Serum neurofilament light chain for individual prognostication of disease activity in multiple sclerosis: a retrospective modelling and validation study

Pascal Benkert[†] PhD, Stephanie Meier[†] MSc, Sabine Schaedelin MSc, Ali Manouchehrinia PhD, Özgür Yaldizli MD, Aleksandra Maceski MSc, Johanna Oechtering MD, Lutz Achtnichts MD, David Conen MD, Tobias Derfuss MD, Patrice H. Lalive MD, Christian Mueller MD, Stefanie Müller MD, Yvonne Naegelin MD, Jorge R. Oksenberg PhD, Caroline Pot MD, Anke Salmen MD, Eline Willemse PhD, Ingrid Kockum PhD, Kaj Blennow PhD, Henrik Zetterberg PhD, Claudio Gobbi MD, Ludwig Kappos MD, Heinz Wiendl MD, Klaus Berger MD, Maria Pia Sormani PhD, Cristina Granziera MD, Fredrik Piehl MD, David Leppert MD, Jens Kuhle* MD for the RDB in the Swiss Multiple Sclerosis Cohort Study Group

[†]Equally contributed

Clinical Trial Unit, Department of Clinical Research (P Benkert PhD, S Schaedelin MSc), **Neurologic Clinic and Policlinic, MS Centre and Research Centre for Clinical Neuroimmunology and Neuroscience Basel (RC2NB)** (S Meier MSc, Ö Yaldizli MD, A Maceski MSc, J Oechtering MD, T Derfuss MD, Y Naegelin MD, E Willemse PhD, L Kappos MD, C Granziera MD, D Leppert MD, J Kuhle MD), **Translational Imaging in Neurology Basel, Department of Medicine and Biomedical Engineering** (C Granziera MD), **Cardiovascular Research Institute Basel, Cardiology Division, Department of Medicine** (C Müller MD),

University Hospital Basel, University of Basel, Basel, Switzerland;

Department of Clinical Neuroscience, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden (A Manouchehrinia PhD, I Kockum PhD, F Piehl MD);

Center for Neurology, Academic Specialist Center, Stockholm Health Services, Stockholm, Sweden (F Piehl MD);

Department of Neurology, Cantonal Hospital Aarau, Aarau, Switzerland (L Achtnichts MD);

Department of Neurology, Bern University Hospital, University of Bern, Bern, Switzerland (A Salmen MD);

Population Health Research Institute, McMaster University, Hamilton, Canada (D Conen MD);

Multiple Sclerosis Center (MSC), Department of Neurology, Neurocenter of Southern Switzerland, Lugano, Switzerland (C Gobbi MD);

Service of neurology, Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland (C Pot MD);

Department of Neurosciences, Division of Neurology, Unit of Neuroimmunology and Neuromuscular Diseases, University Hospital of Geneva and Faculty of Medicine, Geneva, Switzerland (P Lalive MD);

Department of Neurology, Cantonal Hospital St. Gallen, St. Gallen, Switzerland (S Müller MD);

Department of Neurology and Weill Institute for Neurosciences, University of California, San Francisco, United States of America (J Oksenberg PhD);

Faculty of biomedical Sciences, Università della Svizzera Italiana (USI), Lugano, Switzerland (C Gobbi MD);

Clinical Neurochemistry Laboratory, Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska University Hospital, University of Gothenburg, Mölndal, Sweden (K Blennow PhD, H Zetterberg PhD);

Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, and UK Dementia Research Institute at UCL, London, UK (H Zetterberg PhD);

Department of Neurology with Institute of Translational Neurology, University Hospital Münster, Münster, Germany (H Wiendl MD);

Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany (K Berger MD);

University of Genoa, Department of Health Sciences (DISSAL), Genoa, Italy (M P Sormani PhD).

***Corresponding Author**

Jens Kuhle, MD, PhD, Neurology Clinic and Policlinic, MS Centre and Research Centre for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel, University of Basel, Switzerland. Email: jens.kuhle@usb.ch. Fax: +41 (0)612654100. Tel: +41 (0)61 2652525.

Summary

Background

Serum neurofilament light chain (sNfL) is a biomarker of neuronal damage to monitor disease activity and drug response, and to prognosticate disease course in multiple sclerosis (MS) on the *group level*. The lack of representative reference values to correct for the physiologic age-dependent increase of sNfL, has limited so far its diagnostic use for *individual patients*. We aimed at demonstrating the applicability of sNfL to identify individual patients at risk for future disease activity by establishing a reference database (RDB) to derive age- and BMI-corrected reference values. Further, we used the RDB to test the suitability of sNfL as an endpoint for the group level comparison of effectiveness across disease modifying therapies.

