed for cytokine and chemokine concentrations with a multiplex bead assay (Milliplex® MAP kits (EMD Millipore Corporation, St. Charles, Missouri, USA)) according to the manufacturer's instructions, except that an additional lower standard point was added to the standard curve. The measurements were performed using Luminex®200™ (Invitrogen, Merelbeke, Belgium). For data acquisition the software program xPONENT 3.1™ (Luminex Corporation, Austin, Texas, USA) was used, and for data analysis MasterPlex® Reader Fit was used. The detection limits were 16 pg/ml for CXCL1, CXCL10 and CCL22, 3.2 pg/ml for CXCL8, 3.9 pg/ml for CXCL13 and 9.8 pg/ml for CCL20. Values below the detection limit were assigned half the value of the detection limit.

CSF NFL concentrations were measured using the NF-light assay according to instructions from the manufacturer (UmanDiagnostics, Umeå, Sweden). CSF NFH concentrations were measured using the Phosphorylated NEFH (Human) ELISA Kit according to instructions from the manufacturer (Abnova, Taipei City, Taiwan). CSF MMP-9 concentrations were measured using the Human MMP-9 Base Kit according to instructions from the manufacturer (Meso Scale Discovery, Rockville, Maryland). CSF GFAP concentrations were measured using an in house ELISA as previously described[21]. CSF CHI3L1 and OPN concentrations

were measured using commercially available ELISAs (R&D Systems, Inc. Minneapolis, MN). The lower limits of quantification for the NFH and MMP-9 assays were 31.2 and 122 pg/mL, respectively. For the other analytes, all samples had concentrations within the quantifiable range of the assay. All measurements were performed in one round of experiments using one batch of reagents by board-certified laboratory technicians who were blinded to clinical information. Intra-assay coefficients of variation were below 15%.

Magnetic resonance imaging

All MR examinations were performed on a 1.5 T Achieva MRI system (Philips Healthcare, Best, The Netherlands). The examination protocol consisted of a: **Axial T2w FLAIR** Turbo spin echo (TSE), field of view (FOV) 230*183*129 mm, acquired resolution (acq.res.) 0.9*1.14*3 mm, (reconstructed resolution (rec.res.) 0.9*0.9*3 mm, repetition time (TR) 6 s, inversion time (TI) 2 s and echo time (TE) 120 ms. **Sagittal T2W FLAIR**, TSE, FOV 230*230*100 mm, acq.res. 0.9*1.12*4, rec.res. 0.9*0.9*4, TR 6 s, TE 120 ms, TI 2 s. **Axial T2W**, TSE, FOV 230*184*129 mm, SENSE 2, acq.res. 0.6*0.77*3 rec.res. 0.45*0.45*3, TR 4.4 s TE 100 ms. **Axial T1W,** spin echo (SE), FOV 230*183*129 mm, SENSE 1.3, acq.res. 0.9*1.12*3 mm, rec.res 0.9*0.9*3 mm, TR 0.6 s and TE 15 ms. The Axial T1w was performed both before and after contrast infection of 0.1 mmol/kg gadolinium (Bayer Pharma AG, Berlin, Germany).

Statistics

Statistical analyses were performed using SPSS for Windows, version 23. Since the measured parameters did not show a Gaussian distribution, non-parametric tests were primarily used. The Mann-Whitney U-test was used to compare two groups and p-values <0.01 were considered to be significant. To compare multiple study groups, the Kruskal–Wallis test was

first performed, and when $p < 0.05$, the Mann–Whitney U-test was used as a post-hoc test and p-values < 0.01 were considered to be significant. All p-values were based on two-tailed statistical tests. The stringent use of p-values was motivated by multiple testing. Non-parametric bivariate correlation analysis (Spearman) was used when investigating possible association between NFL and variables where normal distribution could not be assumed, except between NFL and a dichotomous variable, in which case point-biserial correlation analysis was used. Logistic regression analysis was used when investigating if NFL alone or in combination with other variables could predict disease activity during follow-up. Receiver operating characteristic (ROC) curves were derived from logistic regression to investigate the discriminatory power of NFL between patients and healthy controls and between patients with and without disease activity during follow-up.

Results

Neurodegenerative and neuroinflammatory markers in patients at baseline and in healthy controls

At baseline, when all 44 patients were untreated, CSF levels of NFL, NFH, CHI3L1, MMP-9, CXCL1, CXCL8, CXCL10, CXCL13 and CCL22 were significantly higher in patients than in HC, whereas GFAP and OPN did not differ between patients and HC (Table 3). CCL20 levels in CSF were below the detection level in both patients and HC and CXCL13 in CSF was below the detection level in HC. Patients and HC did not differ in plasma levels in any of the cytokines or chemokines.

Table 3. Neurodegenerative and neuroinflammatory markers in patients at baseline and in healthy controls.

