Serum Neurofilament Light Trajectories and Their Relation to Subclinical Radiological Disease Activity in Relapsing Multiple Sclerosis Patients in the APLIOS Trial

Amit Bar-Or · Xavier Montalban · Xixi Hu · Harald Kropshofer ·

Petra Kukkaro · Neva Coello · Inga Ludwig · Roman Willi ·

Martin Zalesak · Krishnan Ramanathan · Bernd C. Kieseier ·

Dieter A. Häring · Morten Bagger · Edward Fox

ABSTRACT

Introduction: Several studies have described prognostic value of serum neurofilament light chain (sNfL) at the group level in relapsing multiple sclerosis (RMS) patients. Here, we aimed to explore the temporal association between sNfL and development of subclinical disease activity as assessed by magnetic resonance imaging (MRI) at the group level and evaluate the potential of sNfL as a biomarker for capturing subclinical disease activity in individual RMS patients.
Methods: In the 12-week APLIOS study, patients ($N = 284$) received subcutaneous ofatumumab 20 mg. Frequent sNfL sampling (14 time points over 12 weeks) and monthly MRI scans enabled key analyses including assessment of the group-level temporal relationship of sNfL levels with on-study subclinical development of gadolinium-enhancing (Gd +)T1 lesions. Prognostic value of baseline sNfL ("high" vs. "low") level for subsequent on-study clinical relapse or Gd + T1 activity was assessed. Individual patient-level development of on-study Gd + T1 lesions was compared across three predictors: baseline Gd + T1 lesion number, baseline sNfL ("high" vs. "low"), and time-matched sNfL.
Results: In patients developing Gd + T1 lesions at week 4 (absent at baseline), sNfL levels increased during the month preceding the week-4 MRI scan and then gradually decreased back to baseline. High versus low baseline sNfL conferred

A. Bar-Or (✉)
Center for Neuroinflammation and Experimental Therapeutics and Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
e-mail: amitbar@pennmedicine.upenn.edu

X. Montalban
Department of Neurology-Neuroimmunology and Centre d'Esclerosi Múltiple de Catalunya (Cemcat), Hospital Universitari Vall d'Hebron, Barcelona, Spain

X. Hu
Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA

H. Kropshofer · P. Kukkaro · N. Coello · I. Ludwig · R. Willi · M. Zalesak · K. Ramanathan · B. C. Kieseier · D. A. Häring · M. Bagger
Novartis Pharma AG, Basel, Switzerland

B. C. Kieseier
Department of Neurology, Medical Faculty, Heinrich-Heine University, Dusseldorf, Germany

E. Fox
Central Texas Neurology Consultants and Dell Medical School, The University of Texas at Austin, Round Rock, Austin, TX, USA

increased risk of subsequent on-study clinical relapse or Gd + T1 activity (HR, 2.81; $p < 0.0001$) in the overall population and, notably, also in the patients without baseline Gd + T1 lesions (HR, 2.48; $p = 0.0213$). Individual patient trajectories revealed a marked difference in Gd + T1 lesions between patients with the ten highest vs. lowest baseline sNfL levels (119 vs. 19 lesions). Prognostic value of baseline or time-matched sNfL for on-study Gd + T1 lesions was comparable to that of the number of baseline MRI Gd + T1 lesions.

Conclusions: sNfL measurement may have utility in capturing and monitoring subclinical disease activity in RMS patients. sNfL assessments could complement regular MRI scans and may provide an alternative when MRI assessment is not feasible.

ClinicalTrials.gov: NCT03560739.

Classification of Evidence: This study provides class I evidence that serum neurofilament light may be used as a biomarker for monitoring subclinical disease activity in relapsing multiple sclerosis patients, as shown by its elevation in the weeks preceding the development of new gadolinium-enhancing T1 lesion activity.

Keywords: Biomarker; Ofatumumab; Relapsing multiple sclerosis; Serum neurofilament light chain; Subclinical disease activity

Key Summary Points

Why carry out this study?

Multiple studies have demonstrated that serum neurofilament light chain (sNfL) levels are on average abnormally increased in untreated multiple sclerosis (MS) patients compared to controls, mainly driven by focal inflammatory central nervous system activity, and are reduced by effective therapies. However, insights are lacking into the kinetics of sNfL abnormalities in relation to newly developing disease activity, the potential use of sNfL to capture sub-clinical disease activity, and how sNfL measurements may complement magnetic resonance imaging (MRI) scans in clinical practice in order to monitor disease activity

The present study explored both the temporal association between sNfL and development of subclinical disease activity as assessed by MRI at the group level and evaluated the potential of sNfL as a biomarker for capturing newly developing subclinical disease activity in individual relapsing MS (RMS) patients

What was learned from the study?

Our findings indicate that sNfL may have utility in monitoring newly developing subclinical disease activity in RMS patients as shown by its temporal association with new gadolinium-enhancing (Gd +) T1 lesion activity, its overall comparable prognostic value to baseline MRI, and its added value over MRI in prognostic value of on-study Gd + T1 lesion formation in patients free of brain MRI disease activity at baseline

Overall, our results suggest that when frequent MRI scans are not feasible, serial sNfL measurements may provide an attractive alternative to MRI monitoring if a standardized test becomes available to physicians

INTRODUCTION

Multiple sclerosis (MS) is a chronic inflammatory, demyelinating, autoimmune disorder of the central nervous system, leading to accumulation of disability. The clinical disease course of MS is heterogeneous, and much of the disease is subclinical [1]. The monitoring and detection of suboptimal response to treatment, prognosis of clinical outcome, and therapeutic decision-making in individual patients are key challenges in the management of MS and are currently assessed only with clinical and magnetic resonance imaging (MRI) measures.

Neurofilament light chain (NfL) is a cytoskeletal protein exclusively expressed by neurons and their axons, and its release into the cerebrospinal fluid (CSF) and blood is a highly specific sign of neuroaxonal injury [2–4]. A

strong correlation has been demonstrated between CSF and serum/plasma levels of NfL as a blood biomarker for monitoring neuronal injury and treatment response in MS [2, 3]. At the group level, elevated NfL concentrations at baseline are associated with subsequent on-study relapses, MRI lesions, brain volume loss (BVL), spinal cord atrophy, and disability worsening [2, 4]. While such studies have provided evidence supporting serum NfL (sNfL) as a potential biomarker of both current and future disease activity in patients with relapsing MS (RMS) [5], investigations of the temporal association between sNfL and the development of subclinical disease activity as assessed by MRI and evaluation of the potential of sNfL as a biomarker for capturing subclinical disease activity in individual RMS patients are needed.

