**ORIGINAL COMMUNICATION**

# Diagnostic utility of neurofilament markers for MND is limited in restricted disease phenotype and for differentiation from compressive myeloradiculopathies

Daniel Baumgartner[1] · Radim Mazanec[1] · Jitka Hanzalová[2]

## Abstract

Misdiagnosis is frequent in early motor neuron disease (MND), typically compressive radiculopathy, or in patients with restricted MND phenotype. In this retrospective, single tertiary centre study, we measured levels of neurofilament light (NfL) and phosphorylated neurofilament heavy (p-NfH) chain in cerebrospinal fluid (CSF) and of p-NfH in serum with commercially available ELISA kits and assessed their respective diagnostic performance as a marker of MND. The entire study population ($n = 164$) comprised 71 MND patients, 30 patients with compressive myelo- or radiculopathy, and 63 disease controls (DC). Among MND patients, we specified subgroups with only lower motoneuron involvement (MND-LMN, $n = 15$) and with confounding nerve roots or spinal cord compression (MND-C, $n = 18$), representing clinical diagnostic pitfalls. MND-LMN displayed significantly lower CSF NfL ($p = 0.003$) and p-NFH ($p = 0.017$), but not serum p-NfH ($p = 0.347$) levels compared to other MND patients ($n = 56$). The discriminative ability (area under the curve—AUC) of both CSF Nfs towards all MND patients was comparable to each other but significantly higher than that of p-NfH in serum ($ps < 0.001$). AUC of both CSF Nfs between MND-LMN and DC and also between MND-C and myelo-/radiculopathies were reduced, as compared to AUC between other MND and DC or myelo-/radiculopathies, respectively. Our results suggest that both Nfs in CSF represent a reliable diagnostic marker in a general MND population, fulfilling Awaji criteria. As for diagnostic pitfalls, and also for p-NfH in serum, their discriminative ability and, therefore, clinical utility appears to be limited.

**Keywords** Neurofilament · Biomarker · Cerebrospinal fluid · Amyotrophic lateral sclerosis · Motor neuron disease

## Introduction

Motoneuron disease (MND) is a heterogeneous group of fatal neurodegenerative diseases affecting adults. The clinical [1] and pathophysiological [2] heterogeneity of this group complicates early diagnosis and personalized prognosis for a particular patient. Misdiagnosis is relatively common in the early stages of the disease. The spinal form of MND is often mistaken with compressive radiculopathies or myelopathy, which have the highest incidence in a similar age group as MND [3]. As a result, about 13% of patients undergo unnecessary surgical treatment [4], which, in addition, may lead to an acceleration of MND progression [5]. The differential diagnosis is relatively complicated if MND involvement is limited to lower motoneurons (LMN) [6], and in such cases, the diagnostic uncertainty can lead to costly and potentially risky therapies (immunosuppression) without benefit to the patient [7]. The search for new biomarkers of MND, enabling a timely and reliable diagnosis and personalized prognosis, therefore became one of the main priorities of this research area.

Over the last 20 years, several studies have consistently shown increased levels of neurofilament light (NfL) and phosphorylated heavy chain (p-NfH) in cerebrospinal fluid (CSF) and serum of MND patients. Accordingly, the role of these proteins as a potential diagnostic biomarker for MND has been proposed [8] (among other possible roles, as their utility as prognostic and potentially pharmacodynamic biomarkers has been parallelly explored as well).

✉ Daniel Baumgartner
daniel.baumgartner@fnmotol.cz

1 Neuromuscular Centre, Department of Neurology, Charles University, 2nd Faculty of Medicine and University Hospital Motol, V Uvalu 84, 150 06 Prague, Czech Republic

2 Department of Immunology, Charles University, 2nd Faculty of Medicine and University Hospital Motol, V Uvalu 84, 150 06 Prague, Czech Republic

These neuron-specific cytoskeletal proteins from the family of intermediate filaments are predominantly expressed in large, myelinated axons and are classified according to the molecular weight of their subunits as light (NfL), medium (NfM), and heavy (NfH) chains, forming a triplet (within a tetramer structure) [9]. The approximate intracellular ratio of NfL/NfM/NfH is 5:3:1 [10]. In CSF and serum, increased levels of NfL [11, 12], as well as p-NfH [13–15], have been reported in several other neurological disorders (multiple sclerosis, neuroinflammation, infarction, traumatic lesion, and other neurodegenerative diseases), suggesting that neurofilaments are non-specific markers of neuroaxonal damage. When comparing the levels of Nfs in MND and these disorders, the discriminative ability of both Nfs towards MND seems to be significantly lower [15]. Furthermore, some studies have demonstrated significantly increased levels of neurofilaments in compressive myelopathies and radiculopathies as well [16, 17]. So far, other studies have not particularly focused on the discriminative ability of neurofilaments between compressive root lesions and MND, and the group of myelo-/radiculopathies has been underrepresented [15, 18–20]. Furthermore, although several studies have shown significantly lower levels of NfL and p-NfH in CSF and serum in variants of MND with LMN-restricted involvement as compared to other MND phenotypes [18, 21, 22], data on the discriminative ability of Nfs in these variants, when compared to MND mimics, are lacking. Such data would be very meaningful from a clinical diagnostic perspective.

Therefore, in this retrospective, monocentric study, we evaluated the discriminative ability of CSF NfL, p-NfH, and p-NfH in serum in MND patients and patients with other neurological disorders relevant to the differential diagnosis of MND. In detail, we compared the levels of these proteins (1) between MND and compressive myelo-/radiculopathies and (2) between MND patients with isolated LMN involvement and patients with other clinically relevant conditions. Our secondary goal was to determine the clinical parameters in MND influencing levels of Nfs in serum and CSF.

## Methods

### Patients

Between 2012 and 2019, a total of 164 patients undergoing CSF examination as a part of the routine diagnostic work-up in our neurological department were enrolled in a retrospective manner. The study has been approved by our institutional Ethics Committee.

A total of 71 patients with a diagnosis of MND established at our neuromuscular reference centre according to El Escorial criteria revision from 2015 (patients with only LMN-isolated lesion in $\geq 2$ regions were also included) [23] were enrolled (Table 1). First symptoms (bulbar, weakness, fasciculations) as referred by the patients, and ALSFRS-R score closest to the time of lumbar puncture (LP) have been ascertained from medical records. The time from onset (of MND) to LP has been determined. The disease progression rate ($\Delta$FS) was calculated, as described elsewhere [24].

Further, the MND population was divided into different categories according to three criteria. The first category was composed of MND patients for whom the differential diagnosis of compressive myelopathy or radiculopathy had been considered based on clinical-radiological correlation

**Table 1** Comparison of characteristics between MND ($n=71$) and myelo-/radiculopathies ($n=30$) and between MND and disease controls ($n=63$)

|  | MND ($n=71$) | Myelo-/radiculopathies ($n=30$) | Disease controls ($n=63$) | $p^A$ | $p^B$ |
|---|---|---|---|---|---|
| $N$ (CSF/serum) | 71/62 | 30/26 | 63/36 |  |  |
| Females | 34 (47.9%) | 11 (36.7%) | 25 (39.7%) | 0.300 | 0.340 |
| Age (years) | 59 (27; 81) | 63.5 (33; 83) | 58 (21; 81) | 0.218 | 0.675 |
| p-NfH (pg/ml), CSF | 2 660 (270; 40 000) | 676 (155; 33 170) | 380 (30; 2 560) | **< 0.001** | **< 0.001** |
| p-NfH (pg/ml), serum | 59 (8; 1 455) | 27 (8; 1 883) | 25 (9; 230) | **0.038** | **0.010** |
| NfL (pg/ml) | 4 730 (446; 36 300) | 1 470 (350; 10 000) | 670 (90; 2 100) | **< 0.001** | **< 0.001** |
| Qalb | 5.8 (3.0; 27.3) | 9.6 (3.4; 19.4) | 6.7 (2.0; 55.0) | **< 0.001** | **0.011** |
| Total protein (mg/l) | 420 (215; 1 740) | 623 (230; 1 200) | 480 (245; 3 200) | **< 0.001** | **0.009** |
| Time to LP (months) | 15 (2; 39) | – | – |  |  |

*MND* motor neuron disease, *NfL* neurofilament light chain, *p-NfH* phosphorylated neurofilament heavy chain, *CSF* cerebrospinal fluid, *Qalb* albumin quotient

Absolute frequencies (percentages) are presented for categorical variables. Medians (minimum; maximum) are presented continuous variables

$p^A$—testing differences in parameters between MND and myelo-/radiculopathies with Mann–Whitney test (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

$p^B$—testing differences in parameters between MND and disease controls with Mann–Whitney (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

or who had been misdiagnosed before referral to our centre (MND-C, $n = 18$, Table 2). The criteria of a suspicious radiological finding were: (a) for the cervical region, a finding of myelopathy or cervical compression index (CCI) grade 3 [25] and/or $\geq 2$ levels with foraminal stenosis, corresponding to the documented clinical deficit; (b) for the lumbar region, a dural sac cross-sectional area $< 100$ mm$^2$ [26] and/or $\geq 2$ levels with foraminal stenosis (Table 3). The diagnosis of MND was confirmed by EMG (evidence of LMN dysfunction according to Awaji criteria found in regions without

compressive lesion) and/or based on further clinical progression of the disease. The rest of the MND patients did not fulfil these criteria ($n = 53$, Table 2). The second category was composed of patients with LMN-isolated involvement (in $\geq 2$ regions) at the time of LP (MND-LMN, $n = 15$, Table 4). The other two subgroups comprised patients fulfilling the Awaji criteria of at least a possible ALS, either with spinal (MND-ALS, $n = 36$), or bulbar (MND-B, $n = 20$) involvement at the onset. Third category was composed of MND patients divided according to their progression rate