	Patients n=44		Healthy controls n=23		p
	Concentration pg/ml	Measurable level n (%)	Concentration pg/ml	Measurable level n (%)	
CSF NFL	838 (322-2035)	44 (100)	222 (154-291)	23 (100)	**<0.001**
CSF NFH	36 (16-67)	28 (64)	16 (16-16)	5 (22)	**≤0.01**
CSF GFAP	471 (353-646)	44 (100)	442 (344-500)	23 (100)	0.20
CSF CHI3L1	100663 (75972-181132)	44 (100)	73549 (61412-101061)	23 (100)	**≤0.01**
CSF MMP-9	226 (61-611)	31 (70)	61 (61-61)	1 (4)	**<0.001**
CSF OPN	83 (56-122)	44 (100)	67 (44-99)	19 (83)	0.24
CSF CXCL1	16 (8-24)	25 (57)	8 (8-8)	4 (17)	**≤0.01**
CSF CXCL8	25 (22-29)	44 (100)	17 (14-21)	23 (100)	**<0.001**
CSF CXCL10	1053 (806-1996)	44 (100)	540 (450-818)	23 (100)	**<0.001**
CSF CXCL13	5 (2-18)	26 (59)	UD	0 (0)	**<0.001**
CSF CCL20	UD	0 (0)	UD	0 (0)	
CSF CCL22	21 (8-53)	30 (68)	8 (8-8)	1 (4)	**<0.001**
P CXCL1	413 (212-556)	44 (100)	287 (145-548)	23 (100)	0.30
P CXCL8	8 (2-31)	32 (73)	6 (2-19)	16 (70)	0.54
P CXCL10	327 (237-556)	44 (100)	395 (236-553)	23 (100)	0.64
P CXCL13	29 (21-41)	44 (100)	27 (20-32)	13 (56)	0.33
P CCL20	5 (5-5)	9 (20)	5 (5-11)	6 (26)	0.50
P CCL22	980 (766-1248)	44 (100)	1075 (736-1166)	23 (100)	0.82

Concentrations are presented as median and within brackets interquartile range.
UD = Under detection limit.
P-values from Mann-Whitney U test for independent samples, patients compared to healthy controls, significant differences are in bold.

The patients were also stratified by diagnosis (CIS and RRMS) and relapse status, see supplementary text and Tables S1 and S2. Of note is that NFL in CSF at baseline was not

significantly higher in patients with relapse than in relapse-free patients, regardless if time from onset of relapse was within two months before CSF collection (Table S2) or whether it was defined as within one or three months (data not shown). NFL at baseline did not correlate to patient age, EDSS or MSSS, but there was a correlation to mononuclear cells in CSF at baseline (Spearman´s rho 0.41).

There were moderate to strong positive correlations between several of the neurodegenerative and neuroinflammatory markers in CSF at baseline, see supplementary Table S3.

To evaluate how well NFL at baseline discriminated between patients and HC, we performed logistic regression analysis and constructed an ROC curve (Figure 1), from which sensitivity and specificity were calculated. An NFL value ≥453 pg/ml captured 73% of patients and had a specificity of 83%. An NFL value ≥847 pg/ml captured only 50% of patients but had a specificity of 100% in our cohort. With logistic regression analysis based on NFL, 73% of patients and 83% of HC could be correctly classified, with an overall correctly classified proportion of 76%.

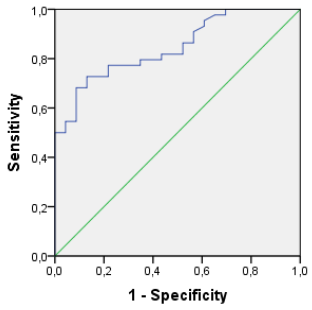

Figure 1. *Receiver operating characteristics (ROC) curve for NFL in CSF regarding separation of CIS and MS patients (n = 44) and HC (n = 23). Area under curve (AUC), representing the proportion of correctly classified patients versus controls, is 0.85, p<0.001.*

Neurodegenerative and neuroinflammatory markers in patients at baseline in relation to disease activity during follow-up

CSF levels of NFL, NFH, OPN, CXCL1, CXCL10, CXCL13 and CCL22 at baseline were significantly lower in patients with NEDA than EDA (evidence of disease activity) at one year (data not shown) and two years of follow-up (Supplementary Table S4) (Figure 2). Baseline CSF levels of CHI3L1, MMP-9 and CXCL8 were significantly lower in patients with NEDA than EDA at one year of follow-up (data not shown). 29 of 42 evaluable patients at 2 year follow-up (1 patient was pregnant and 1 patient left the study) had signs of disease activity and were classified as EDA, whereas 13 had not and were classified as NEDA.