Ofatumumab, a fully human anti-CD20 monoclonal antibody, with a monthly 20 mg subcutaneous (SC) dosing regimen, demonstrated superior efficacy versus teriflunomide 14 mg oral once daily and a favorable safety profile in patients with RMS in the phase 3 ASCLEPIOS I and II trials. Ofatumumab significantly lowered sNfL levels versus teriflunomide from the first post-baseline assessment at month 3 through 24 [6].

The 12-week, open-label, phase 2 APLIOS study in RMS [7] incorporated frequent blood sampling (14 time points) and monthly MRI scans. This design offered a unique opportunity to study sNfL trajectories, explore the temporal association between sNfL levels and subclinical disease activity, and evaluate the potential of sNfL as a biomarker for capturing and monitoring subclinical disease activity in individual patients with RMS.

MATERIALS AND METHODS

APLIOS Study Design and Patient Population

The APLIOS study was a 12-week, randomized, open-label, parallel-group, multicenter, phase 2 study in RMS (NCT03560739) that demonstrated the bioequivalence of two formulations: a subcutaneous injection (20 mg) pre-filled

syringe and SensoReady® auto-injector pen. In total, 284 patients received ofatumumab 20 mg (0.4 ml) subcutaneous injections on days 1, 7, and 14 (initial doses) and thereafter every 4 weeks from week 4 onwards (subsequent doses) [8].

Eligibility criteria at screening included an age of 18–55 years; a diagnosis of RMS; an Expanded Disability Status Scale score of 0–5.5; ≥ 1 relapse in the year before screening or ≥ 2 relapses in the 2 years before screening or a positive gadolinium-enhancing (Gd +) scan during the year before randomization; and neurologically stable disease in the month prior to randomization. Further details of the study design and patient population are reported elsewhere [7].

sNfL Sampling and Assay Characteristics

Blood sampling for sNfL measurement was done in the full study population following the sampling scheme for the evaluation of pharmacokinetics for the bioequivalence assessment at 14 time points over 12 weeks, i.e., at baseline/pre-treatment period and on study days 1, 4, 7, 14, 28, 42, 56, 57, 59, 63, 70, 77, and 84 (Fig. 1) [8].

Samples were analyzed using the Siemens sNfL assay, on ADVIA Centaur®, a novel, validated assay. Briefly, the ADVIA Centaur NfL assay is an automated two-site sandwich immunoassay using direct chemiluminometric technology. The assay uses two mouse monoclonal antibodies (mAb) that bind to the NfL, each recognizing a unique epitope. The antibodies are licensed from Quanterix and thus have the same specificity as antibodies in the Simoa NF-light™ assay. The solid phase reagent contains a biotinylated anti-NfL mAb, coupled to a paramagnetic particle conjugated with streptavidin. The Lite reagent contains a monoclonal anti-NfL antibody conjugated with acridinium ester for chemiluminescent detection. The accumulated light signal is directly related to the sample NfL concentration. This assay uses a full 7-point standard curve for calibration. The analytical sensitivity of LoB was determined to be 0.89 pg/ml, LoD was 1.49 pg/
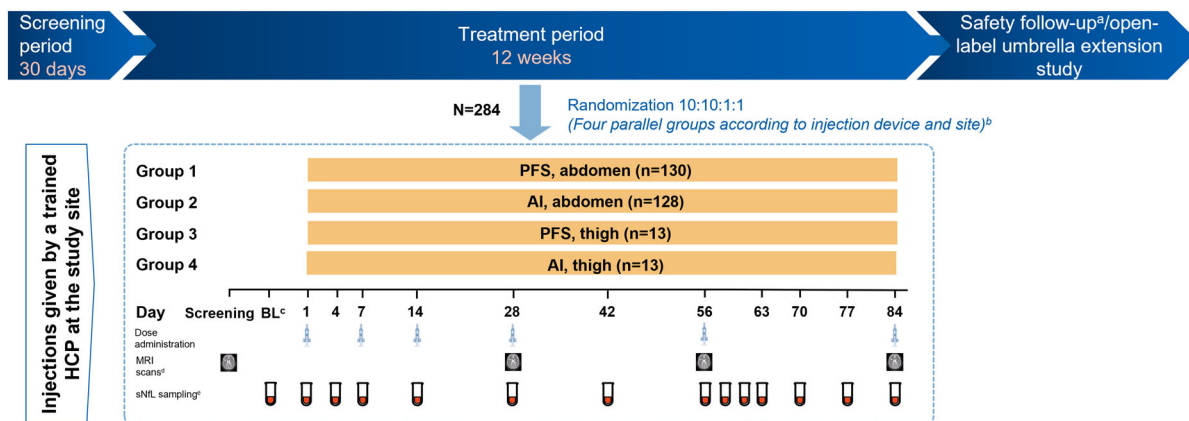
Fig. 1 APLIOS study design with sNfL sampling time points. [a]Nine months or until the B cells returned to their baseline value or to LLN; [b]randomization was stratified by body weight (< 60 kg, 60–90 kg, and > 90 kg); [c]day 7 to day -1; [d]in case of relapse, if an MRI for efficacy was to be scheduled within 14 days of the initiation of steroid treatment, the MRI was performed before steroid treatment was initiated. No MRI for efficacy was performed while a patient was on steroid therapy for relapse and within 14 days upon termination of steroid therap. [c]On non-injection visits the PK (Plasma) and neurofilament samples were drawn at any time during the visit. *AI* autoinjector; *BL* baseline; *HCP* healthcare professional; *LLN* lower limit normal; *MRI* magnetic resonance imaging; *PFS* prefilled syringe; *PK* pharmacokinetics; *sNfL* serum neurofilament light chain

ml, and LLoQ was 1.85 pg/ml. The linearity was established from 1 to 646 pg/ml in the serum. The linear regression result was $R^2 = 0.996$ ($p < 0.001$) using JMP (version 14.0, SAS Institute).

MRI Assessments and Analyses

MRI scans, including assessment of Gd + T1 lesions, were performed at baseline and weeks 4, 8, and 12, with timing of the baseline scans ranging from 57 days to 1 day prior to treatment initiation. Gd + T1 lesions from scans collected within 14 days after termination of steroid therapy were excluded from the analyses.