**Table 2** Comparison of characteristics between MND-C and other MNDs, between MND-C and myelo-/radiculopathies and between other MNDs and myelo-/radiculopathies

|  | MND-C ($n = 18$) | Other MNDs ($n = 53$) | Myelo-/radiculopathies ($n = 30$) | $p^A$ | $p^B$ | $p^C$ |
|---|---|---|---|---|---|---|
| Females | 6 (33.3%) | 28 (52.8%) | 11 (36.7%) | 0.153 | 0.815 | 0.156 |
| Age (years) | 66 (9) | 56 (12) | 64 (33; 83) | **0.004** | 0.449 | 0.057 |
| Time to LP (months) | 16 (2; 38) | 15 (4; 39) | N/A | 0.706 | | |
| Δ FS | 0.475 (0.050; 1.375) | 0.600 (0.050; 4.000) | N/A | 0.109 | | |
| p-NfH (pg/ml), CSF | 2225 (535; 16 800) | 2660 (270; 40 000) | 676 (155; 33 170) | 0.658 | **0.007** | **< 0.001** |
| p-NfH (pg/ml), serum | 67 (8; 456) | 55 (9; 1 455) | 27 (8; 1 883) | 0.687 | 0.164 | **0.041** |
| NfL (pg/ml) | 4085 (480; 10 000) | 5790 (446; 36 300) | 1470 (350; 10,000) | **0.050** | **0.048** | **< 0.001** |

Absolute frequencies (percentages) are presented for categorical variables. Medians (minimum; maximum) are presented for non-normal continuous variables; mean (standard deviation) for normal variables. For testing, non-normal variables were transformed with a natural logarithm to construct normal data. Transformed values of Δ FS followed a normal distribution

$p^A$—testing differences in parameters between MND-C and other MNDs through the Mann–Whitney/t test (for non-normal/normal continuous data) or the Pearson Chi-squared test (for categorical data); level of statistical significance 5%

$p^B$—testing differences in parameters between MND-C and myelo-/radiculopathies with Mann–Whitney test (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

$p^C$—testing differences in parameters between other MNDs and myelo-/radiculopathies with Mann–Whitney (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

**Table 3** Association between radiological and clinical findings and biomarker levels among MND-C patients ($n = 18$)

| Radiological findings | Clinical phenotype | | $p$ |
|---|---|---|---|
| | MND-LMN ($n = 7$) | MND-ALS ($n = 11$) | |
| Myelopathy | 1 | 3 | 0.623 |
| CCI grade 3 | 1 | 8 | **0.0498[A]** |
| Pluriradicular compression | 6* | 7* | 0.596 |
| Surgically treated | 2 (28.5%) | 3 (27.3%) | 1.00 |
| **Biomarkers** | | | |
| p-NfH (pg/ml), CSF | 865 (535; 16,750) | 2700 (1250; 16,800) | 0.147 |
| p-NfH (pg/ml), serum | 19 (9; 271) | 70 (15; 456) | 0.378 |
| NfL (pg/ml) | 2240 (480; 5910) | 4350 (1450; 10,000) | **0.038[A]** |

$p$—testing differences in parameters between MND-LMN and MND-ALS with Mann–Whitney test (for continuous variables) and Fisher's exact test (for categorical parameters); 5% level of statistical significance

*Patients had combined radiological findings, i.e. of foraminal stenosis AND CCI grade 3/myelopathy

[A]Statistically significant difference between MND-LMN and MND-ALS

**Table 4** Comparison of characteristics between MND-ALS, MND-LMN, and MND-B

| | MND-ALS ($n=36$) | MND-LMN ($n=15$) | MND-B ($n=20$) | $p$ |
|---|---|---|---|---|
| Females | 13 (36.1%) | 8 (53.3%) | 13 (65.0%) | 0.104 |
| Age (years) | 57 (27; 81) | 62 (38; 76) | 62 (33; 79) | 0.336 |
| Slow progressors | 11 (30.6%) | 8 (53.3%) | 1 (5.0%) | **0.018[A,B]** |
| Intermediate progressors | 13 (36.1%) | 6 (40.0%) | 11 (55.0%) | |
| Fast progressors | 12 (33.3%) | 1 (6.7%) | 8 (40.0%) | |
| Time to LP (months) | 16 (4; 39) | 15 (2; 38) | 10 (5; 36) | 0.295 |
| Δ FS | 0.565 (0.050; 2.300) | 0.300 (0.050; 3.000) | 0.775 (0.300; 4.000) | **0.011[A]** |
| p-NfH (pg/ml), CSF | 2850 (340; 16,800) | 865 (270; 40,000) | 2535 (640; 6800) | **0.017[A]** |
| p-NfH (pg/ml), serum* | 62 (11; 1 455) | 21 (8; 619) | 75 (9; 1102) | 0.347 |
| NfL (pg/ml) | 4830 (1450; 11,220) | 2354 (446; 36,300) | 5977 (1,560; 10,000) | **0.003[A,C]** |

$p$—testing differences in parameters between MND-ALS, MND-LMN and MND-B with Kruskal–Wallis test (for continuous variables) and Pearson Chi-squared test/Fischer's exact test (for categorical parameters); 5% level of statistical significance

[A]Statistically significant difference between MND-LMN and MND-B after Bonferroni correction

[B]Statistically significant difference between MND-ALS and MND-B after Bonferroni correction

[C]Statistically significant difference between MND-LMN and MND-ALS after Bonferroni correction

*Serum p-NfH values were not available in all parameters: MND-ALS ($n=33$), MND-LMN ($n=14$), MND-B ($n=16$) significance

into groups of slow (ΔFS < 0.3; 25. percentile, $n=20$), fast (ΔFS ≥ 1; 75. percentile, $n=21$) and intermediate (ΔFS 0.3–1, $n=30$) progressors (Supplementary Table 1).

There were 93 patients without MND. (Table 1). The first cohort consisted of 30 patients with cervical spondylotic myelopathy or pluriradiculopathy due to lumbar spinal stenosis, multilevel foraminal stenosis, or discogenic compression (for details, see Supplementary Table 2). Medical records of each patient could be tracked for at least one year after LP without signs of developing MND. The second cohort (63 patients, disease controls—DC) comprised subjects with several peripheral nervous system disorders or functional neurological disorders (Table 5).

## Neurofilament assays

Samples from all 71 MND patients and 93 other patients were tested for CSF p-NfH and NfL levels, and samples from 62 MND patients and 62 other patients (26 myelo-/radiculopathies and 36 DC) were tested for p-NfH in serum with ELISA kits. CSF was obtained by lumbar puncture, centrifuged, aliquoted, and stored within 2 h at −80 °C until

**Table 5** Demographics, clinical data, and biomarker concentration of the disease controls (DC)

| | n (CSF/serum) | Sex (M/F) | Age | TP (mg/l) | Qalb | NfL (pg/ml) | p-NfH (pg/ml) CSF | Serum |
|---|---|---|---|---|---|---|---|---|
| CIDP and MMN | 23/17 | 16/7 | 56 (21; 77) | 820 (250; 3200) | 10.6 (2; 55) | 825 (380; 2100) | 450 (150; 2120) | 30 (9; 230) |
| CIAP | 19/12 | 15/5 | 59 (48; 81) | 437.5 (280; 900) | 5.85 (3.5; 25.4) | 725 (520; 1690) | 518 (217; 2560) | 19.5 (9; 129) |
| Polyneuropathy of defined aetiology | 6/0 | 2/4 | 61 (47; 73) | 398 (310; 1130) | 6.65 (4.7; 15.5) | 530 (340;810) | 425 (170; 830) | NA |
| Myasthenia gravis | 5/4 | 4/1 | 71 (28; 74) | 550 (245; 720) | 6.2 (3.5; 9.8) | 785 (450; 1400) | 452 (196; 730) | 17.5 (9; 73) |
| Myopathy | 4/3 | 1/3 | 51 (30; 62) | 360 (250; 475) | 5.05 (3.3; 9.6) | NA | 155 (30; 230) | 22 (11; 43) |
| Functional disorder | 6/0 | 1/5 | 50 (42; 61) | 435 (295; 635) | 5 (3.1; 6.7) | 475 (280; 590) | 213 (150; 380) | NA |

*CIAP* chronic idiopathic axonal polyneuropathy, *CIDP* chronic i demyelinating polyneuropathy, *MMN* multifocal motor neuropathy, *p-NfH* phosphorylated neurofilament heavy chain, *CSF* cerebrospinal fluid, *Qalb* albumin quotient

Absolute frequencies (percentages) are presented for categorical variables. Medians (minimum; maximum) are presented for continuous variables

analysis. A paired sample of serum, obtained at the time of lumbar puncture, was processed likewise. The cell count and concentration of albumin in CSF and sera, with a calculation of CSF/serum albumin ratio (Qalb), were determined in a standard manner. Commercially available ELISA CE marked kits for in vitro diagnostics (IVD) were used both for NfL (UmanDiagnostics AB, Umea, Sweden, kit No. 10-7001 CE), as well as for p-NfH (Euroimmun AG, Luebeck, Germany), according to manufacturers' instructions. For p-NfH in CSF, kit No. EQ 6561-9601 with a declared limit of detection of 27 pg/ml, and for sera, kit No. EQ 6562-9601 with a declared limit of detection of 1.7 pg/ml were used. Samples were measured in duplicate and the mean intra-assay coefficient of variation was: for NfL at 1%, for p-NfH in CSF at 8%, and for p-NfH in serum at 4.4%.