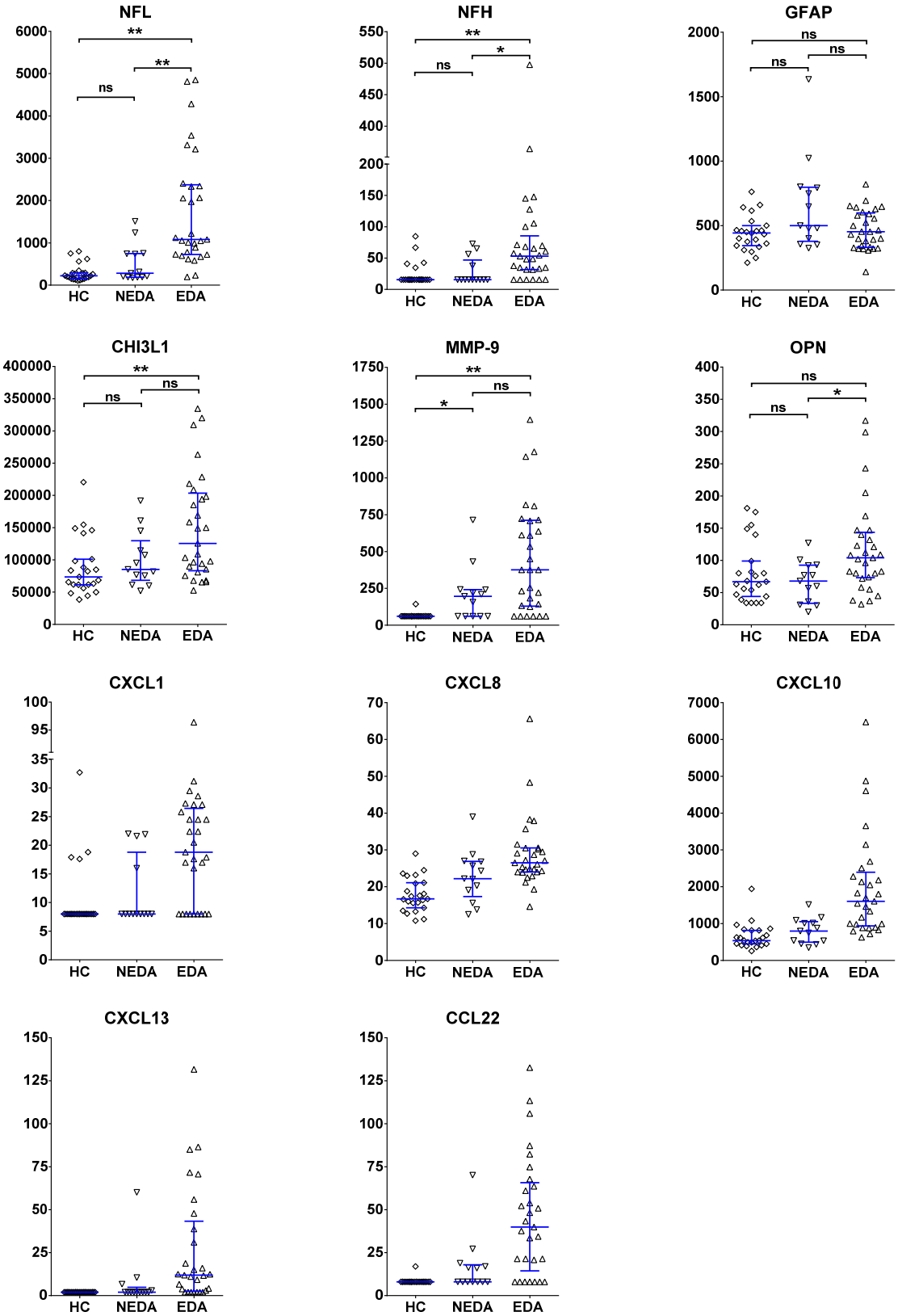
NFL
6000
5000
4000
3000
2000
1000
0
**
**
ns
HC
NEDA
EDA
NFH
550
500
450
400
200
150
100
50
0
**
*
ns
HC
NEDA
EDA
GFAP
2000
1500
1000
500
0
ns
ns
ns
HC
NEDA
EDA
CHI3L1
400000
350000
300000
250000
200000
150000
100000
50000
0
**
ns
ns
HC
NEDA
EDA
MMP-9
1750
1500
1250
1000
750
500
250
0
**
ns
*
HC
NEDA
EDA
OPN
400
350
300
250
200
150
100
50
0
ns
*
ns
HC
NEDA
EDA
CXCL1
100
95
35
30
25
20
15
10
5
0
HC
NEDA
EDA
CXCL8
70
60
50
40
30
20
10
0
HC
NEDA
EDA
CXCL10
7000
6000
5000
4000
3000
2000
1000
0
HC
NEDA
EDA
CXCL13
150
125
100
75
50
25
0
HC
NEDA
EDA
CCL22
150
125
100
75
50
25
0
HC
NEDA
EDA

Figure 2. *Neurodegenerative and neuroinflammatory markers in CSF at baseline in healthy controls and in CIS and MS patients subgrouped by disease activity during follow-up. Lines show median and interquartile range.*
**: P-value from Mann-Whitney U test for independent samples ≤ 0.01.*
***: P-value from Mann-Whitney U test for independent samples <0.001.*
ns: P-value from Mann-Whitney U test for independent samples >0.01.
HC: Healthy controls
NEDA: Patients with no evidence of disease activity during two years of follow-up (no relapses, no brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) and no disability worsening)
EDA: Patients with evidence of disease activity during two years of follow-up (relapses, brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) or disability worsening)

From an ROC curve (Figure 3) sensitivity and specificity for NFL levels at baseline in relation to disease activity at two year follow-up were calculated. With a cut off value of ≥450 pg/ml for NFL, 97% of patients that showed disease activity in this cohort were identified, but the specificity was only 62%, while a cut-off value of ≥749 pg/ml gave a sensitivity of 72% and a specificity of 77%.