Statistical Analyses

Temporal Association Between Serial sNfL Levels and Gd + T1 Lesions

The temporal association of serial sNfL levels and MRI activity was analyzed in patients with no Gd + T1 lesions at baseline and with no missing scans and separately for patients with Gd + T1 lesions on week 4 scan only and patients with Gd + T1 lesions on week 4 and 8 scans. For each case, an additive mixed model was fitted to allow a qualitative assessment of group-level sNfL over time. The model was adjusted for age and log of baseline sNfL levels and included a nonparametric smoother of their time, random intercept, and random slope.

sNfL over Time Versus Disease Activity and Individual Trajectories

Baseline sNfL was defined as the last assessment on or before day 1, and the median baseline sNfL level across all patients was 9.1 pg/ml. The sNfL concentrations were analyzed using spaghetti plots, and patients were grouped by (1) sNfL during the study showing values constantly above baseline sNfL median ($N = 60$), (2) sNfL during the study showing both values above and below ("crossed baseline median") ($N = 146$), and (3) sNfL during the study showing values constantly below the baseline sNfL median ($N = 78$). For the above-defined longitudinal sNfL groups, group-level summaries of confirmed clinical relapses (according to pre-specified criteria [6]), Gd + T1 lesions, and the

proportion of patients with no evidence of disease activity (NEDA-3) were assessed. NEDA-3 was defined as no confirmed clinical relapses, no Gd + T1 lesions on any post-baseline scans and no new/enlarging T2 lesions (relative to baseline) on any post-baseline scans, and no 3-month confirmed disability worsening. The individual sNfL trajectories and disease activities of patients with the ten lowest and ten highest baseline sNfL levels were also presented. Their total numbers of confirmed clinical relapses and Gd + T1 lesions were summarized.

Prognostic Value of Baseline sNfL Levels for On-Study Disease Activity

We dichotomized cohorts as baseline sNfL level above median: "high" (> 9.1 pg/ml) or below median: "low" ($\leq$ 9.1 pg/ml) to analyze the risk of on-study disease activity (confirmed clinical relapses or Gd + T1 lesions) using a Cox regression model of time since treatment start to the first confirmed clinical relapse or first observation of on-study Gd + T1 lesions, whichever was earlier. The analysis was performed in all patients and in the subgroup of patients free of Gd + T1 lesions at baseline. For the analysis of all patients, the model included baseline sNfL category as factor ("high" vs. "low") and age, baseline number of Gd + T1 lesions, baseline volume of T2 lesions, baseline EDSS, and number of relapses in previous year as continuous covariates. For the analysis including only patients free of Gd + T1 lesions at baseline, the model included baseline sNfL category as factor and age, baseline volume of T2 lesions, baseline EDSS, and number of relapses in previous year as continuous covariates.

Patient-level Identification of On-Study Gd + T1 Lesion Activity

Using area under the receiver-operating characteristics curve (ROC AUC), we assessed the accuracy of prediction of whether individual patients would have on-study Gd + T1 lesions on each of week 4, 8, and 12 scans, using logistic regression models. For each of the three post-baseline scans, the development of Gd + T1 lesions was compared across three models:

(1) model "baseline Gd + T1" used the number of Gd + T1 lesions on the baseline scan as the predictor, (2) model "baseline sNfL" used the baseline sNfL measure (> 9.1 pg/ml vs. $\leq$ 9.1 pg/ml) as the predictor, and (3) model "time-matched sNfL" used the sNfL level on the MRI scan day (or if unavailable, the last sNfL measurement before the scan day; > 9.1 pg/ml vs. $\leq$ 9.1 pg/ml) as the predictor. All models were adjusted for age.

AUCs for the development of week 4, 8, and 12 Gd + T1 lesions were calculated from ten-fold cross-validation. Confidence intervals (CIs) of AUCs were obtained through 10,000 bootstrap resampling iterations.

Standard Protocol Approvals, Registrations, and Patient Consents

The study was conducted in accordance with the International Conference on Harmonisation guidelines for Good Clinical Practice [9] and the principles of the Declaration of Helsinki [10]. The protocol was approved by an independent ethics committee or institutional review board for each study site, and all patients provided written informed consent before commencing trial-related procedures. This article is based on previously conducted studies and does not contain any new studies with human participants or animals performed by any of the authors.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

RESULTS

Patient Demographics and Baseline Characteristics

In total, 284 patients were included in the study. At baseline, patients in the high baseline sNfL (> 9.1 pg/ml) group versus those in the low baseline sNfL ($\leq$ 9.1 pg/ml) group were found to

Table 1 Patient demographics and baseline disease characteristics by baseline sNfL low ($\leq$ 9.1 pg/ml) vs. high (> 9.1 pg/ml)

Characteristics	Low sNfL (N = 146)	High sNfL (N = 138)	All patients (N = 284) [7]
Age, years	36.1 ± 8.12	38.6 ± 9.56	37.3 ± 8.92
Sex, female, n (%)	102 (69.9)	97 (70.3)	199 (70.1)
Race, white, n (%)	145 (99.3)	130 (94.2)	275 (96.8)
Weight, kg	76.55 ± 17.81	70.69 ± 18.55	73.70 ± 18.38
BMI, kg/m^2	26.42 ± 5.65	24.58 ± 6.48	25.53 ± 6.13
MS duration since first symptom, years	8.40 ± 7.15	10.19 ± 8.28	9.27 ± 7.75
No. of relapses in the year before the study	1.3 ± 0.70	1.4 ± 0.73	1.3 ± 0.72
No. of relapses in the month 12 to month 24 before the study	0.9 ± 1.03	1.1 ± 2.01	1.0 ± 1.58
EDSS score	2.82 ± 1.32	3.08 ± 1.26	2.95 ± 1.30
No. of Gd + T1 lesions	0.5 ± 2.55	2.6 ± 6.47	1.5 ± 4.97
Patients with Gd + T1 lesions, n (%)	31 (21.2)	76 (55.1)	107 (37.7)
Patients with previous DMTs, n (%)	103 (70.5)	91 (65.9)	194 (68.3)

Data are presented as mean ± SD, unless otherwise specified

BMI body mass index; *DMTs* disease-modifying therapies; *EDSS* Expanded Disability Status Scale; *Gd +* gadolinium enhancing; *MS* multiple sclerosis; *sNfL* serum neurofilament light chain

be slightly older (38.6 vs. 36.1 years) and had a lower BMI (24.58 vs. 26.42 kg/m^2), a longer disease history (10.19 vs. 8.40 years), and a higher number of Gd + T1 lesions (2.6 vs. 0.5) (Table 1).