## Statistical analyses

The normality of data was determined with the Shapiro–Wilk test. In non-normally distributed data, a Kruskal–Wallis test was performed and for statistically significant difference, a pairwise comparison through the 2-tailed Mann–Whitney test at a significance level of 5% with Bonferroni correction was performed. Demographic (age, sex) and clinical data (phenotype, progression rate, and duration of disease, particularly for MND subcohorts) were assessed as covariates with a multiple linear regression model. Categorical parameters were tested with Pearson Chi-squared test (multiple comparisons) or Fischer's exact test (pairwise comparison). Correlation analysis of non-normally distributed data was performed with the Kendall correlation test (due to multiple tied ranks and relatively small sample numbers). The diagnostic performances of p-NfH (CSF, serum) and NfL were tested by receiver-operating characteristic (ROC) curves through area under the curve (AUC) calculation. The optimal cut-off was calculated using three different methods: (1) Youden-J index (2) for target sensitivity of 90% to model a clinical screening situation minimizing false negatives and (3) for target specificity of 90% to model a confirmatory test, minimizing false positives. For each optimal cut-off value, the sensitivity, specificity, and predictive values with corresponding 95% confidence intervals (CI) were calculated. Analysis (generation of tables) was performed in IBM SPSS Statistics 28, MSOffice (generation of boxplot figures), and R 4.2 (pROC and ROCit packages).

## Results

### Clinical characteristics

We found no significant differences in age or sex between MND and the other two groups (Table 1). Among different MND subgroups, MND-C patients were significantly older than the other MND patients (Table 2). As for disease duration, fast progressors displayed the shortest and slow progressors the longest time-to-LP interval (Supplementary Table 1), whereas, among other subgroups, there were no statistically significant differences in this regard. As for the progression rate (Δ FS), the MND-LMN subgroup displayed significantly lower values than MND-B (Table 4). Among MND-LMN patients, slow progressors were significantly more and fast progressors significantly less frequently represented, when compared to the MND-B subgroup (Table 4). Within the entire MND-LMN group ($n = 15$), eight patients (53.3%) manifested involvement of upper motor neurons (UMN) during the course of the disease and fulfilled Awaji criteria for ALS, while in seven patients, the lesion was only limited to LMN signs throughout the entire course of the disease—these patients were diagnosed with progressive muscular atrophy (PMA). Patients with LMN-isolated lesion were significantly more frequently represented in the MND-C group (7/18; 38.9%) than in the rest of the MND population (8/53; 15.1%) ($p = 0.046$). Regarding the radiological findings among MND-C patients (Table 3), CCI grade 3, but not myelopathy, was significantly associated with the presence of a clinical UMN lesion. All but one patient with LMN-isolated lesion (within MND-C subgroup) had radiological evidence of (at least two levels of) radicular compression, neuroanatomically corresponding to the clinical deficit, but this association was not statistically significant.

### Concentrations of biochemical markers

We observed statistically significant differences between MND and myelo-/radiculopathies and between MND and disease controls in the levels of p-NfH (CSF), p-NfH (serum), NfL, total protein, and value of Qalb. Patients from the MND group had significantly higher levels of p-NfH (both CSF and serum) and NfL and significantly lower Qalb values and total protein levels compared to the other two groups (Table 1; Fig. 1a, b; Supplementary Fig. 1a). Furthermore, group of myelo-/radiculopathies had significantly higher levels of NfL ($p < 0.001$) and p-NfH in CSF ($p = 0.003$), but not p-NfH in serum ($p = 0.89$) than DC group. In a multiple linear regression model accounting for age and sex, the significant effect of MND on higher levels of NfL and p-NfH in CSF and serum was confirmed (for both Nfs in CSF $p < 0.001$, for p-NfH in serum $p = 0.04$). In this model, a significant effect of age (increasing level with increasing age) and sex (higher levels in males) on levels of NfL and p-NfH in CSF among the entire study population ($p < 0.001$ for both Nfs and both parameters) was observed. For p-NfH in serum, the effect of age ($p = 0.044$), but not of sex ($p = 0.195$), was significant in this model. When only MND patient was tested in this model, no significant effect of
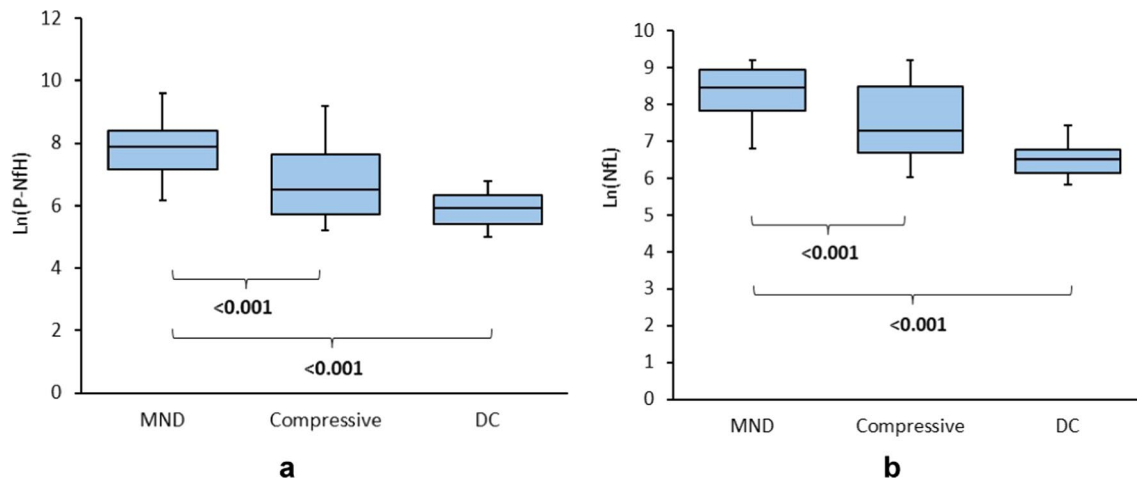
**a**



**b**

**Fig. 1** **a** Comparison of transformed values of p-NfH (CSF, pg/ml) with natural logarithm between MND ($n=71$) and compressive myelo-/radiculopathies ($n=30$) and between MND and disease controls (DC; $n=63$). **b** Comparison of transformed values of NfL (CSF, pg/ml) with natural logarithm between MND ($n=71$) and compressive myelo-/radiculopathies ($n=30$) and between MND and disease controls (DC; $n=63$)

age or sex on levels of CSF NfL (for age: $p=0.654$, for sex: $p=0.223$) or CSF p-NfH (for age: $p=0.138$, for sex: $p=0.552$) was found.

## Correlation between parameters

A correlation analysis of all studied parameters (Supplementary Table 3) revealed a significant positive correlation between age and levels of both CSF Nfs among patients with compressive myelo-/radiculopathies (NfL $\tau_b=0.471$; p-NfH $\tau_b=0.372$) and disease controls (NfL $\tau_b=0.369$; p-NfH $\tau_b=0.313$). We observed no significant correlation between these parameters among MND patients. A significant positive correlation was found between age and p-NfH in serum among myelo-/radiculopathies ($\tau_b=0.440$), but not among disease controls or MND patients. Among all groups, we found a significant positive correlation between levels of p-NfH and NfL in CSF (MND $\tau_b=0.575$, myelo-/radiculopathies $\tau_b=0.526$, DC $\tau_b=0.657$) and a significant, but only moderate correlation between levels of p-NfH in CSF and in serum (MND $\tau_b=0.280$, myelo-/radiculopathies $\tau_b=0.284$, DC $\tau_b=0.283$). Furthermore, between levels of CSF NfL and $\Delta$ FS, a significant strong positive correlation was found ($\tau_b=0.390$), while for CSF p-NfH, a significant moderate positive correlation ($\tau_b=0.265$), and for serum, a significant weak positive correlation ($\tau_b=0.190$) was found in this respect.

## Concentration of neurofilaments in different MND subgroups

Patients from the MND-C group ($n=18$) displayed significantly lower levels of CSF NfL ($p=0.05$), but not of p-NfH (both in CSF and serum) as compared to the rest of the MND patients ($n=53$). Further, MND-C group had significantly higher levels of NfL ($p=0.048$) and p-NfH in CSF ($p=0.007$) (Fig. 2a, b) but not in serum ($p=0.164$), than patients from the group of compressive myelo-/radiculopathies. The rest of the MND patients had significantly higher levels of each Nfs than patients from the myelo-/radiculopathy group (Table 2, Fig. 2a, b).