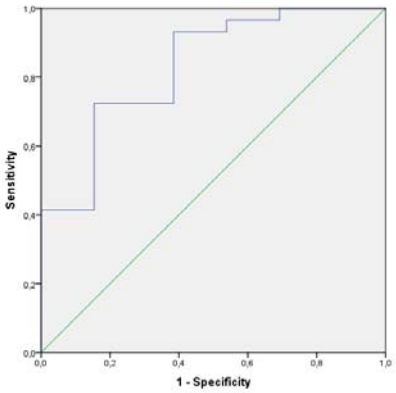

Figure 3. *Reciever operating characteristics (ROC) curve for NFL in CSF at baseline and evidence of disease activity in patients during two years of follow-up. Area under curve (AUC), representing the proportion of patients correctly classified as with or without disease activity, is 0.83, p=0.001.*

The predictive value of each one of the neurodegenerative and neuroinflammatory markers in CSF at baseline with regard to disease activity during follow-up was evaluated with logistic regression, as presented in supplementary Table S5. We found that NFL performed best at predicting disease activity status at two years, classifying 93% of EDA patients and 62% of

NEDA patients correctly. The overall correctly classified proportion of patients was 83% (Table 6). We then combined NFL with the other markers one by one. The only combinations that yielded a slightly higher overall percentage of correctly classified patients were NFL+OPN and NFL+CCL22. Combining NFL with both CCL22 and OPN correctly classified 93 % of patients with disease activity and 85 % of patients without disease activity, with an overall proportion of correctly classified patients of 91 % (Table 6). Of note is that baseline number of T2 hyperintense lesions in the brain was inferior to NFL at predicting disease activity and that combining NFL with number of T2 hyperintense lesions did not improve the model compared to NFL alone. NFL at baseline did not correlate with total number of T2 lesions in the brain at baseline or to new lesions in brain MRI during two years of follow-up.

Table 6. Results from logistic regression analysis showing the ability of markers at baseline to predict disease activity during two years of follow-up.

	Correctly predicted EDA (% out of 29 patients)	Correctly predicted NEDA (% out of 13 patients)	Overall correctly classified (% out of 42 patients)
NFL	93	62	83
NFL + NFH	93	62	83
NFL + GFAP	93	62	83
NFL + MMP-9	90	62	81
NFL + CHI3L1	93	62	83
NFL + OPN	93	69	86
NFL + CXCL1	83	77	81
NFL + CXCL8	90	54	79
NFL + CXCL10	83	69	79
NFL + CXCL13	93	62	83
NFL + CCL22	86	85	86
NFL + CCL22 + OPN	93	85	91
T2	100	8	71
EDSS	100	0	69
NFL + T2	90	62	81

EDA: Patients with evidence of disease activity during two years of follow-up (relapses, brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) or disability worsening)
NEDA: Patients with no evidence of disease activity during two years of follow-up (no relapses, no brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) and no disability worsening)
T2: Total number of T2 hyperintense lesions in brain MRI at baseline
EDSS: Expanded disability status scale at baseline

Discussion

In the present study, we examined a prospective longitudinally followed cohort of patients with CIS and newly diagnosed MS to evaluate a broad set of both established and non-established markers of neurodegeneration and neuroinflammation as potential prognostic markers of disease activity. Importantly, we show that NFL in CSF, with high accuracy predicted whether patients would show signs of disease activity during two years of follow-up. Combining NFL with CSF-levels of OPN and CCL22 increased the overall correctly classified patients from 83% to 93%. Considering the heterogeneity in MS, prediction of disease activity using fluid biomarkers whould be of great value in decision of treatment strategy. Currently, clinical and radiological parameters are used for such decisions, but it should be noted that in the present study, T2 lesions at baseline did not add to the prediction of NEDA status.

CSF levels of NFL in CIS patients have been reported as an independent prognostic marker for conversion from CIS to MS[22], to correlate with multiple sclerosis severity score during long term follow-up in MS[5] and to decrease on treatment[7]. To our knowledge this is the first study to demonstrate that NFL at baseline is a strong predictor of disease activity in the comprehensive form of NEDA-3 status after two years of follow-up in CIS and newly diagnosed MS. Thus, based on our and previous data, NFL in CSF emerges as a useful prognostic biomarker of disease activity. It has been shown that exposure to room temperature or repetitive thawing does not influence measurement of NFL concentrations[23], which is of great value for a potential biomarker. NFL in CSF at baseline did not correlate with patient age at baseline in our study and we believe that this is due to the relatively low age of most patients in the cohort, since otherwise age should be taken into account when evaluating NFL levels[24]. We also evaluated several other suggested CSF biomarkers related to the process of

CNS tissue damage (NFH, GFAP, CHI3L1 and MMP-9), but in our setting, NFL in CSF stood out as the most useful prognostic biomarker of NEDA, with the highest discriminative capability also in relation to markers of neuroinflammation.