Methods

We measured sNfL levels in 10,133 blood samples from 5,390 US and European control persons. We modelled the distribution of sNfL levels in function of physiologic age-related increase and BMI-dependent modulation to derive percentile/Z-score values via a generalized additive model for location, scale and shape (GAMLSS). Based on 7,769 longitudinally collected samples obtained from 1,313 MS patients participating in the Swiss MS Cohort, we compared the association of sNfL Z-scores with clinical and MRI characteristics in view of their respective disease prognostic capacity. In a further step, we validated our findings in an independent sample of 4341 MS patients followed in the Swedish MS registry.

Findings

In controls, sNfL increases exponentially along age with an increased rate above approximately 50 years of age. In MS, sNfL percentiles/Z-scores identify a gradually

increased risk for future acute (relapse, lesion formation) and chronic (disability worsening) disease activity: a sNfL Z-score above 1.5 was associated with a 3·15-fold increased risk of future clinical or MRI disease activity in all MS patients (OR: 3·15; 95% CI 2·35-4·23; p<0·0001) and in those considered stable with NEDA-3 (2·66; 95% CI 1·08-6·55; p=0·034); they outperform absolute raw sNfL cut-off values for diagnostic accuracy. On the group level, the longitudinal course sNfL Z-score values of MS patients decreased to those of controls under monoclonal antibody (alemtuzumab, natalizumab, ocrelizumab, rituximab) and to lesser extent oral therapy (dimethyl fumarate, fingolimod, siponimod, teriflunomide), but remained elevated with platform (interferons, glatiramer acetate) compounds (p<0·0001 for the interaction term between treatment category and treatment duration). Results were fully confirmed in the independent validation cohort.

Interpretation

The use of sNfL percentiles/Z-scores allows identifying individual patients at risk for a detrimental disease course and suboptimal therapy response beyond clinical and MRI measures, specifically in those with NEDA-3 status. Second, sNfL may be used as endpoint for comparing effectiveness across drug classes in pragmatic trials. We provide an internet-based application for the calculation of sNfL percentile/Z-score values enabling the interpretation of individual measurements (under construction).

Introduction

Multiple sclerosis (MS) is a chronic inflammatory and neurodegenerative disease of the central nervous system characterized by acute deterioration of neurological functions (relapses) and chronic accumulation of relapse-independent disability ('progression'). In the past three decades, increasingly effective disease-modifying therapies (DMT) have led to ground-breaking success in suppressing relapse activity, and its MRI correlate, focal brain lesion formation, while the effect on the course of progression has been modest at best.[1] Disease activity-free status or 'no evidence of disease activity-3' (NEDA-3, i.e. absence of relapses, clinically significant increase in EDSS, of new or enlarging T2 weighted (w) lesions, and T1w contrast-enhancing lesions on brain MRI) has become a treatment goal for MS and a new outcome measure in clinical trials.[1–4] However, less than 8% of patients keep NEDA-3 status. Moreover, this was not associated with statistically significant better EDSS outcomes 7 and 8 years later, respectively[4,5] and Cree et al. "call into question the utility of annual MRI assessments as a treat-to-target approach for [long-term] MS care".[5] Similarly, there is no biofluid marker available for clinical practice to monitor drug response, or to predict the course of progression in individual patients.[6] Accordingly, there is no 'common denominator endpoint' established for an objective evaluation of the relative effectiveness across disease modifying therapies (DMTs), on the background that head-to-head comparisons of modern high-efficacy DMTs are lacking.[2]

Neurofilament light chain (NfL) is a neuro-axonal cytoskeletal protein that is released into cerebrospinal fluid (CSF) and eventually into blood upon neuronal injury.[7] In MS, it has been established as the first serum biomarker to reflect acute disease activity (relapses and lesion formation), to correlate with therapy response and to predict the course of disability-worsening[7–14] on the group level. Other than clinical measures, but equivalent to MRI, serum NfL (sNfL) provides a rater-independent quantification of the intensity of ongoing neuronal