Temporal Association of sNfL Levels and Gd + T1 Lesions

The unique approach of frequent blood sampling and imaging in APLIOS enabled us to assess the temporal relationship between sNfL and the development of subclinical new disease activity, assessed as new Gd + T1 lesions. We observed that in patients without Gd + T1 lesions at baseline who developed one or more Gd + T1 lesion/s at week 4, the group-level average sNfL levels increased during the month before the week 4 MRI scan, reached maximum levels over a few weeks, and then gradually decreased to baseline levels (Fig. 2A). This was also the case in patients who had no Gd + T1

lesions at baseline and were observed to develop Gd + T1 lesions at both week 4 and 8 scans (most likely lesions persisting for > 4 weeks after they were first observed) (Fig. 2B), except that in this group, sNfL did not decrease to baseline levels by week 12. Analysis that accounts further for sex was also considered and the findings remain the same.

sNfL Concentrations over Time and Disease Activity

At the group level, ofatumumab treatment was associated with a consistent decline in sNfL levels from baseline up to 12 weeks of study treatment (Supplementary Material; data on file), similar to the trend observed in the pivotal phase 3 studies [6].

We first considered whether sNfL levels over time were associated with measures of disease activity, as previously reported [2, 4], by dividing patients into three groups based on their
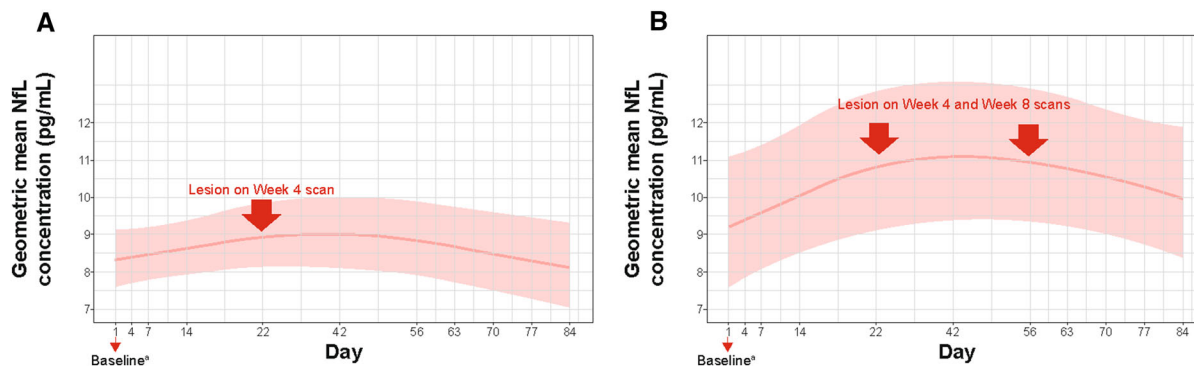
Fig. 2 Temporal association of sNfL levels and Gd + T1 lesions in patients free of Gd + T1 lesions at baseline: **A** patients (*N* = 13) Who only had Gd + T1 lesions on the week 4 scan and **B** patients (*N* = 7) who had Gd + T1 lesions on both week 4 and week 8 scans.

Additive mixed model with a nonparametric smoother of time adjusted for age and log of baseline sNfL. [a]Patients were free of Gd + T1 lesions at baseline. *Gd* + gadolinium enhancing; *sNfL* serum neurofilament light chain

longitudinal sNfL levels. sNfLs over time by longitudinal sNfL group are presented as spaghetti plots in Fig. 3. We found that the proportion of patients with confirmed clinical relapses was higher in the group of patients with sNfL levels constantly above the baseline median (9/60; 15%) versus groups with sNfL levels that crossed the baseline median (4/146; 2.7%) or constantly below the baseline median (2/78; 2.6%).

Over time, the proportion of patients with Gd + T1 lesions was also higher in the group with sNfL levels above the baseline median versus those with sNfL levels that crossed the baseline median or below the baseline median, with 64.7% vs. 31.1% and 16.2% at week 4, 25.9% vs. 10.8% and 9.0% at week 8, and 11.9% vs. 4.4% and 3.9% at week 12, respectively.

The mean number of Gd + T1 lesions by longitudinal sNfL group over 12 weeks is

Gd+ T1 lesions	Week 4	Week 8	Week 12
Above baseline median	2.29 ± 3.60	0.63 ± 2.35	0.36 ± 1.23
Crossed baseline median	0.61 ± 1.23	0.17 ± 0.61	0.05 ± 0.25
Below baseline median	0.24 ± 0.59	0.12 ± 0.43	0.07 ± 0.38
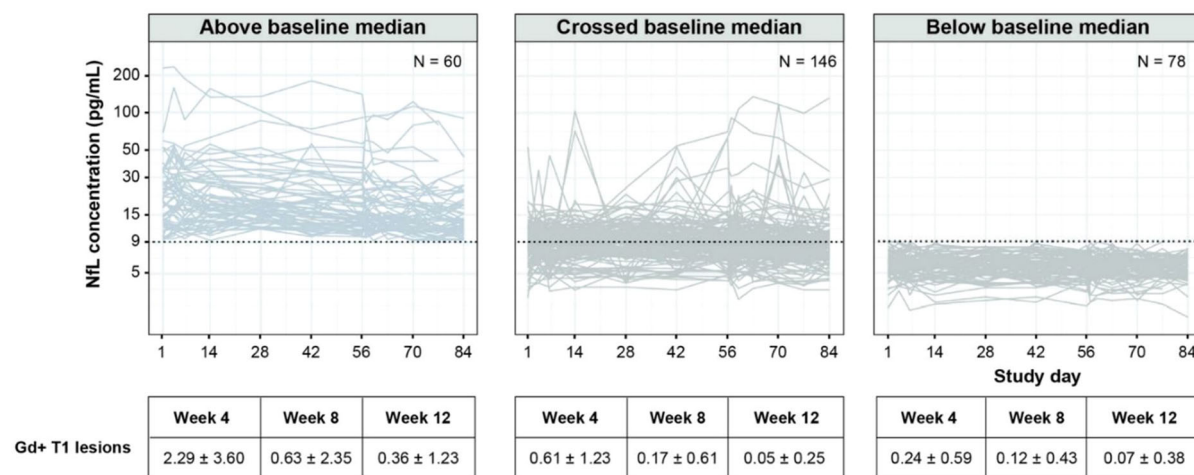
Fig. 3 sNfL over time versus disease activity and individual trajectories. Dotted line: baseline median (9.1 pg/ml). sNfL, serum neurofilament light chain. Day 1 represents baseline. The group 'above baseline median' includes patients whose serial sNfL values were all > baseline median; the group 'below baseline median' includes patients whose serial sNfL values were all ≤ baseline median; the group 'crossed baseline median' includes the rest of the patients

presented in Table 2. As shown, the mean number of Gd + T1 lesions over time was higher in the group of patients with sNfL levels above the baseline median versus those with sNfL levels that crossed or were below the baseline median.