After finding NFL at baseline as the sole marker with the highest potential to distinguish patients with disease activity after two years, we tested if combinations with markers of CSF inflammation could add to the discriminative power. OPN and CCL22 then emerged as useful since they, both alone and in combination, increased the proportion correctly predicted patients according to NEDA status at follow-up. In previous studies of OPN in MS, it was associated with ongoing disease activity[15], but could also predict time to conversion from optic neuritis to MS, although it was outperformed by CHI3L1 in that setting[25]. In our study, OPN levels in CSF did not differ between patients and controls at baseline, and it was not increased at time of relapse. Still, it added to the prediction of disease activity after two years. OPN is a pleiotropic protein involved in several processes and it is not known by which mechanism it may contribute as a predictive biomarker. Interestingly, OPN was recently shown to be part of a module of dysregulated genes in activation of T helper cells in MS, indicating the involvement of T helper cell-associated mechanisms in the ability of OPN to contribute as a prognostic biomarker in MS. Although chemokines have been increasingly recognized as potential biomarkers, there has been little attention to CCL22 in MS, probably because it is induced by IL4 and IL-13 and involved in the recruitment of Th2 cells and in Th2 immunity[26] which has not been considered relevant in MS. However, when measured in CSF of MS patients, CCL22 levels were increased in comparison with controls[27, 28] and furthermore, CCL22 levels decreased after natalizumab treatment[27]. The present study is the first to demonstrate that higher levels at baseline are associated with disease activity during follow-up.

NFL, OPN and CCL22 levels did not differ between patients in relapse and patients not in relapse at the time of CSF collection in our study. As for NFL[22, 29] and OPN[15, 30], data on levels in relation to relapse status are inconsistent, possibly because of differences in sample size and selection. Our finding of similar levels of NFL, OPN and CCL22 irrespective of relapse implies positive qualities as clinical biomarkers.

A strength of this study is that the control group consisted of sex- and age-matched healthy individuals, not patients with "other neurological diseases" or patients examined due to a suspicion of neurological disease but without clinical findings indicating any CNS disorder. Another strength of this study is that the patient group consisted of are well-characterized MS patients examined and thoroughly followed-up by the same neurologist in a standardized way. A limitation of this study is the low sample size when stratifying into subgroups. Still, data were clear and consistent and we used a stringent statistical approach to compensate for multiple testing. The present study was designed to detect short-term effects, *i.e.* disease activity after two years, and future studies will address biomarkers in relation to long-term outcome including development of atrophy. Some of the analytes were not detected in CSF, which precluded evaluation as biomarker. Of note, none of the inflammatory markers in plasma were increased in patients versus controls and could not predict disease activity after two years, showing that CSF sampling is necessary to mirror the CNS inflammation. Regarding neuronal markers, it will be interesting to see whether super-sensitive assays of serum levels[31] can add as biomarkers in MS.

In conclusion, we show that NFL in CSF at diagnostic lumbar puncture in CIS and MS is able to predict disease activity as measured by NEDA status after two years. In combination with

previous studies, there is evidence to include NFL in CSF as a biomarker in clinical routine. We also show that combining NFL in CSF with markers of neuroinflammation seem to increase prognostic ability, which should be confirmed in larger studies. The use of biomarkers will be crucial in clinical practice as we want to individualize treatment early in the disease course.

Acknowledgements

We thank Patrik Fägerstam at the Department of Radiology, Linköping University Hospital, for systematically reviewing MRI scans and detailed MRI reports. We thank the staff at the Department of Neurology, Linköping University Hospital, for their help in collecting samples. We thank the staff at Center for Medical Image Science and Visualization (CMIV), Linköping University, for their help in carrying out MRI examinations.

Funding

The study was supported by grants from the Swedish Research Council (K2013-61X-22310-01-4). KB is supported by the Torsten Söderberg foundation and the Royal Academy of Sciences, Sweden. HZ is a Wallenberg Academy Fellow.