damage based on a standardized assay platform[7] and could therefore serve as a common denominator for the objective comparative assessment of drug effectiveness across all DMTs.[15] However, sNfL is not a stable measure but increases physiologically with age[7] and decreases with body-mass-index (BMI).[16,17] These physiological modulators hamper the validity of fixed cut-off values to define 'pathological' levels for individuals and limit the use of sNfL as a biomarker to group level comparisons where through randomisation or other ways of adjustment these confounding factors may be neutralised. Hence, for individual use and to compare across treatment groups in real-world settings reference values are needed that control for age, BMI and potentially for comorbidities that impact on sNfL levels. We aimed at establishing percentiles/Z-scores for sNfL based on a large reference database (RDB) from a general population to define levels of pathological increase independent of BMI and age. Our objective was to test whether these adjusted sNfL measures, in two large and independent cohorts of MS patients, predict the risk for the future disease activity (relapses, significant increase in EDSS, new/enlarging T2w lesions or T1w contrast-enhancing lesions) both in patient groups and in clinical use cases of individual patients, and whether they allow to quantitate the long-term effectiveness of and across DMT classes.

Methods

Swiss Multiple Sclerosis Cohort (SMSC), Swedish MS registry cohorts and source of data of persons included into the reference database

The SMSC is a prospective multicentre cohort study performed across eight Swiss academic medical centres. All patients with a diagnosis of relapsing or secondary progressive MS were included. For independent validation, we investigated 4341 MS cases participating in three prospective partly overlapping large cohorts in Sweden, the, the Epidemiological Investigation of Multiple Sclerosis (EIMS), Immunomodulation and Multiple Sclerosis Epidemiology (IMSE) and Comparison Between All immuno-Therapies for Multiple Sclerosis (COMBAT-MS) (validation cohort) (**Suppl Table 1**).[18–20] Institutional review boards at the respective SMSC centres and the Stockholm regional ethical committee approved the study, and written informed consent was obtained from all participants.

Therapies were categorized into "HemAb" (high efficacy monoclonal antibody therapies: alemtuzumab, natalizumab, ocrelizumab and rituximab), "oral" (dimethyl fumarate, fingolimod, siponimod and teriflunomide), "platform" (interferon beta, glatiramer acetate) and "untreated" (see Suppl Methods for further details).

The origin and characteristics of the four cohorts of control persons included in the RDB are described in **Suppl Table 2**.

Serum/plasma Neurofilament light chain measurements

sNfL was measured in duplicate with the NF-light® assay (Quanterix, Billerica, USA) according to the protocol provided by the company. Intra- and inter-assay variability were evaluated with three native quality control serum samples (QC) in each of the runs. All samples produced signals above the analytical sensitivity of the assay. Measurements of the few samples with intra-assay coefficients of variation (CV) >20% were repeated. The mean

CVs of duplicate determinations for concentration were 5·2% (6·2pg/ml, QC 1), 3·1% (18·8pg/ml, QC 2), and 3·0% (37·1pg/ml, QC 3). Interassay CVs were 6·9% (QC 1), 5·5% (QC 2), and 5·8% (QC 3). In the validation cohort, NfL was measured in duplicate in ethylenediaminetetraacetic acid (EDTA)-treated plasma samples (pNfL) by NF-light® assay as described[18,19] (see Suppl Methods).

Statistical analysis

Modelling of the sNfL-BMI-age relationship and creation of the reference database

10,133 serum samples of 5,390 persons without evidence of CNS disease were available for the creation of the RDB (**Suppl Table 2**). This cohort was assembled from European and US population-based studies and control groups of genetic MS studies spanning over 6 decades of life. See Suppl Methods for details on reasoning for inclusion of BMI and age, but not diabetes mellitus and for excluding few samples with an estimated glomerular filtration rate (eGFR) <60 mL/min/1·73 m^2 from the RDB. Suppl Methods also describe selection of one sample from each control person (n: 4532), modelling and how generalisability of the resulting RDB was tested and overtraining of the final RDB was ruled out (**Suppl Tables 3 and 4** and **Suppl Figures 2-6**).

The relationship between sNfL, BMI and age in control persons was modelled using a generalized additive model for location, scale and shape (GAMLSS). From this model, percentiles and Z-scores as two interchangeable measures quantifying the deviation of sNfL values from control persons were calculated.[21] Percentiles express the percentage of persons in the general population that are expected to have a sNfL value (adjusted for age and BMI) as high or lower than a given value. Z-scores express the deviation of the adjusted sNfL from values in the control population in terms of number of standard deviations from the mean.

We report reference values in three different formats: as a Figure (**Figure 1**), as a reference table (**Suppl Table 3**) and as an internet-based App (**Suppl Figure 17**).

Multivariable mixed-effects models with sNfL Z-score as dependent variable

Multivariable linear mixed-effects model with a random intercept for the patient were used to investigate associations between sex, clinical and MRI parameters, DMT and longitudinal sNfL Z-scores as dependent variable. The estimates represent additive effects on the sNfL Z-score.