Patients from the MND-LMN subgroup had significantly lower levels of NfL than patients from the MND-ALS ($p=0.003$) and the MND-B ($p=0.003$) subgroups (Table 4, Fig. 2c). MND-LMN group had significantly lower levels of p-NfH in CSF than MND-ALS ($p=0.008$), but not MND-B ($p=0.027$, NS after Bonferroni correction) groups (Fig. 2d). We found no significant difference in serum p-NfH levels among the three groups ($p=0.347$). The effect of the MND-LMN subtype on NfL levels was also confirmed in a multiple linear regression model including age, sex, $\Delta$FS, and time-to-LP (between MND-LMN and MND-B: $p=0.01$; between MND-ALS and MND-LMN: $p=0.006$). The same model confirmed a significant effect of MND-LMN on CSF p-NfH levels in comparison to MND-ALS ($p=0.042$), but not to MND-B ($p=0.075$). Within the MND-LMN group, slow progressors ($n=8$) had significantly lower levels of NfL ($p=0.002$) and p-NfH in CSF ($p=0.008$) than intermediate and fast progressors. No significant differences in levels of p-NfH (CSF) ($p=0.95$) or NfL ($p=0.68$) were found between patients with PMA ($n=7$) and those, who later fulfilled ALS criteria ($n=8$). Furthermore, the MND-LMN group had significantly higher levels of CSF NfL and p-NfH than disease controls, ($ps < 0.001$), but not of p-NfH in serum ($p=0.642$) (Table 6). No
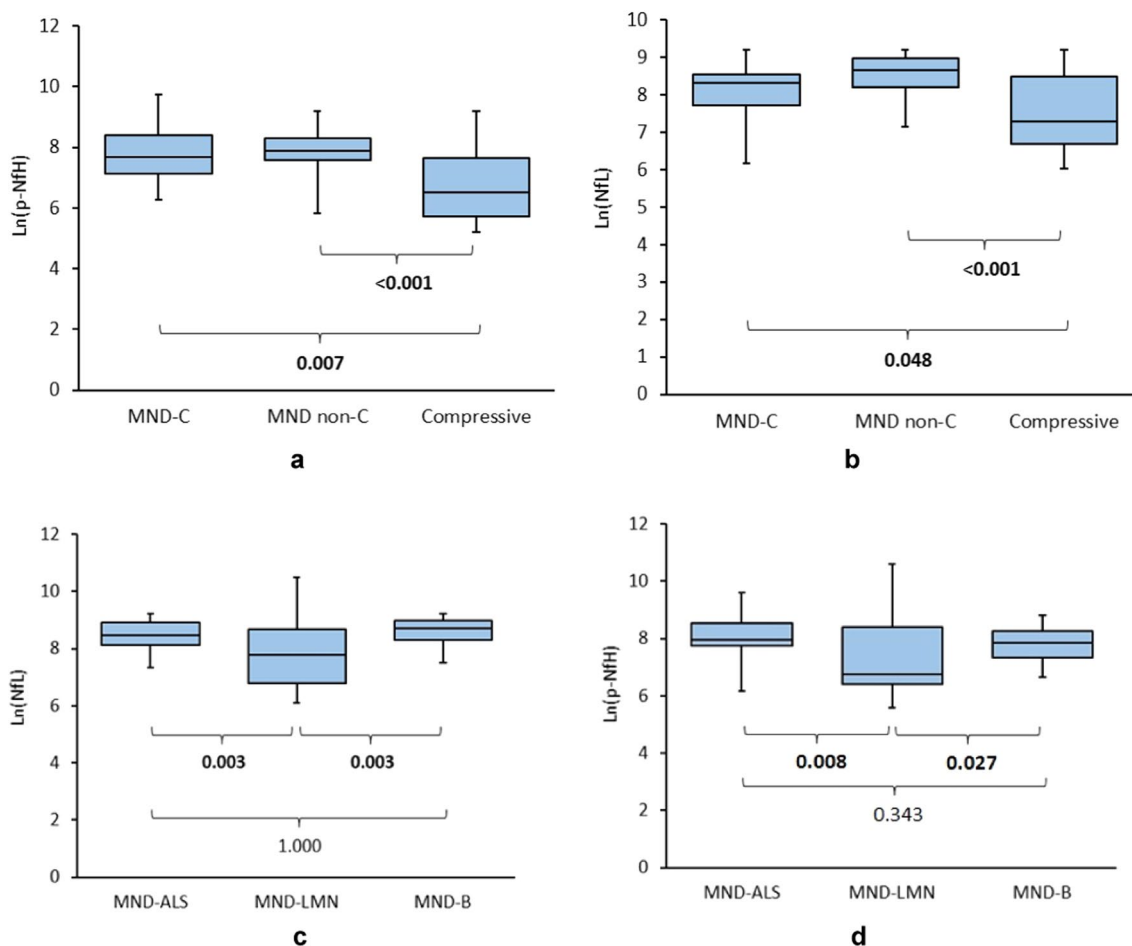
**Fig. 2** **a** Comparison of transformed values (natural logarithm) of p-NfH (CSF, pg/ml) between MND-C ($n=18$), MND non-C ($n=53$), and compressive myelo-/radiculopathies ($n=30$). **b** Comparison of transformed values (natural logarithm) of NfL (CSF, pg/ml) between MND-C ($n=18$), MND non-C ($n=53$), and compressive myelo-/ radiculopathies ($n=30$). **c** Comparison of transformed values (natural logarithm) of NfL (CSF, pg/ml) between MND-ALS ($n=36$), MND-LMN ($n=15$), and MND-B ($n=20$). **d** Comparison of transformed values (natural logarithm) of p-NfH (CSF, pg/ml) between MND-ALS ($n=36$), MND-LMN ($n=15$), and MND-B ($n=20$)

significant difference in any Nfs levels between MND-LMN and myelo-/radiculopathies was observed (Table 6, Supplementary Fig. 1b, c).

Among the cohorts divided according to progression rate, fast progressors had the highest and slow progressors the lowest NfL levels in the CSF (Supplementary Table 1). The independent effect of a group on the levels of NfL was confirmed in a multiple linear regression model including age, sex, phenotype, and time-to-LP (between slow and fast progressors: $p = 0.005$, between intermediate and slow progressors: $p = 0.049$). The difference in levels of CSF p-NfH was significant only between the fast and slow progressor group, confirmed by the multiple linear regression model ($p = 0.038$). No significant difference in p-NfH levels in serum among the three groups was found.

## Diagnostic performance of Nfs in discriminating MND from other disorders

In the entire study population ($n = 164$), the ability of both CSF Nfs to discriminate patients with MND from the other two groups combined was good (CSF p-NfH: AUC 0.885, CI 0.832–0.938; CSF NfL: AUC 0.892, CI 0.839–0.946), while for p-NfH in serum, it was poor (AUC 0.650, CI 0.553–0.746) (Table 7a; Fig. 3a, b; Supplementary Fig. 1d). The discriminative ability of CSF NfL and CSF p-NfH was comparable (AUC comparison between ROC curves $p = 0.723$) and significantly higher than that of p-NfH in serum (for both $p < 0.001$). All the 3 cut-off values (Youden-J index, target sensitivity, and specificity of 90%, both respectively) were calculated (Table 7a). For CSF p-NfH, a Youden-J cut-off value of 1205 pg/ml reached

**Table 6** Comparison of characteristics between MND-LMN, myelo-/radiculopathies, and DC

| | MND-LMN (*n* = 15) | Myelo-/radiculopathies (*n* = 30) | DC (*n* = 63) | $p^A$ | $p^B$ |
|---|---|---|---|---|---|
| Females | 8 (53.3%) | 11 (36.7%) | 25 (39.7%) | 0.286 | 0.336 |
| Age (years) | 62 (38; 76) | 64 (33; 83) | 58 (21; 81) | 0.621 | 0.530 |
| p-NfH (pg/ml), CSF | 865 (270; 40 000) | 676 (155; 33,170) | 380 (30; 2560) | 0.178 | **< 0.001** |
| p-NfH (pg/ml), serum | 21 (8; 619) | 27 (8; 1883) | 25 (9; 230) | 0.887 | 0.642 |
| NfL (pg/ml) | 2354 (446; 36,300) | 1470 (350; 10000) | 670 (90; 2100) | 0.393 | **< 0.001** |
| Qalb | 6.1 (3.1; 27.3) | 9.6 (3.4; 19.4) | 6.7 (2.0; 55.0) | 0.053 | 0.270 |
| Total protein (mg/l) | 415 (250; 1740) | 623 (230; 1200) | 480 (245; 3200) | **0.014** | 0.179 |

Absolute frequencies (percentages) are presented for categorical variables. Medians (minimum; maximum) are presented for continuous variables

$p^A$—testing differences in parameters between MND-LMN and myelo-/radiculopathies with Mann–Whitney test (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

$p^B$—testing differences in parameters between MND-LMN and disease controls with Mann–Whitney (for continuous variables) and Pearson Chi-squared test (for categorical parameters); 5% level of statistical significance

a sensitivity of 81.7 (CI 73.2–90.1%), specificity of 86.0 (CI 78.5–92.5%), with PPV of 81.7 (CI 78.5–92.5%) and NPV of 86.0 (CI 76.9–92.3%). For target sensitivity of 90%, the cut-off value was 625 pg/ml, with a specificity of 71.0 (CI 61.3–80.7%), while for target specificity of 90%, a cut-off value of 2101.5 pg/ml reached a sensitivity of 67.6 (CI 55.5–78.2%). For CSF NfL, the Youden-J cut-off value of 2180 pg/ml was identical to that of the target specificity of 90%, with a sensitivity of 83.1 (CI 74.7–90.1%), PPV of 86.8 (CI 77.0–93.4%), and NPV 87.5 (CI 79.9–94.5%). For the target sensitivity of 90%, a cut-off value of 1525 pg/ml reached a specificity of 80.6 (CI 73.1–88.2%). When only MND and DC groups were included in the analysis (Supplementary Table 4a), diagnostic performance of CSF p-NfH, as well as of CSF NfL, (but not of p-NfH in serum) further significantly increased (for CSF p-NfH: AUC 0.954, CI 0.904–0.981, AUC comparison between ROC curves *p* = 0.044; for NfL: AUC 0.962, CI: 0.928–0.996, *p* = 0.014; for p-NfH in serum: AUC 0.656, CI 0.545–0.767, *p* = 0.46) becoming excellent for both CSF Nfs. Further, all the respective calculated cut-off values were lower, with higher sensitivities, specificities, PPV, and NPV as compared to the analysis with myelo-/radiculopathies included (Supplementary Table 4a).