The proportion of patients with NEDA-3 assessed by longitudinal sNfL subgroup was higher in patients with below baseline median sNfL (51/78; 65.4%) compared to those with above (13/60; 21.7%) or crossing the baseline sNfL median levels (73/146; 50.0%; Fig. 4).

To illustrate individual patient sNfL levels over the course of the study and the associated relapses and Gd + T1 lesions, sNfL trajectories in patients with the ten lowest and ten highest baseline sNfL levels are presented in Fig. 5A and 5B, respectively.

Gd + T1 lesions and clinical relapses were more frequently observed over the course of the study in patients with the highest baseline sNfL levels than in those with the lowest baseline sNfL levels. Over time (including baseline), in patients with the ten highest versus ten lowest baseline sNfL levels, seven versus four patients reported at least one Gd + T1 lesion and seven vs. one patients had a scan with at least two lesions. The total number of Gd + T1 lesions was notably higher in patients with the ten highest vs. ten lowest baseline sNfL levels (116 vs. 19); in patients with the ten lowest baseline sNfL levels, most of the lesions were from one patient. One notes that due to the frequent MRI scans, the total number may include lesions observed in more than one scan. Two clinical relapses were observed among those ten patients with highest baseline sNfL and none in patients with the ten lowest baseline sNfL levels.

Prognostic Value of Baseline sNfL Levels for On-Study Disease Activity

In both the overall patient population and the subset of patients who were free of Gd + T1 lesions at baseline, we confirmed that baseline sNfL level had a prognostic value for on-study disease activity (confirmed clinical relapses or Gd + T1 lesions). Specifically, high baseline sNfL levels conferred an increased risk of on-study confirmed clinical relapses or Gd + T1 lesions versus low baseline sNfL levels (hazard ratio [HR], 2.81 [95% CI: 1.78, 4.42]; $p < 0.0001$; Fig. 6A). Among the patients free of Gd + T1 lesions at baseline ($N = 176$), the high baseline sNfL levels also predicted an increased risk of on-study confirmed clinical relapses or Gd + T1 lesions compared with low baseline sNfL levels (HR, 2.48 [95% CI: 1.15, 5.39]; $p = 0.0213$) (Fig. 6B).

Analysis that accounts further for sex was also considered, and the findings remain the same.

An important question relates to the ability of sNfL, measured in an individual patient, to capture the presence of subclinical disease activity. Using ROC curves, we observed that the identification of on-study Gd + T1 lesions based on sNfL levels in individual patients, either at baseline or time matched, was comparable in performance to the prediction based on the number of Gd + T1 lesions on the baseline MRI scan, an observation which was consistent at all time points over 12 weeks

Table 2 Number of Gd + T1 lesions by longitudinal sNfL group

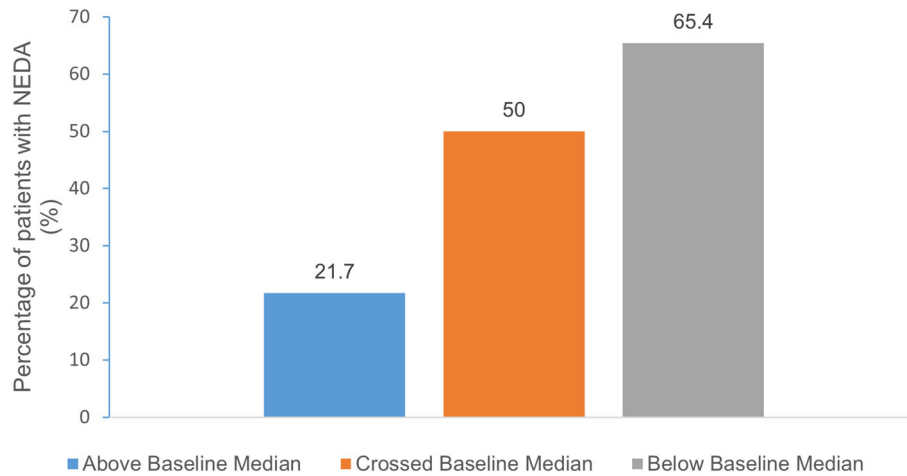| Visit (mean ± SD)	Above baseline median ($N = 60$)	Below baseline median ($N = 78$)	Crossed baseline median ($N = 146$)	All patients ($N = 284$)
Baseline	4.88 ± 9.10	0.24 ± 0.68	0.82 ± 2.76	1.54 ± 4.98
Week 4	2.29 ± 3.60	0.24 ± 0.59	0.61 ± 1.23	0.83 ± 1.98
Week 8	0.63 ± 2.35	0.12 ± 0.43	0.17 ± 0.61	0.25 ± 1.17
Week 12	0.36 ± 1.23	0.07 ± 0.38	0.05 ± 0.25	0.12 ± 0.64

Fig. 4 Proportion of patients with NEDA by longitudinal sNfL subgroup
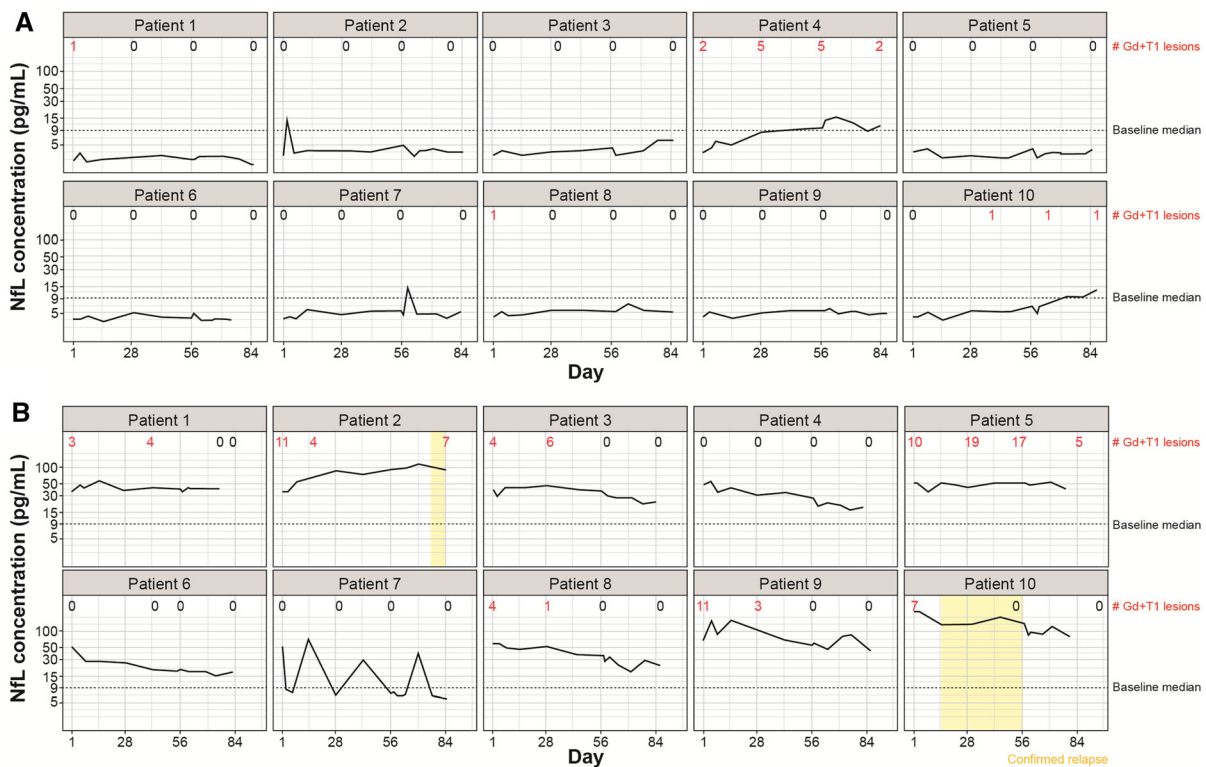