References

1. Hartel M, Kluczewska E, Pierzchala K, Adamczyk-Sowa M and Karpe J. What you cannot get from routine MRI of MS patient and why - The growing need for atrophy assessment and seeing beyond the plaque. *Neurol Neurochir Pol*. 2016; 50: 123-30.
2. Fitzner B, Hecker M and Zettl UK. Molecular biomarkers in cerebrospinal fluid of multiple sclerosis patients. *Autoimmun Rev*. 2015; 14: 903-13.
3. Housley WJ, Pitt D and Hafler DA. Biomarkers in multiple sclerosis. *Clin Immunol*. 2015.
4. Harris VK and Sadiq SA. Biomarkers of therapeutic response in multiple sclerosis: current status. *Mol Diagn Ther*. 2014; 18: 605-17.
5. Salzer J, Svenningsson A and Sundstrom P. Neurofilament light as a prognostic marker in multiple sclerosis. *Mult Scler*. 2010; 16: 287-92.
6. Kuhle J, Disanto G, Lorscheider J, et al. Fingolimod and CSF neurofilament light chain levels in relapsing-remitting multiple sclerosis. *Neurology*. 2015; 84: 1639-43.
7. Gunnarsson M, Malmestrom C, Axelsson M, et al. Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab. *Ann Neurol*. 2011; 69: 83-9.
8. Salzer J. The only certain measure of the effectiveness of multiple sclerosis therapy is cerebrospinal neurofilament level - YES. *Mult Scler*. 2015; 21: 1239-40.
9. Arrambide G, Espejo C and Tintore M. The only certain measure of the effectiveness of multiple sclerosis therapy is cerebrospinal neurofilament level-NO. *Mult Scler*. 2015; 21: 1240-2.
10. Romme Christensen J, Bornsen L, Khademi M, et al. CSF inflammation and axonal damage are increased and correlate in progressive multiple sclerosis. *Mult Scler*. 2013; 19: 877-84.
11. Teunissen CE and Khalil M. Neurofilaments as biomarkers in multiple sclerosis. *Mult Scler*. 2012; 18: 552-6.
12. Sellebjerg F and Sorensen TL. Chemokines and matrix metalloproteinase-9 in leukocyte recruitment to the central nervous system. *Brain Res Bull*. 2003; 61: 347-55.
13. Hinsinger G, Galeotti N, Nabholz N, et al. Chitinase 3-like proteins as diagnostic and prognostic biomarkers of multiple sclerosis. *Mult Scler*. 2015; 21: 1251-61.
14. Chabas D, Baranzini SE, Mitchell D, et al. The influence of the proinflammatory cytokine, osteopontin, on autoimmune demyelinating disease. *Science*. 2001; 294: 1731-5.
15. Bornsen L, Khademi M, Olsson T, Sorensen PS and Sellebjerg F. Osteopontin concentrations are increased in cerebrospinal fluid during attacks of multiple sclerosis. *Mult Scler*. 2011; 17: 32-42.
16. Krumbholz M, Theil D, Cepok S, et al. Chemokines in multiple sclerosis: CXCL12 and CXCL13 up-regulation is differentially linked to CNS immune cell recruitment. *Brain*. 2006; 129: 200-11.
17. Khademi M, Kockum I, Andersson ML, et al. Cerebrospinal fluid CXCL13 in multiple sclerosis: a suggestive prognostic marker for the disease course. *Mult Scler*. 2011; 17: 335-43.
18. Henningsson AJ, Tjernberg I, Malmvall BE, Forsberg P and Ernerudh J. Indications of Th1 and Th17 responses in cerebrospinal fluid from patients with Lyme neuroborreliosis: a large retrospective study. *J Neuroinflammation*. 2011; 8: 36.
19. Giovannoni G, Turner B, Gnanapavan S, Offiah C, Schmierer K and Marta M. Is it time to target no evident disease activity (NEDA) in multiple sclerosis? *Mult Scler Relat Disord*. 2015; 4: 329-33.

20. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol*. 2011; 69: 292-302.
21. Rosengren LE, Wikkelso C and Hagberg L. A sensitive ELISA for glial fibrillary acidic protein: application in CSF of adults. *J Neurosci Methods*. 1994; 51: 197-204.
22. Martinez MA, Olsson B, Bau L, et al. Glial and neuronal markers in cerebrospinal fluid predict progression in multiple sclerosis. *Mult Scler*. 2015; 21: 550-61.
23. Kuhle J, Plattner K, Bestwick JP, et al. A comparative study of CSF neurofilament light and heavy chain protein in MS. *Mult Scler*. 2013; 19: 1597-603.
24. Vagberg M, Norgren N, Dring A, et al. Levels and Age Dependency of Neurofilament Light and Glial Fibrillary Acidic Protein in Healthy Individuals and Their Relation to the Brain Parenchymal Fraction. *PLoS One*. 2015; 10: e0135886.
25. Modvig S, Degn M, Roed H, et al. Cerebrospinal fluid levels of chitinase 3-like 1 and neurofilament light chain predict multiple sclerosis development and disability after optic neuritis. *Mult Scler*. 2015; 21: 1761-70.
26. Yamashita U and Kuroda E. Regulation of macrophage-derived chemokine (MDC, CCL22) production. *Crit Rev Immunol*. 2002; 22: 105-14.
27. Mellergard J, Edstrom M, Vrethem M, Ernerudh J and Dahle C. Natalizumab treatment in multiple sclerosis: marked decline of chemokines and cytokines in cerebrospinal fluid. *Mult Scler*. 2010; 16: 208-17.
28. Burman J, Svensson E, Fransson M, et al. The cerebrospinal fluid cytokine signature of multiple sclerosis: a homogenous response that does not conform to the Th1/Th2/Th17 convention. *J Neuroimmunol*. 2014; 277: 153-9.
29. Kuhle J, Barro C, Disanto G, et al. Serum neurofilament light chain in early relapsing remitting MS is increased and correlates with CSF levels and with MRI measures of disease severity. *Mult Scler*. 2016.
30. Braitch M, Nunan R, Niepel G, Edwards LJ and Constantinescu CS. Increased osteopontin levels in the cerebrospinal fluid of patients with multiple sclerosis. *Arch Neurol*. 2008; 65: 633-5.
31. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. *Clin Chem Lab Med*. 2016.

Supplemental materials

Patients stratified by diagnosis

CSF levels of NFL, NFH, GFAP, MMP-9, CHI3L1, OPN, CXCL1, CXCL8, CXCL10, CCL13 and CCL22 in patients at baseline did not differ with regard to diagnosis CIS (n = 19), RRMS (n = 22) or PPMS (n = 3). Compared to HC (n = 23), CSF levels of NFL, MMP-9, CXCL8 and CXCL13 were higher in all patient subgroups, NFH, CXCL10 and CCL22 were higher in the CIS group and the RRMS group, CHI3L1 and CXCL1 were higher in the RRMS group and GFAP and OPN did not differ from HC in any patient group. Detailed data are presented in supplemental Table S1. Plasma levels of CXCL1, CXCL8, CXCL10, CXCL13, CCL20 and CCL22 at baseline did not differ between patient groups or between patient groups and HC (data not shown).