The ability of CSF p-NfH to discriminate the MND-C group from myelo-/radiculopathies was fair (Table 7b; AUC 0.733, CI 0.592–0.875; Fig. 3c) and not significantly lower than in the respective comparison between the rest of the MND group (*n* = 53) and myelo-/radiculopathies (Supplementary Table 4b; AUC 0.774, CI 0.627–0.869; AUC comparison between ROC curves: *p* = 0.341). For p-NfH in serum (AUC 0.627, CI 0.444–0.809) and NfL (AUC 0.672, CI 0.514–0.830; Fig. 3d), the discriminative ability was poor (Table 7b), but not significantly different as compared

to other MND patients in this respect (Supplementary Table 4b; CSF NfL: AUC 0.768, CI 0.621–0.862; AUC comparison *p* = 0.341). For CSF p-NfH, a Youden-J cut-off value of 607.5 pg/ml reached a sensitivity of 94.4 (CI 83.3–100%) and specificity of 50 (CI 33.3–66.7%), with PPV of 53.1 (CI 35.0–70.5%) and NPV of 93.7 (CI 67.7–99.7%). For target sensitivity of 90%, a cut-off value was 702 pg/ml, with a specificity of 50 (CI 33.3–66.7%) and for target 90% specificity, a cut-off value of 5014 pg/ml reached a sensitivity of 22.2 (CI 5.6–44.4%). For NfL, the Youden-J cut-off value was 2180 pg/ml, with a sensitivity of 77.8 (CI 55.6–94.4%), specificity of 70 (CI 53.3–86.7%), PPV of 58.3 (CI 36.9–77.2%) and NPV of 83.3 (CI 61.8–94.5%). For the target sensitivity of 90%, a cut-off value was 895 pg/ml, with a specificity of 33.3 (CI 16.7–50.0%), and for the target 90% specificity, a cut-off value of 7682 pg/ml reached a sensitivity of 11.1 (CI 1.9–36.0%). All data combined, we could not reach an optimal cut-off value without substantial loss of either sensitivity or specificity, especially for p-NfH in CSF. Therefore, cut-off values derived from the entire study population have been applied (for p-NfH in CSF: 1205 pg/ml, for NfL: 2180 pg/ml). When these values were applied, 11 (36.7%) and 9 (30%) patients from the myeloradiculopathy group had levels above the cut-offs for CSF p-NfH and NfL, respectively. In seven out of ten patients (70%) from the myelo-/radiculopathy group, who underwent surgical decompression, the measured levels of both CSF Nfs were simultaneously above these cut-offs. On the other hand, five patients from the MND-C group (27.7%) underwent surgical treatment due to suspected compressive aetiology before a referral to our centre. In two out of these five patients (40%), the levels of both Nfs in CSF were below the cut-off values.

When comparing MND-LMN patients with the disease control group, the discriminative ability of CSF p-NfH

**Table 7** Results of receiver-operating characteristic (ROC) analysis discriminating patients

(a) from MND group ($n=71$) and from myelo-/radiculopathy and disease control groups ($n=93$)

| Parameter | Cut-off | Sensitivity (95% CI) | Specificity (95% CI) | PPV | NPV | AUC (95% CI) |
|---|---|---|---|---|---|---|
| p-NfH (pg/ml), CSF | 1205* | **81.7**% (73.2–90.1%) | **86.0**% (78.5–92.5%) | **81.7**% (78.5–92.5%) | **86.0**% (76.9–92.3%) | **0.885** (0.832–0.938) |
| 90% sensitivity | 625 | **90.1**% (83.1–95.8%) | **71.0**% (61.3–80.7%) | **70.3**% (59.9–86.0%) | **90.4**% (81.2–93.9%) | |
| 90% specificity | 2101.5 | **67.6**% (55.5–78.2%) | **90.3**% (82.4–95.5%) | **84.2**% (72.8–90.2%) | **78.5**% (68.5–89.2%) | |
| p-NfH (pg/ml), serum | 41.75* | **61.3**% (48.4–72.6%) | **64.5**% (53.2–75.8%) | **63.3**% (49.8–75.1%) | **62.5**% (49.5–74.0%) | **0.650** (0.553–0.746) |
| 90% sensitivity | 12.5 | **90.3**% (82.3–96.8%) | **32.3**% (21.0–45.2%) | **57.1**% (46.7–67.0%) | **76.9**% (55.9–90.3%) | |
| 90% specificity | 130 | **25.8**% (14.5–37.1%) | **90.3**% (82.3–96.8%) | **72.7**% (49.6–88.0%) | **54.9**% (44.7–64.6%) | |
| NfL (pg/ml) | 2180* | **83.1**% (74.7–90.1%) | **90.3**% (83.9–95.7%) | **86.8**% (77.0–93.4%) | **87.5**% (79.9–94.5%) | **0.892** (0.839–0.946) |
| 90% sensitivity | 1525 | **90.1**% (83.1–95.8%) | **80.6**% (73.1–88.2%) | **78.0**% (73.1–90.7%) | **91.5**% (84.0–94.7%) | |
| 90% specificity | 2180 | **83.1**% (74.7–90.1%) | **90.3**% (83.9–95.7%) | **86.8**% (77.0–93.4%) | **87.5**% (79.9–94.5%) | |

(b) from MND-C group ($n=18$) and from myelo-/radiculopathies ($n=30$)

| Parameter | Cut-off | Sensitivity (95% CI) | Specificity (95% CI) | PPV | NPV | AUC (95% CI) |
|---|---|---|---|---|---|---|
| p-NfH (pg/ml), CSF | 607.5* | **94.4**% (83.3–100%) | **50.0**% (33.3–66.7%) | **53.1**% (35.0–70.5%) | **93.7**% (67.7–99.7%) | **0.733** (0.592–0.875) |
| 90% sensitivity | 702 | **88.9**% (72.2–100%) | **50.0**% (33.3–66.7%) | **51.6**% (33.4–69.4%) | **88.2**% (62.2–97.9%) | |
| 90% specificity | 5014 | **22.2**% (5.6–44.4%) | **90.0**% (80.0–100%) | **57.1**% (20.2–88.2%) | **65.8**% (49.3–79.4%) | |
| p-NfH (pg/ml), serum | 48* | **64.7**% (41.2–88.2%) | **73.1**% (53.9–88.5%) | **61.1**% (36.1–81.7%) | **76.0**% (54.4–89.8%) | **0.627** (0.444–0.809) |
| 90% sensitivity | 9.5 | **82.3**% (55.8–95.3%) | **7.7**% (1.3–26.6%) | **36.8**% (22.3–54.0%) | **40.0**% (7.2–82.9%) | |
| 90% specificity | 150.5 | **17.6**% (4.6–44.2%) | **88.5**% (68.7–96.9%) | **50.0**% (13.9–86.1) | **62.1**% (44.8–77.1%) | |
| NfL (pg/ml) | 2180* | **77.8**% (55.6–94.4%) | **70.0**% (53.3–86.7%) | **58.3**% (36.9–77.2%) | **83.3**% (61.8–94.5%) | **0.672** (0.514–0.830) |
| 90% sensitivity | 895 | **88.9**% (72.2–100%) | **33.3**% (16.7–50.0%) | **44.4**% (28.3–61.7%) | **83.3**% (50.9–97.1%) | |
| 90% specificity | 7616 | **11.1**% (1.9–36.0%) | **90.0**% (72.3–97.4%) | **40.0**% (7.2–82.9%) | **62.8**% (46.7–76.6%) | |

(c) from MND-LMN group ($n=15$) and from disease controls ($n=63$)

| Parameter | Cut-off | Sensitivity (95% CI) | Specificity (95% CI) | PPV | NPV | AUC (95% CI) |
|---|---|---|---|---|---|---|
| p-NfH (pg/ml), CSF | 605* | **80.0**% (60.0–95.7%) | **81.0**% (71.4–90.5%) | **50.0**% (34.5–84.7%) | **94.4**% (82.1–97.2%) | **0.848** (0.671–0.934) |
| 90% sensitivity | 535 | **86.7**% (59.5–98.3%) | **66.7**% (71.4–90.5%) | **37.9**% (26.4–85.0%) | **95.5**% (82.6–97.4%) | |
| 90% specificity | 727 | **66.7**% (38.4–88.2%) | **90.5**% (82.5–96.8%) | **62.5**% (41.9–86.0%) | **92.1**% (78.0–97.0%) | |
| NfL (pg/ml) | 2170* | **60.0**% (33.3–86.7%) | **100**% (94.3–100%) | **100**% (70.3–100%) | **91.4**% (76.9–100%) | **0.838** (0.694–0.981) |
| 90% sensitivity | 800 | **87.7**% (59.5–98.3%) | **68.3**% (15.9–36.5%) | **39.4**% (29.3–85.6%) | **94.2**% (83.0–97.5%) | |
| 90% specificity | 1216.5 | **66.7**% (46.5–86.7%) | **90.5**% (82.5–96.8%) | **62.5**% (41.9–86.1%) | **92.1**% (78.0–97.0%) | |

*PPV* positive predictive value, *NPV* negative predictive value, *AUC* area under the curve, *CI* confidence interval

*Cut-off value was obtained through maximum value of Youden's index for the ROC curve

Confidence interval for AUC is computed through DeLong method of variance. Confidence interval for sensitivity and specificity is computed through 2000 stratified bootstrap replicates.