Fig. 5 Individual patient trajectories: **A** Patients with the ten lowest baseline sNfL levels and **B** patients with the ten highest baseline sNfL levels. Dotted line: baseline median (9.1 pg/ml). The number of Gd + T1 lesions annotated refers to those visible on brain MRI scans. Additional lesions may have occurred outside the field of view, e.g., in the spinal cord. *Gd +* gadolinium-enhancing; *sNfL* serum neurofilament light chain

(Fig. 7). We note that the AUCs are compatible within the estimated confidence intervals, and the ROC curves are particularly close in the high-sensitivity range, which is of most interest to identify patients at risk. Analysis that
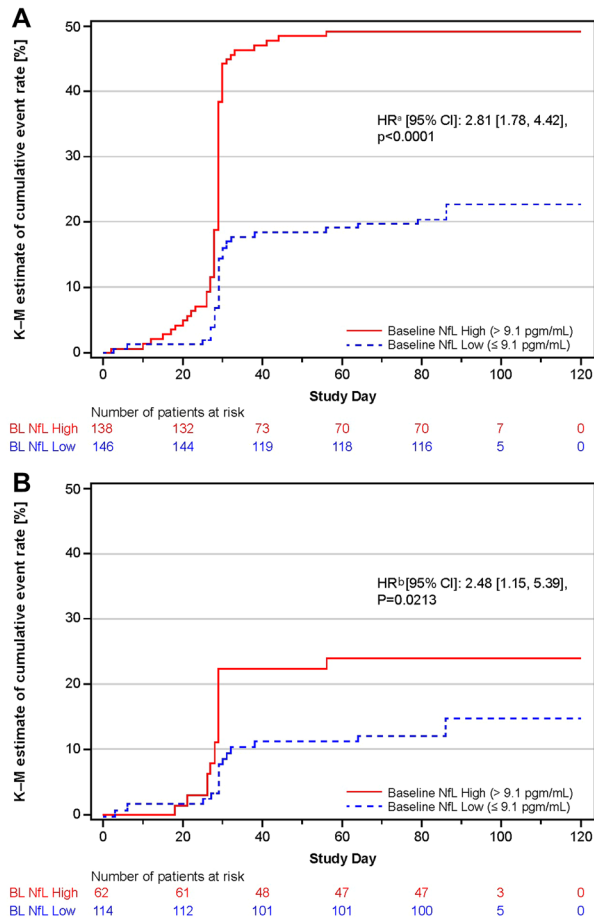
Fig. 6 Baseline sNfL level predicts risk of on-study disease activity: **A** All patients and **B** patients free of Gd + T1 lesions at baseline. The stepped lines show Kaplan-Meier estimates of the cumulative proportion of patients with either confirmed clinical relapses or Gd + T1 lesions by time (days). In the figure, the steep increase in the estimated proportion of patients with on-study confirmed clinical relapses or Gd + T1 lesions at day 28 was observed because day 28 is the day of the first-post baseline MRI scan, uncovering the proportion of patients with radiological disease activity. [a]The Cox regression model included the baseline sNfL category as factor, age, baseline number of Gd + T1 lesions, baseline volume of T2 lesions, baseline EDSS, and number of clinical relapses in previous year as continuous covariates. [b]The Cox regression model included baseline sNfL category as factor and age, baseline volume of T2 lesions, baseline EDSS and number of clinical relapses in the previous year as continuous covariates. *Gd* + gadolinium enhancing, *HR* hazard ratio, *sNfL* serum neurofilament light chain

accounts further for sex was also considered, and the findings remain the same.

DISCUSSION

APLIOS was an open-label study with all patients initiating treatment with ofatumumab, which is known to substantially limit new focal MS disease activity [6]. In keeping with multiple prior studies reporting longitudinal reductions in sNfL following initiation of most approved MS treatments [4, 6, 11–16], ofatumumab treatment was associated with a consistent decline in sNfL levels from baseline up to 12 weeks [17].

Confirming prior studies [2–5], we observed that elevated sNfL levels could reflect and predict the development of new disease activity at the group level. The high-frequency MRI and high-frequency sNfL sampling in this study provided the unique opportunity to assess the temporal association between sNfL and disease activity confirmed by subclinical newly developing Gd + T1 lesions. We observed that sNfL levels increased before new lesions were observed on monthly scheduled MRI scans, reached maximal levels over a period of few weeks, and then gradually decreased thereafter to baseline. While one cannot be sure exactly when a new Gd + T1 lesion formed, our findings indicate that the half-life of sNfL is longer than that of the Gd + T1 lesions that contribute to the rise in sNfL. These observations provide new insights into the dynamic changes and evolution of neuroaxonal injury and resulting sNfL measures associated with the development of new focal inflammatory lesions in patients with RMS. They also indicate that sNfL assessment may capture subclinical disease activity that might otherwise be missed without frequent MRI scans.