Patients stratified by relapse status

CSF levels of NFH were higher in patients with relapse at the time of baseline CSF collection, whereas the other markers in CSF did not differ between patients with regard to relapse status. CSF levels of NFL, MMP-9, CXCL8, CLCL10, CXCL13 and CCL22 at baseline were significantly higher in patients than in HC, regardless of relapse status. Levels of NFH and CXCL1 in CSF at baseline were higher in patients in relapse than in HC, but not in relapse-free patients. Levels of CHI3L1 were higher in relapse-free patients than in HC (p 0.01) and showed a tendency to be higher in patients in relapse than in HC (p 0.03). Levels of GFAP and OPN in patients did not differ from HC, regardless of relapse status. CSF data for patients, subgrouped by relapse status, with time from onset of relapse within two months, and healthy controls are presented in supplemental Table S2.

Table S1. Neurodegenerative and neuroinflammatory markers in CSF in healthy controls and in patients at baseline, subgrouped by diagnosis.

	CIS n=19	RRMS n=22	PPMS n=3	HC n=23	p MW (HC-CIS, HC-RRMS, HC-PPMS)
NFL	724 (207-2332)	998 (650-1447)	743 (616-1121)	222 (154-291)	**≤0.01, <0.001, ≤0.01**
NFH	48 (16-73)	33 (16-56)	65 (33-71)	16 (16-16)	**≤0.01**, **≤0.01**, 0.03
GFAP	453 (324-642)	489 (374-627)	749 (453-820)	442 (344-500)	0.71, 0.18, 0.06
CHI3L1	85222 (75436-114464)	117224 (81168-193442)	149885 (144979-194116)	73549 (61412-101061)	0.21, **≤0.01**, 0.02
MMP-9	143 (61-376)	404 (110-717)	255 (196-534)	61 (61-61)	**<0.001, <0.001, <0.001**
OPN	68 (36-127)	98 (72-127)	77 (74-121)	67 (44-99)	0.76, 0.04, 0.44
CXCL1	8 (8-18)	21 (8-25)	8 (8-27)	8 (8-8)	0.11, **<0.001**, 0.65
CXCL8	24 (22-27)	27 (23-31)	29 (22-30)	17 (14-21)	**≤0.01, <0.001, ≤0.01**
CXCL10	885 (750-1336)	1533 (886-2207)	980 (537-1606)	540 (450-818)	**≤0.01, <0.001,** 0.16
CXCL13	2 (2-12)	11 (2-35)	9 (3-71)	UD	**≤0.01, <0.001, <0.001**
CCL22	17 (8-27)	44 (14-71)	34 (8-52)	8 (8-8)	**<0.001, <0.001,** 0.78

Concentrations are given in pg/ml and presented as median and within brackets interquartile range, except for PPMS where minimum value and maximum values are given within brackets.
UD = Under detection limit.
p MW: P-values from Mann-Whitney U test for independent samples, comparing healthy controls to patients with CIS, RRMS and PPMS, respectively. Please note that to compare data from patients with different diagnoses and healthy controls, the Kruskal–Wallis test was first performed. When significant differences were found (p-values <0.05 in the Kruskal–Wallis test), the Mann–Whitney U-test was used as a post-hoc test to compare groups. P-values <0.01 in the Mann–Whitney U-test were considered to be significant.
Please also note that p-values from Kruskal-Wallis test for independent samples comparing patients with diagnosis CIS, RRMS and PPMS were all >0.05 and therefore no Mann–Whitney U-test was used to compare these groups.
HC-CIS = Healthy controls compared to patients with diagnosis clinically isolated syndrome
HC-RRMS = Healthy controls compared to patients with diagnosis relapsing remitting multiple sclerosis
HC-PPMS = Healthy controls compared to patients with diagnosis primary progressive multiple sclerosis

Table S2. Neurodegenerative and neuroinflammatory markers in CSF in healthy controls and in patients, subgrouped by relapse status at time of baseline CSF collection.

	Relapse within two months n=23	Relapse free n=21	p MW[a]	Healthy controls n=23	p MW[b] r-HC, rf-HC
NFL	1049 (724-2332)	697 (258-1181)	0.13	222 (154-291)	**<0.001, <0.001**
NFH	58 (32-105)	16 (16-51)	**≤0.01**	16 (16-16)	**<0.001**, 0.09
GFAP	499 (381-642)	453 (334-698)	0.85	442 (344-500)	0.17, 0.42
CHI3L1	97919	103407	0.86	73549	0.03, **≤0.01**

	(75647-185171)	(76192-180416)		(61412-101061)	
MMP-9	376 (61-612)	219 (61-622)	0.77	61 (61-61)	**<0.001, <0.001**
OPN	83 (58-132)	92 (44-110)	0.59	67 (44-99)	0.20, 0.46
CXCL1	18 (8-26)	8 (8-22)	0.22	8 (8-8)	**0.001,** 0.04
CXCL8	26 (24-30)	24 (21-29)	0.13	17 (14-21)	**<0.001, <0.001**
CXCL10	1336 (797-2280)	980 (790-1829)	0.25	540 (450-818)	**<0.001, <0.001**
CXCL13	7 (2-48)	4 (2-16)	0.71	UD	**<0.001, <0.001**
CCL22	21 (8-61)	21 (8-50)	0.42	8 (8-8)	**<0.001, <0.001**