(AUC 0.848, CI 0.671–0.934) and NfL (AUC 0.838, CI 0.694–0.981) was good (Table 7c). For CSF NfL, this result was significantly lower than in the comparison between other MND patients ($n=56$) and DC (Supplementary Table 4 c; AUC 0.995, CI 0.981–0.999; AUC comparison between ROC curves: $p=0.032$), while for CSF p-NfH, the difference was not significant in this respect (AUC 0.968, CI 0.920–0.987, AUC comparison between ROC curves: $p=0.064$). Serum p-NfH failed to discriminate between MND-LMN and DC (AUC 0.457, CI 0.254–0.661). For p-NfH (CSF), the Youden-J cut-off value of 605 pg/ml reached a sensitivity of 80 (CI 60.0–95.7)% and specificity of 81 (CI 71.4–90.5)%, with PPV of 50 (CI 34.5–84.7%) and NPV of 94.4 (CI 82.1–97.2%). For target sensitivity of 90% a cut-off 535 pg/ml had specificity of 66.7 (CI 71.4–90.5%), and for target specificity of 90%, a cut-off value of 727 pg/ml reached sensitivity 66.7 (CI 38.4–88.2%), with PPV 62.5 (CI 41.9–86.1%) and NPV 92.1 (CI 78.0–97.0%). For NfL, the Youden-J cut-off value of 2170 pg/ml reached a sensitivity of 60 (CI 33.3–86.7%), with a specificity of 100 (CI 94.3–100%), PPV of 100 (CI 70.3–100%) and NPV of 91.4 (CI 76.9–100%). A cut-off value for target sensitivity of 90% was 800 pg/ml, with a specificity of 68.3 (CI 55.3–79.3%). For target specificity of 90%, a cut-off value of 1216.5 pg/ml reached a sensitivity of 66.7 (CI 46.5–86.7%). Regarding the choice of an optimal cut-off value, in a clinical situation
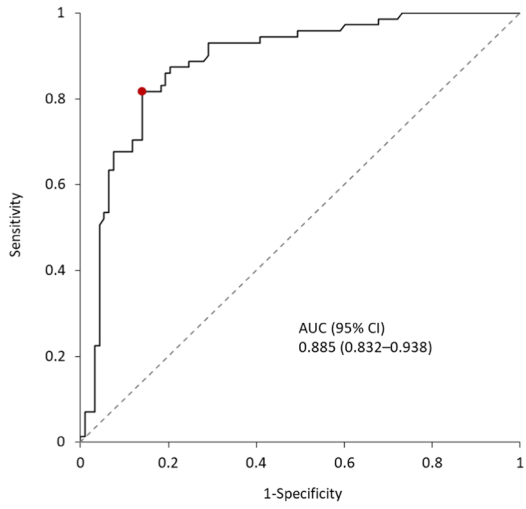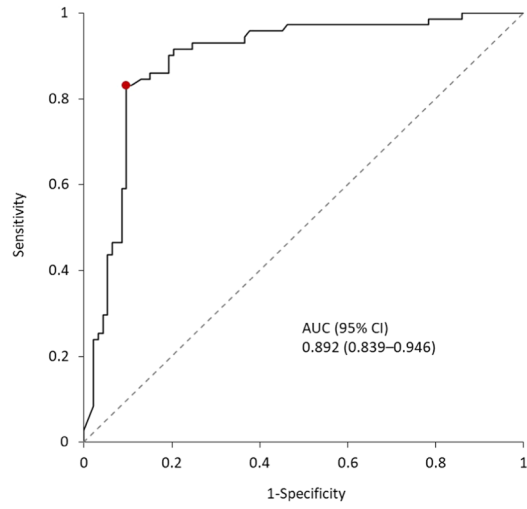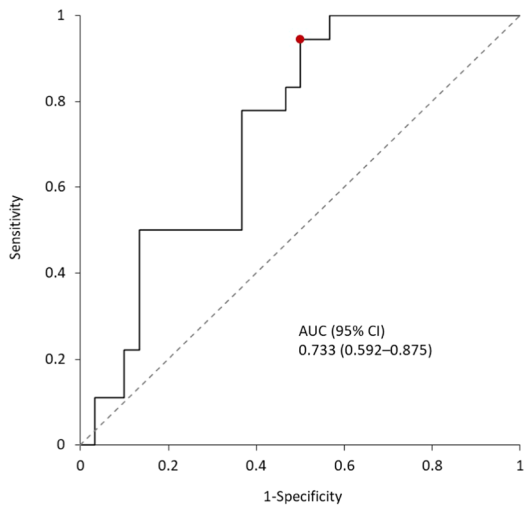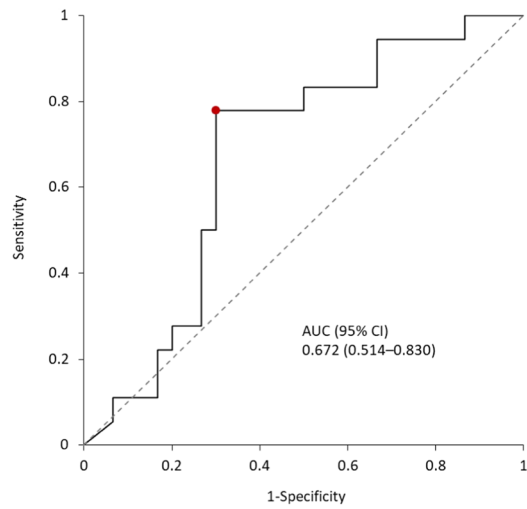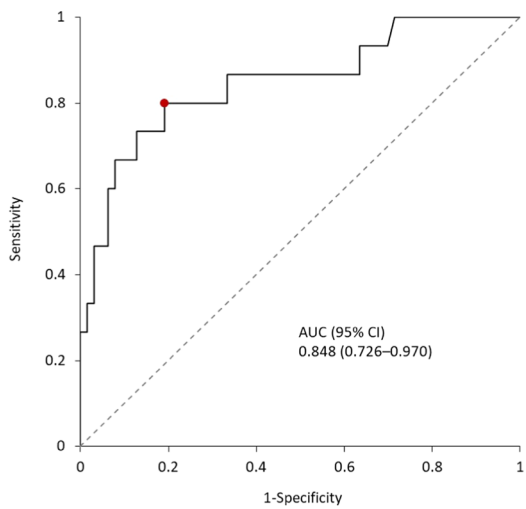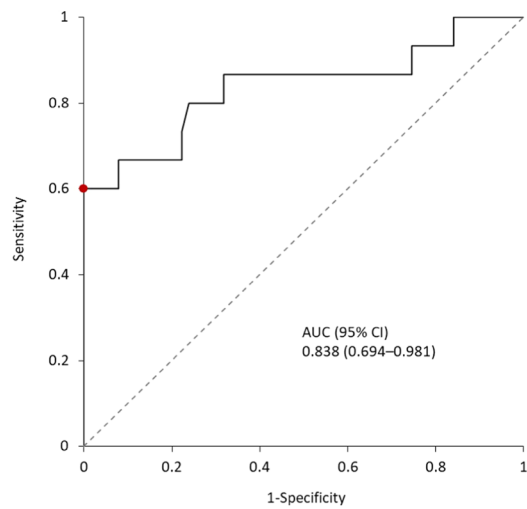
◄**Fig. 3** **a** ROC curve for p-NfH in CSF (pg/ml) discriminating MND from myelo-/radiculopathies and disease control group; red point shows p-NfH cut-off value obtained as maximum value of Youden's index; *AUC* area under the curve, *CI* confidence interval. **b** ROC curve for NfL (pg/ml) discriminating MND from myelo-/radiculopathies and disease control group; **c** ROC curve for p-NfH in CSF(pg/ml) discriminating MND-C from myelo-/radiculopathies group. **d** ROC curve for NfL (pg/ml) discriminating MND-C from myelo-/radiculopathies group. **e** ROC curve for p-NfH in CSF(pg/ml) discriminating MND-LMN from disease controls. **f** ROC curve for NfL (pg/ml) discriminating MND-LMN from disease controls

where diagnostic confirmation is requested, the high specificity of a test has to be preferred. For discrimination between MND-LMN and disease controls, we, therefore, selected the optimal cut-off for p-NfH (CSF) as 727 pg/ml and NfL as 2170 pg/ml based on the highest achievable test specificity and positive predictive value, while maintaining acceptable sensitivity and negative predictive values When analysing the discriminative ability between all MND (=71) and DC groups, for CSF p-NfH, identical cut-off value has been achieved, while for NfL, a lower value of 1425 pg/ml has yielded the best diagnostic performance (Supplementary Table 4a). As there were seven patients with LMN-restricted phenotype with a confounding radiological finding of multiple nerve roots compression (46.6% of the MND-LMN subgroup), we have compared the MND-LMN group to all controls combined (*n*=93), but could not reach an optimal cut-off value without a substantial loss of either sensitivity or specificity, and most importantly, with very poor positive predictive values (Supplementary Table 4d). Therefore, we have applied the cut-off values established from the entire study population (all MND, *n*=71 compared to all controls, *n*=93), i.e. for CSF p-NfH 1205 pg/ml, and for CSF NfL 2180 pg/ml, on the MND-LMN group, to account for a clinical scenario of differentiation between MND-LMN and compressive myelo-/radiculopathies as well. These cut-off values yielded 8 (53.3%) and 6 (40%) false negative MND-LMN patients, respectively.

## Discussion

To the best of our knowledge, this is the first study analysing the discriminative ability of Nfs towards specific MND subgroups, where the clinician is effectively in doubt due to a restricted phenotype (limited clinical signs of the disease) or confounding radiological findings of a nerve root or spinal cord compression.

Our study demonstrated that both p-NfH and NfL in CSF have an overall good discriminative ability towards the group of all MND patients in the entire study population. When patients with myelo-/radiculopathy—a group with NfL and p-NfH levels in CSF significantly higher than in the

group of disease controls—were excluded from the analysis, the discriminative ability of both Nfs further increased (Supplementary Table 4a). In addition, after this exclusion, our diagnostic performance and respective calculated cut-off values towards MND strongly resembled those of previous studies, where the structure of disease control (or mimic) groups was comparable to our DC group [18–20]. Given the unspecific nature of Nfs elevation, reflecting rather the extent and pace of axonal degeneration, we assume, the main reason for the observed difference (myelo-/radiculopathies compared to disease controls) is based on a higher burden of neuroaxonal damage associated with acute (multiple) nerve root or myelon compression. Our own data support this assumption, as we have observed significantly higher levels of both CSF Nfs among acute myelopathies, as well as pluriradiculopathies, compared to chronic myelopathies or monoradiculopathies (Supplementary Table 2). The effect of pace and extent of neuroaxonal damage on p-NfH levels among compressive myelopathies or radiculopathies has been documented by other authors as well [16, 17]. On the other hand, the group of disease controls comprised aetiologies with chronic, slow, and predominantly peripheral nerve involvement or chronic, predominantly demyelinating polyradiculoneuropathies, where the degree of neuroaxonal damage seems to be less pronounced. After the inclusion of myelo-/radiculopathies, the discriminative ability of both Nfs remained good, although significantly lower than that of the former analysis, but the respective cut-off values increased (i.e. for p-NfH in CSF from 840 to 1205 pg/ml), and positive predictive values were reduced (i.e. for p-NfH in CSF from 92.5%, CI 82.7–97.2% to 81.7%, CI 78.5–92.5%). Our data confirm the difficulty in homogenizing cut-off values among different studies (even when identical ELISA kits are used) and highlight the importance of the control group structure. When a wide range of disorders with a higher burden of neuroaxonal damage are included as controls, the cut-off values may vary substantially, given that elevation of Nfs levels is independent of causal pathways.