The unique frequent assessments of sNfL levels and new disease activity in the relatively large number of patients monitored in this study showed that the proportion of patients with NEDA-3 was higher in the group whose
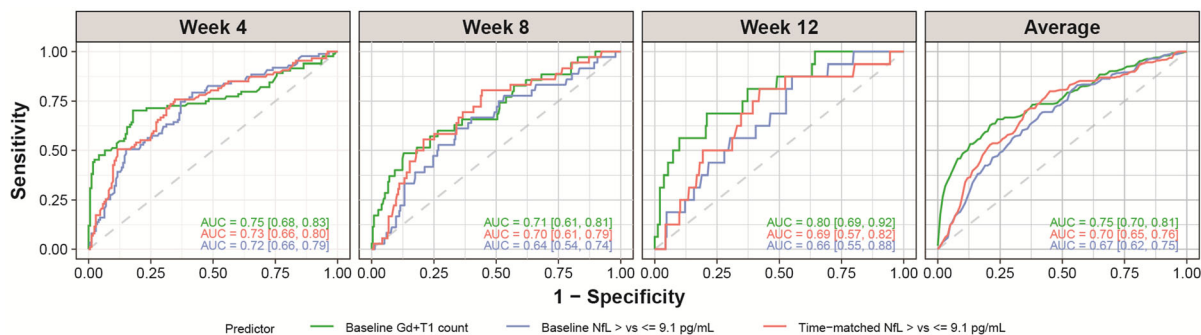
Fig. 7 ROC curves and AUC for the identification of on-study Gd + T1 lesions[a]; 9.1 pg/ml is the median baseline sNfL concentration for all patients in the safety set whose baseline sNfL levels are available. [a]Prediction of whether a patient would have a Gd + T1 lesion on the scan of the specified week using logistic regressions based on the following linear predictors: Model "baseline Gd + T1" (green) used the number of Gd + T1 on the baseline scan as the predictor, model "baseline sNfL" (blue) used the baseline sNfL measure (> 9.1 pg/ml vs. ≤ 9.1 pg/ml) as predictor, and model "time-matched sNfL" (red) used time-matched sNfL (if not available, the last sNfL value before the scan day) (> 9.1 pg/ml vs. ≤ 9.1 pg/ml) as the predictor of whether there would be lesions on the respective scan. All models were adjusted for age. AUCs from week 4, week 8, and week 12 scans were calculated from tenfold cross-validation, and average AUC is the average across the three scans. For each scan and each model, 10,000 bootstrap resampling on the predicted probabilities and the true labels is performed, and a residual bootstrap interval is obtained from the bootstrapped AUCs. *AUC* area under the curve; *Gd +* gadolinium-enhancing; *MRI* magnetic resonance imaging; *ROC AUC* area under the receiver-operating characteristics curve; *sNfL* serum neurofilament light chain

sNfL levels were constantly below the baseline median throughout the study period compared to both patients with longitudinal sNfL levels that crossed the baseline median on one or more occasion, particularly those whose longitudinal sNfL levels were constantly above the baseline median. These results suggest that persistently low longitudinal sNfL levels associate with NEDA-3. As expected, we also observed that baseline sNfL level showed prognostic value for on-study disease activity (confirmed clinical relapses or Gd + T1 lesions) as those patients who had a high baseline sNfL level were at significantly higher risk of developing on-study lesions or confirmed clinical relapses than those who had a low baseline sNfL value. We further noted that sNfL measurements provided added value over MRI among patients who appeared free of Gd + T1 lesions based on MRI scan at baseline.

At the individual patient level, the prognostic value of baseline sNfL for on-study Gd + T1 lesions over 12 weeks was comparable to that of the prognostic value of the number of Gd + T1 lesions on the baseline brain MRI scan, even in a population where patients were on highly efficacious treatment such as ofatumumab. Time-matched sNfL (which utilizes the more frequent sampling) was also assessed but did not substantially improve the prediction compared to baseline sNfL over the 12-week study period. Although not better than baseline sNfL, the use of on-treatment sNfL did not decrease the performance, suggesting the value of a more recent sNfL over a potentially distant "baseline," e.g., in cases where patients have been on treatment for years.

While at the group level low sNfL was associated with fewer Gd + T1 lesions, variability was observed at the individual patient level. Some individual patient trajectories showed high baseline sNfL levels but no baseline or on-study Gd + T1 lesions, which could potentially be explained by recently resolved lesions that developed relatively close to study entry. New subclinical lesions and clinical relapses were nevertheless more frequent over time in patients with the highest compared to those with lowest baseline sNfL levels, indicating prognostic potential of baseline sNfL even in a

population where patients were on highly efficacious treatment. These findings may not be surprising as high baseline sNfL is likely indicative of a higher level of disease activity at baseline, and during the relatively short study duration, patients with a high level of disease activity at baseline are more likely to manifest new disease activity post-baseline. One notes that patients enrolled in this study had to have documented disease activity sometime within 2 years (see "Methods"), and the half-life of NfL is likely at least 6–8 weeks [18]. High baseline sNfL levels might also reflect focal inflammatory MS disease activity in regions not covered by brain MRI (e.g., the spinal cord), damage in the normal-appearing white/gray matter, or other sources such as brain trauma [19]. Future studies may help to further understand the source of elevated NfL levels and their relationship to disease activity.

The present analysis, which extends previous studies [2, 6, 13, 20, 21], supports the utility of sNfL in monitoring subclinical disease activity and points to the potential of individual sNfL measures to capture such activity in patients with RMS, as determined by MRI. The analysis indicates that even in RMS patients without evidence of MRI or clinical disease activity, presence of high sNfL levels may capture the development of new disease activity as well as flag patients who may benefit from closer monitoring and/or reconsideration of treatment response.

Together, our findings suggest that, in certain settings, sNfL measurements may have utility in replacing or complementing standard MRI monitoring of subclinical disease activity in individual patients. This may be particularly helpful when serial MRIs are prohibitively costly, burdensome, or otherwise inaccessible to patients. The kinetics and evolution of sNfL trajectories that we observed in relation to new focal inflammatory disease activity in individual patients over this 3-month study indicate that sNfL sampling on a quarterly basis may offer adequate monitoring and surveillance of subclinical disease activity (at least involving the brain).

Limitations of this post hoc analysis of data from the APLIOS study include that this was a single-arm trial with a short study duration, where treatment (ofatumumab) was highly efficacious in suppressing lesion formation and reducing sNfL levels. Further findings from a larger independent cohort study could provide a more complete perspective. Currently, the research is being conducted with different assays and cut-offs and with different endpoints in mind [22]. The present APLIOS study used the median sNfL value at baseline (9.1 pg/ml) as an analysis threshold to divide the recruited RMS patient population into 'high' vs. 'low' groups (without the use of external controls). The data do not allow claims of optimality regarding this choice of analysis threshold. A threshold optimization should be performed in the context of an intended use of an in vitro diagnostic device development for NfL to ensure the utility of such a tool in clinical practice.