Concentrations are given in pg/ml and presented as median and within brackets interquartile range.
UD = Under detection limit.
p MW[a]: P-values from Mann-Whitney U test for independent samples, comparing patients in relapse and relapse free patients.
p MW[b]: P-values from Mann-Whitney U test for independent samples, comparing healthy controls to patients in relapse and relapse free patients, respectively.
r-HC = Patients in relapse compared to healthy controls
rf-HC= Relapse free patients compared to healthy controls

Table S3. Map of correlation analyses between neurodegenerative and neuroinflammatory markers in CSF at baseline.

	NFL	NFH	GFAP	CHI3L1	MMP-9	OPN	CXCL1	CXCL8	CXCL10	CXCL13	CCL22
NFL	1.00	**0.57****	**0.44***	**0.60****	**0.49***	**0.43***	**0.40***	**0.47***	**0.58****	**0.54****	**0.50***
NFH		1.00	**0.51****	0.34	0.18	0.35	0.08	0.28	0.32	0.36	0.11
GFAP			1.00	**0.56****	0.18	**0.41***	0.15	0.25	0.22	0.34	0.12
CHI3L1				1.00	**0.44***	**0.58****	**0.50***	**0.51****	**0.57****	**0.47***	**0.43***
MMP-9					1.00	**0.40***	**0.48***	**0.50****	**0.75****	**0.82****	**0.74****
OPN						1.00	**0.40***	0.38	**0.48***	**0.44***	0.31
CXCL1							1.00	**0.64****	**0.62****	0.34	**0.51****
CXCL8								1.00	**0.67****	0.33	**0.40***
CXCL10									1.00	**0.62****	**0.68****
CXCL13										1.00	**0.74****
CCL22											1.00

Magnitude of correlation coefficients (r) are shown, p-values according to bivariate non-parametric correlation analyses (Spearman).
* $p < 0.01$
** $p < 0.001$

Table S4. Neurodegenerative and neuroinflammatory markers in CSF in healthy controls and at baseline in patients with NEDA/EDA for two years of follow-up.

	NEDA n=13	EDA n=29	p[a]	HC n=23	p[b]
NFL	283 (190-749)	1081 (726-2374)	**<0.001**	222 (154-291)	0.11, **<0.001**
NFH	16	53	**≤0.01**	16	0.65, **<0.001**

	(16-47)	(32-86)		(16-16)	
GFAP	352 (322-528)	499 (395-650)	0.03	442 (344-500)	0.52, 0.05
CHI3L1	85222 (68416-129721)	125532 (83201-203501)	0.06	73549 (61412-101061)	0.33, **0.001**
MMP-9	196 (61-242)	376 (130-712)	0.05	61 (61-61)	**≤0.01, <0.001**
OPN	68 (34-92)	104 (73-144)	**≤0.01**	67 (44-99)	0.58, 0.03
CXCL1	8 (8-19)	19 (8-26)	**≤0.01**	8 (8-8)	0.49, **<0.001**
CXCL8	22 (17-27)	26 (24-30)	0.02	17 (14-21)	0.04, **<0.001**
CXCL10	797 (495-1053)	1606 (939-2394)	**<0.001**	540 (450-818)	0.15, **<0.001**
CXCL13	2 (2-5)	12 (3-43)	**≤0.01**	ND	0.13, **<0.001**
CCL22	8 (8-18)	40 (14-66)	**≤0.01**	8 (8-8)	0.04, **<0.001**

Concentrations are given in pg/ml and presented as median and within brackets interquartile range.
NEDA: No evidence of disease activity, defined as not having had any new relapse, not having progressed on EDSS and not having had any new T2 lesions or gadolinium enhancing T1 lesions for two years.
EDA: Patients that do not fulfil criteria for NEDA as specified above.
p^a: *P-values from Mann-Whitney U test for independent samples, comparing NEDA and EDA patients.*
p^b: *P-values from Mann-Whitney U test for independent samples, comparing first NEDA patients and healthy controls and then EDA patients and healthy controls.*

Table S5. Results from logistic regression analysis showing the ability of markers in CSF at baseline to predict disease activity during two years of follow-up.

	Correctly predicted EDA (% out of 29 patients)	Correctly predicted NEDA (% out of 13 patients)	Overall correctly classified (% out of 42 patients)
NFL	93	62	83
NFH	79	69	76
GFAP	100	0	69
CHI3L1	97	8	69
MMP-9	100	0	69
OPN	86	31	69
CXCL1	69	69	69
CXCL8	97	23	74
CXCL10	86	54	76
CXCL13	100	0	69
CCL22	76	54	69

EDA: Patients with evidence of disease activity during two years of follow-up (relapses, brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) or disability worsening)
NEDA: Patients with no evidence of disease activity during two years of follow-up (no relapses, no brain MRI activity (new or enlarging T2 lesions or Gd-enhancing lesions) and no disability worsening)
T2: Total number of T2 hyperintense lesions in brain MRI at baseline
EDSS: Expanded disability status scale at baseline