On the other hand, the diagnostic performance of p-NfH in serum was significantly lower than that of its CSF counterpart, as well as than that of NfL in CSF, and was overall poor. The main rationale for using blood-based Nf measurement was its better accessibility, e.g. in the setting of an outpatient clinic. On the other hand, CSF could more directly reflect CNS pathophysiology. Therefore, we examined how well the value of serum measurements could serve as a proxy for CSF measurements. We used a commercially available CE marked ELISA Kit for p-NfH assessment in sera, as this would represent a widely accessible solution for routine laboratory testing outside of an academic setting. Furthermore, promising results from other research group [19] using this type of ELISA kit prompted us to replicate their results. Regarding NfL measurement in sera, the

commercially available ELISA kit has been deemed unsuitable for analysis of blood samples due to its limited sensitivity combined with significantly lower NfL blood levels, compared to CSF. Meanwhile, ultrasensitive single-molecule array (SiMoA) digital immunoassay is about 100-fold more sensitive than ELISA for quantification of NfL. Although proposed as a potential gold standard for blood NfL quantitation, its high cost and designation as for research use only limit the accesibility of SIMoA. Due to limited funding, we therefore choose only p-NfH quantitation in our study. Our data show significantly lower diagnostic performances of p-NfH in serum as compared to similarly designed studies [19, 27]. Nevertheless, in a previous, differently designed study, combining the cross-sectional and longitudinal design of p-NfH measurement in plasma, a broad overlap of values between MND and controls was observed [28]. The authors attributed this finding to a documented high variability of p-NfH levels over time among MND patients, depending on the actual stage of disease and progression rate. The authors argued that an inter-individual difference in antibody response to p-NfH, as well as antibody-independent aggregation (matrix effect) of p-NfH, could be responsible for the p-NfH blood levels variation among MND patients. Consequently, they deemed the cross-sectional measurement of p-NfH levels in plasma unreliable. Furthermore, our data are in line with those of another study that compared the performance of three analytical approaches in serum (namely, ELISA and two different types of the single molecular array—SIMoA kits) [29]. This comparative study found that SIMoA was superior to ELISA regarding analytical sensitivity in serum. Given these and our data, it can be speculated that p-NfH measured with, as for now available, second-generation ELISA kits has a clearly limited value as a diagnostic blood marker.

Our secondary goal was to determine the clinical parameters in MND influencing levels of Nfs. While among MND, we found no significant effect of age or sex on NfL or p-NfH levels in CSF or serum, the opposite was observed among other neurological diagnoses. Both findings are in accordance with a meta-analysis of studies on NfL in CSF [12], as well as with other studies on p-NfH [15, 20]. We found two significant factors affecting the levels of NfL and p-NfH in CSF in MND patients: (1) clinical phenotype (lower levels in LMN-isolated disease, a model of the limited extent of involvement) and (2) disease progression rate. Previous studies also observed a significant effect of clinical phenotype, with lower CSF levels of NfL in isolated LMN phenotypes [21] and of both Nfs in patients with a limited extent of simultaneous UMN and LMN involvement [18]. Although we could speculate a correlation between CSF NfL levels and the burden of corticospinal tract involvement in light of our data, a different design including MRI markers (fractional anisotropy, radial diffusivity) would be needed to

add to existing evidence [30, 31], and such an assessment was beyond the scope of our study. The primary purpose to include the MND-LMN subgroup in our design was to represent a real-life clinical setting, where missing signs of UMN involvement in the early diagnostic phase broaden the spectrum of differential diagnoses. In these cases, a reliable diagnostic marker would have a great added value. Furthermore, this group provided evidence for the limited reliability of phenotyping patients in the early MND phases, given that half of its patients manifested UMN signs later on and fulfilled the criteria of classic ALS, while the other half fulfilled the PMA criteria (as a true restricted disease phenotype). In such conditions, a difference in levels of Nfs between these patients would have been of great significance. Unfortunately, we did not observe any significant difference in this respect. A plausible explanation would be a very similar progression rate, which we observed between these two subcohorts.

The ability of both CSF Nfs to differentiate MND patients with isolated LMN involvement from disease controls was reduced as compared to their respective ability towards other MND patients. This effect was more prominent in the case of CSF NfL, reaching a statistically significant difference. For CSF p-NfH, although statistically not significant, the difference was also substantial. The significant association of the MND-LMN group with lower levels of both CSF Nfs (compared to the rest of the MND patients) resulted in a higher overlap of values between the MND-LMN and DC groups. Although the diagnostic performance of both CSF Nfs remained good from the statistical perspective, the further in-depth analysis revealed the problematic point from the clinical perspective. Our cut-off values for the discrimination of MND-LMN from DC (as well as for discrimination of all MND from DC only) were similar to other published data [18–20] Looking in detail at false negative MND-LMN patients after our cut-offs were applied, all of them (5/15, 33.3% with p-NfH, 6/15, 40% with NfL cut-off) were simultaneously slow progressors, a sub-cohort with significantly lower levels of both CSF Nfs even within the MND-LMN group. In these patients, the correct diagnosis had to be established by means of a rather extensive diagnostic work-up, involving electrophysiology (exclusion of a motor conduction block, confirmation of a denervation–reinnervation syndrome in multiple myotomes), laboratory diagnostics (e.g. antiganglioside antibodies), molecular genetic testing (e.g. CAG expansion in AR gene, SMN1 deletion, mutation in HEXA gene, etc.), and clinical monitoring for a confirmatory disease progression pattern. Furthermore, as a substantial proportion of MND-LMN group (46.6%) simultaneously displayed a confounding radiological finding, we analysed the discriminative ability of CSF Nfs between MND-LMN and both control groups combined (Supplementary Table 4d). In this analysis, most

prominently, the positive predictive values of any of the respective cut-off value for both CSF Nfs have not even reached 50%, a major limitation for the clinical utility of these biomarkers. After the cut-off values established from the entire study population have been applied, 6 (40%) and 8 (53.3%) MND-LMN patients, respectively, had levels below the cut-off value for CSF NfL and p-NfH. Again, all these false negative patients were simultaneously slow progressors and four of them had confounding radiological finding. Thus, for the discrimination of LMN-predominant MND patients with slow disease progression rate from disease mimics (and, among them, particularly from compressive radiculopathies), quantification of Nfs appears to be rather unsuitable. For the discrimination of all MND patients with suspicious radiological findings (MND-C) from patients with only compressive aetiology, the diagnostic performance of CSF p-NfH was fair, but in the case of p-NfH in serum and CSF NfL, it was poor. Significantly higher levels of both Nfs in myelo-/radiculopathies (compared to the DC group) have, again, led to a broader overlap of values with MND patients. Consequently, the applied cut-off values for both CSF Nfs resulted in high false positive rate among myelo-/radiculopathy patients. Looking at the clinical implications of these data in further detail, 70% of the surgically treated patients from the myelo-/radiculopathy group (all of whom subsequently improved after the treatment) tested above the cut-off levels for both Nfs. Labelled as false positives by the test, these patients would be at risk of being withheld a therapeutic benefit. On the other hand, there were still four (22.2%) MND patients with levels of both CSF Nfs below their respective cut-off. Two of these false negative MND patients underwent surgical treatment (out of five surgically treated MND patients altogether) before a referral to our centre. The correct diagnosis in all the patients from MND-C group has been established by means of a careful clinico-radiological correlation (absence of radicular pain and sensitive deficit, signs of a lower or upper motor neuron lesion in regions, unaffected by the documented compression) and/or electrophysiological examination (proof of denervation–reinnervation syndrome in regions, unaffected by the documented compression). Given the substantially higher prevalence of compressive myelo-/radiculopathies than MND in the general population, and the risk of withholding or delaying a potential therapeutical benefit by means of misdiagnosis, these data highlight a limited utility of Nfs as diagnostic markers in clinically difficult to distinguish cases.

The main limitation of our study was its retrospective design, leading to a bias towards MND patients who underwent CSF examination at our department. This represented only about 40% of all patients examined at our centre during the time frame. Other MND patients were only examined at the outpatient clinic, and CSF examination was already performed elsewhere (and therefore not repeated). Thus, this limitation has led to a reduced number of recruited patients when compared to similar studies. The effect of insufficient sample size was most pronounced in the comparison between subgroups of interest (MND-C, MND-LMN). A future study, preferably with multi-centric design to allow for adequate enrichment in MND patients with concomitant radiological findings or restricted phenotype to either confirm or contradict our preliminary findings, is therefore warranted.