CONCLUSIONS

In conclusion, this study suggests that sNfL may have utility in monitoring subclinical disease activity in RMS patients as shown by its temporal association with Gd + T1 lesion activity, its overall comparable prognostic value to baseline MRI, and its added value over MRI in prognostic value of on-study confirmed clinical relapses or Gd + T1 lesions in patients free of brain MRI disease activity. Therefore, when frequent MRI scans are not feasible, sNfL may provide an attractive alternative to MRI if a standardized test becomes available to physicians.

ACKNOWLEDGEMENTS

interpretation, and preparation, review, and approval of the manuscript. The Novartis Pharma AG also sponsored the publication of this manuscript along with its associated rapid service fee.

Medical Writing, Editorial, and Other Assistance. The authors thank Saimithra Thammera and Anuja Shah (Novartis Healthcare Pvt. Ltd., Hyderabad, India) for providing medical writing support, which encompassed writing of the manuscript, formatting, referencing, preparation of tables and figures as per journal guidelines, and incorporating the authors' revisions and finalizing the draft for submission, all under the direction of the authors. All authors edited the manuscript for intellectual content, provided guidance during manuscript development and approved the final version submitted for publication.

Data Availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the work, and have given final approval to the version to be published. All authors are responsible for intellectual content and data accuracy.

Author Contributions. Amit Bar-Or: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Xavier Montalban: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Xixi Hu: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Harald Kropshofer: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Petra Kukkaro: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Neva Coello: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Inga Ludwig: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Roman Willi: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Martin Zalesak: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Krishnan Ramanathan: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Bernd Kieseier: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Dieter A. Häring: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Morten Bagger: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data. Edward Fox: Drafting/revision of the manuscript for content, including medical writing for content; Major role in the acquisition of data; Study concept or design; Analysis or interpretation of data.

Disclosures. Amit Bar-Or has participated as a speaker in meetings sponsored by and

REFERENCES

1. Joy JE, Johnston RB Jr, editors. Committee on multiple sclerosis: current status and strategies for the future, board on neuroscience and behavioral health. Washington: National Academies Press (US); 2001.

2. Kuhle J, Kropshofer H, Haering DA, et al. Blood neurofilament light chain as a biomarker of MS disease activity and treatment response. Neurology. 2019;92(10):e1007–15.

3. Berger T, Stuve O. Neurofilament light chain: an important step toward a disease biomarker in multiple sclerosis. Neurology. 2019;92:1–2.

4. Disanto G, Barro C, Benkert P, et al. Serum Neurofilament light: a biomarker of neuronal damage in multiple sclerosis. Ann Neurol. 2017;81:857–70.

5. Haering DA, Kropshofer H, Kappos L, et al. Long-term prognostic value of longitudinal

measurements of blood neurofilament levels. Neurol Neuroimmunol Neuroinflamm. 2020;7: e856.

6. Hauser SL, Bar-Or A, Cohen JA, et al. Ofatumumab versus teriflunomide in multiple sclerosis. N Engl J Med. 2020;383:546–57.

7. Bar-Or A, Wiendl H, Montalban X, et al. Rapid and sustained B-cell depletion with subcutaneous ofatumumab in relapsing multiple sclerosis: APLIOS, a randomized phase-2 study. Mult Scler. 2021. https://doi.org/10.1177/13524585211044479.

8. Jones B, et al. Statistical methodology for highly variable compounds: a novel design approach for the ofatumumab phase 2 bioequivalence study. Pharm Stat. 2022. https://doi.org/10.1002/pst.2233.

9. Dixon JR Jr. The International Conference on Harmonization Good Clinical Practice guideline. Qual Assur. 1998;6(2):65–74.

10. Declaration of Helsinki: ethical principles for medical research involving human subjects. Fortaleza: World Medical Association; 2013. p. 388.

11. Akgün K, Kretschmann N, Haase R, et al. Profiling individual clinical responses by high-frequency serum neurofilament assessment in MS. Neurol Neuroimmunol Neuroinflamm. 2019;6(3): e555.

12. Sejbaek T, Nielsen HH, Penner N, et al. Dimethyl fumarate decreases neurofilament light chain in CSF and blood of treatment naïve relapsing MS patients. J Neurol Neurosurg Psychiatry. 2019;90(12):1324–30.

13. Jakimovski D, Zivadinov R, Ramanthan M, et al. Serum neurofilament light chain level associations with clinical and cognitive performance in multiple sclerosis: a longitudinal retrospective 5-year study. Mult Scler J. 2019;26:1670–81.

14. de Flon P, Laurell K, Sundstrom P, et al. Comparison of plasma and cerebrospinal fluid neurofilament light in a multiple sclerosis trial. Acta Neurol Scand. 2019;139(5):462–8.

15. Cross A. Ocrelizumab treatment reduced levels of neurofilament light chain and numbers of B cells in the cerebrospinal fluid of patients with relapsing multiple sclerosis in the OBOE study S56.008. Neurology. 2019;92(Supplement 15):52.

16. Thebault S, Tessier D, Lee H, et al. High serum neurofilament light chain normalises after haematopoietic stem cell transplant for MS. Neurol Neuroimmunol Neuroinflamm. 2019;6(5):e598.

17. Data on file. Ofatumumab (OMB157G). Exploratory analysis. COMB157G2102. Adjusted geometric mean NfL concentrations over time. Novartis Pharma AG; 2022.

18. Bergman J, Dring A, Zetterberg H, et al. Neurofilament light in CSF and serum is a sensitive marker for axonal white matter injury in MS. Neurol Neuroimmunol Neuroinflamm. 2016;3(5): e271.

19. McKee AC, Daneshwar DH. The neuropathology of traumatic brain injury. Handb Clin Neurol. 2015;127:45–66.

20. Novakova L, Zetterberg H, Sundström P, et al. Monitoring disease activity in multiple sclerosis using serum neurofilament light protein. Neurology. 2017;89(22):2230–7.

21. Varhaug KN, Barro C, Bjørnevik K, et al. Neurofilament light chain predicts disease activity in relapsing-remitting MS. Neurol Neuroimmunol NeuroInflamm. 2018;5(1): e422.

22. Bittner S, Oh J, Havrdová EK, Tintoré M, Zipp F. The potential of serum neurofilament as biomarker for multiple sclerosis. Brain. 2021. https://doi.org/10.1093/brain/awab241.