In conclusion, from the perspective of a general MND population, both Nfs in CSF performed well as diagnostic markers. Despite this, caution must be applied, as higher levels of Nfs are not specific towards MND, reflecting rather a burden of an ongoing unspecific neuroaxonal damage. Nevertheless, when the clinical picture is highly suggestive of (and respecting exclusion criteria for) MND and is further supported by electrophysiology and not contradicted by radiological findings, Nfs levels above their respective cut-off values could be a strong support to the definitive diagnosis. Both Nfs in CSF displayed comparable diagnostic performance. On the other hand, the significant covariates that render the clinical picture less specific towards MND (LMN-restricted phenotype, slow disease progression) are simultaneously associated with lower levels of Nfs, which limits the utility of Nfs as a reliable diagnostic marker for the LMN-restricted disease. Furthermore, due to the high overlap of values between MND patients and patients with compressive radiculopathies and myelopathies (mainly those with acute progression of deficits, who may benefit from surgical treatment), the diagnostic performance of Nfs is insufficient for a reliable discrimination in this situation. Finally, although measurement of p-NfH from sera using ELISA would be a feasible alternative for an outpatient clinic, a broad overlap of values between MND and other neurological disorders, and its overall low sensitivity towards MND, render this test unsuitable for a screening task. In the era of fourth-generation assays (SIMoA) with superior analytical sensitivity, the mere better accesibility of the blood-based ELISA method does not seem to outweigh its low diagnostic performance.

## Declarations

**Conflicts of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Patient consent** Informed consent was obtained from the patients who participated in the study.

**Ethical approval** The study was approved by the Ethics Committee of the University Hospital Motol, Prague and has, therefore, been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.

## References

1. Chiò A, Calvo A, Moglia C, Mazzini L, Mora G, PARALS study group (2011) Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study. J Neurol Neurosurg Psychiatry 82(7):740–746. https://doi.org/10.1136/jnnp.2010.235952

2. Bendotti C, Bonetto V, Pupillo E et al (2020) Focus on the heterogeneity of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 21(7–8):485–495. https://doi.org/10.1080/21678421.2020.1779298

3. Cellura E, Spataro R, Taiello AC, La Bella V (2012) Factors affecting the diagnostic delay in amyotrophic lateral sclerosis. Clin Neurol Neurosurg 114(6):550–554. https://doi.org/10.1016/j.clineuro.2011.11.026

4. Srinivasan J, Scala S, Jones HR, Saleh F, Russell JA (2006) Inappropriate surgeries resulting from misdiagnosis of early amyotrophic lateral sclerosis. Muscle Nerve 34(3):359–360. https://doi.org/10.1002/mus.20555

5. Pinto S, Swash M, de Carvalho M (2014) Does surgery accelerate progression of amyotrophic lateral sclerosis? J Neurol Neurosurg Psychiatry 85(6):643–646. https://doi.org/10.1136/jnnp-2013-305770

6. Sanderson AB, Arnold WD, Elsheikh B, Kissel JT (2015) The clinical spectrum of isolated peripheral motor dysfunction. Muscle Nerve 51(3):358–362. https://doi.org/10.1002/mus.24326

7. Simon NG, Ayer G, Lomen-Hoerth C (2013) Is IVIg therapy warranted in progressive lower motor neuron syndromes without conduction block? Neurology 81(24):2116–2120. https://doi.org/10.1212/01.wnl.0000437301.28441.7e

8. Poesen K, Van Damme P (2019) Diagnostic and prognostic performance of neurofilaments in ALS. Front Neurol 18(9):1167. https://doi.org/10.3389/fneur.2018.01167

9. Yuan A, Rao MV, Veeranna, Nixon RA (2017) Neurofilaments and neurofilament proteins in health and disease. Cold Spring Harb Perspect Biol 9(4):a018309. https://doi.org/10.1101/cshperspect.a018309 (**Published 2017 Apr 3**)

10. Lobsiger CS, Cleveland DW (2009) Neurofilaments: organization and function in neurons. In: Squire LR (ed) Encyclopedia of neuroscience. Elsevier, Amsterdam, pp 433–436. https://doi.org/10.1016/B978-008045046-9.00728-2

11. Gaiottino J, Norgren N, Dobson R et al (2013) Increased neurofilament light chain blood levels in neurodegenerative neurological diseases. PLoS ONE 8(9):e75091. https://doi.org/10.1371/journal.pone.0075091

12. Bridel C, van Wieringen WN, Zetterberg H et al (2019) Diagnostic value of cerebrospinal fluid neurofilament light protein in neurology: a systematic review and meta-analysis. JAMA Neurol 76(9):1035–1048. https://doi.org/10.1001/jamaneurol.2019.1534

13. Kuhle J, Regeniter A, Leppert D, Mehling M, Kappos L, Lindberg RL, Petzold A (2010) A highly sensitive electrochemiluminescence immunoassay for the neurofilament heavy chain protein. J Neuroimmunol 220(1–2):114–119. https://doi.org/10.1016/j.jneuroim.2010.01.004

14. Kušnierová P, Zeman D, Hradílek P, Čábal M, Zapletalová O (2019) Neurofilament levels in patients with neurological diseases: a comparison of neurofilament light and heavy chain levels. J Clin Lab Anal 33(7):e22948. https://doi.org/10.1002/jcla.22948

15. Rossi D, Volanti P, Brambilla L, Colletti T, Spataro R, La Bella V (2018) CSF neurofilament proteins as diagnostic and prognostic biomarkers for amyotrophic lateral sclerosis. J Neurol 265(3):510–521. https://doi.org/10.1007/s00415-017-8730-6

16. Ohya J, Chikuda H, Kato S, Hayakawa K, Oka H, Takeshita K, Tanaka S, Ogata T (2015) Elevated levels of phosphorylated neurofilament heavy subunit in the cerebrospinal fluid of patients with lumbar spinal stenosis: preliminary findings. Spine J 15(7):1587–1592. https://doi.org/10.1016/j.spinee.2015.03.013

17. Takahashi H, Aoki Y, Nakajima A et al (2018) Axonal damage is remarkable in patients with acutely worsening symptoms of compression myelopathy: biomarkers in cerebrospinal fluid samples. Eur Spine J 27(8):1824–1830. https://doi.org/10.1007/s00586-018-5549-5

18. Poesen K, De Schaepdryver M, Stubendorff B et al (2017) Neurofilament markers for ALS correlate with extent of upper and lower motor neuron disease. Neurology 88(24):2302–2309. https://doi.org/10.1212/WNL.0000000000004029

19. De Schaepdryver M, Jeromin A, Gille B, Claeys KG, Herbst V, Brix B, Van Damme P, Poesen K (2018) Comparison of elevated phosphorylated neurofilament heavy chains in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 89(4):367–373. https://doi.org/10.1136/jnnp-2017-316605

20. Steinacker P, Feneberg E, Weishaupt J et al (2016) Neurofilaments in the diagnosis of motoneuron diseases: a prospective study on 455 patients. J Neurol Neurosurg Psychiatry 87(1):12–20. https://doi.org/10.1136/jnnp-2015-311387

21. Gaiani A, Martinelli I, Bello L et al (2017) Diagnostic and prognostic biomarkers in amyotrophic lateral sclerosis: neurofilament light chain levels in definite subtypes of disease. JAMA Neurol 74(5):525–532. https://doi.org/10.1001/jamaneurol.2016.5398

22. Falzone YM, Domi T, Agosta F et al (2020) Serum phosphorylated neurofilament heavy-chain levels reflect phenotypic heterogeneity and are an independent predictor of survival in motor neuron disease. J Neurol 267(8):2272–2280. https://doi.org/10.1007/s00415-020-09838-9

23. Ludolph A, Drory V, Hardiman O, Nakano I, Ravits J, Robberecht W, Shefner J, WFN Research Group On ALS/MND (2015) A revision of the El Escorial criteria 2015. Amyotroph Lateral Scler Frontotemporal Degener. 16(56):291–292. https://doi.org/10.3109/21678421.2015.1049183

24. Kimura F, Fujimura C, Ishida S, Nakajima H, Furutama D, Uehara H, Shinoda K, Sugino M, Hanafusa T (2006) Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. Neurology 66(2):265–267. https://doi.org/10.1212/01.wnl.0000194316.91908.8a

25. Lee SH, Kim KT, Suk KS, Lee JH, Shin JH, So DH, Kwack YH (2010) Asymptomatic cervical cord compression in lumbar spinal stenosis patients: a whole spine magnetic resonance imaging

study. Spine (Phila Pa). 35(23):2057–2063. https://doi.org/10.1097/BRS.0b013e3181f4588a

26. Schönström N, Willén J (2001) Imaging lumbar spinal stenosis. Radiol Clin N Am 39(1):31–53. https://doi.org/10.1016/s0033-8389(05)70262-1

27. Halbgebauer S, Steinacker P, Verde F, Weishaupt J, Oeckl P, von Arnim C, Dorst J, Feneberg E, Mayer B, Rosenbohm A, Silani V, Ludolph AC, Otto M (2022) Comparison of CSF and serum neurofilament light and heavy chain as differential diagnostic biomarkers for ALS. J Neurol Neurosurg Psychiatry 93(1):68–74. https://doi.org/10.1136/jnnp-2021-327129

28. Lu CH, Petzold A, Topping J et al (2015) Plasma neurofilament heavy chain levels and disease progression in amyotrophic lateral sclerosis: insights from a longitudinal study. J Neurol Neurosurg Psychiatry 86(5):565–573. https://doi.org/10.1136/jnnp-2014-307672

29. Wilke C, Pujol-Calderón F, Barro C et al (2019) Correlations between serum and CSF pNfH levels in ALS, FTD and controls: a comparison of three analytical approaches. Clin Chem Lab Med 57(10):1556–1564. https://doi.org/10.1515/cclm-2019-0015

30. Schreiber S, Spotorno N, Schreiber F et al (2018) Significance of CSF NfL and tau in ALS. J Neurol 265(11):2633–2645. https://doi.org/10.1007/s00415-018-9043-0

31. Menke RA, Gray E, Lu CH, Kuhle J, Talbot K, Malaspina A, Turner MR (2015) CSF neurofilament light chain reflects corticospinal tract degeneration in ALS. Ann Clin Transl Neurol 2(7):748–755. https://doi.org/10.1002/acn3.212