Comparison of absolute values of sNfL concentration (pg/ml) and sNfL Z-score in terms of association with disease activity

We compared the performance of absolute sNfL concentration with that of sNfL Z-score in terms of association with future disease activity ('evidence of disease activity-3' (EDA-3), Suppl Methods) or recent disease activity (relapse ≤4 months). 'High' sNfL was defined as the portion of samples with highest values (separately based on absolute sNfL level and based on sNfL Z-score) using three different cut-offs: top 25% (i.e. 1st quartile), top 10% and top 5% of all samples. Generalised linear (logistic) mixed-effects models with future and recent disease activity as dependent variable were generated with the dichotomized variable (based on absolute sNfL levels or sNfL Z-scores) as the only predictor and odds ratios are presented. For comparison between the SMSC and the validation cohort identical absolute values of NfL for cut-offs were used.

sNfL percentiles/Z-scores as predictors of future disease activity in MS

We analysed the performance of sNfL Z-scores to quantify the risk of future disease activity using this age- and BMI-adjusted measure of disease activity as a) continuous variable and b)

using cut-offs (sNfL Z-scores above versus below 1 or 1·5 or 2) in univariable generalised linear (logistic) mixed-effects models predicting future disease activity (occurrence of relapses, EDSS worsening or EDA-3) in the following year.

In a next step, we combined disease activity measures currently used in clinical practice (EDSS worsening (Suppl Methods) and rate of relapses in the last year, new/enlarging T2w lesions in the last year, current contrast enhancing lesions) with sNfL Z-scores in multivariable generalised linear (logistic) mixed-effects models to quantify the potentially added contribution of sNfL Z-score to predict the risk of future (following year) EDA-3 status. The fit of the two alternative multivariable models (including and excluding sNfL Z-score) was compared using Chi-square test.

Finally, we analysed the performance of sNfL Z-score cut-offs (dichotomising in above/below respective cut-offs) in patients currently (past year and present) fulfilling NEDA-3 criteria (i.e. without clinical or MRI evidence of disease activity, Suppl Methods) to quantify the risk of future (following year) EDA-3 in univariable generalised linear (logistic) mixed-effects models in these clinical and according to conventional MRI stable patients.

Modelling of the evolution of sNfL Z-scores in four treatment categories using mixed-effects models

To model disease activity as expressed by sNfL Z-scores under specific DMT categories, a multivariable model with sNfL Z-score as dependent variable was built using treatment regimen (DMT categories or untreated) and time since its start (or time untreated, respectively) as explanatory variables. Further, the interaction term between time since start and treatment category was included to assess whether the evolution of sNfL Z-scores differs between the DMT groups. The non-linear dynamics in disease activity over time was modelled by using spline terms for time under treatment/time untreated. The optimal number

of degrees of freedom of the splines (5 in the final model) was chosen based on the model's Akaike information criterion. From the final model, marginal effects for DMT groups along time were extracted and plotted together with the 95% confidence bands using the R-package 'sjPlot'.[22] As a sensitivity analysis, a model adjusted for demographic and clinical covariates (sex, age, disease duration, SPMS vs. RMS, presence of relapse in the last 4 months, EDSS) was built (**Suppl Figure 7**).

All analyses were done using the statistical software package R (version 4.0.4) using two-sided tests.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of this report.

Results

1. Reference Database (RDB) for sNfL in a healthy persons population spanning six decades of life: effect of age and BMI on sNfL percentiles and Z-scores

Figure 1 shows that the age-related increase of sNfL percentiles and Z-scores in control persons is not linear; further analysis revealed that the increase is exponential but with an inflection point around 50 years of age with a steeper increase thereafter (**Suppl Figure 8**). Instead, the BMI shows a constant but inverse correlation with sNfL after age-adjustment (**Suppl Figure 2B**) (reference curves for BMI 20, 25 and 30; **Figure 1**).

As 68% control persons contributed several serum samples at different time points, sensitivity analyses confirmed that shape and position of percentile/Z-score reference curves are insensitive to alterations of the underlying RDB data set (using alternative selections of samples per control person; **Suppl Figure 5** and using bootstrapping; **Suppl Figure 6**).

2. sNfL Z-scores as a continuous endpoint reflecting disease activity

All patients participating in the SMSC with a disease course classified as relapsing (n=1,238; 94·3%) or secondary progressive (n=75; 5·7%) MS, according to Lublin et al.,[23] were included (**Table 1**). The age distribution of patients was congruent to that seen for the RDB population (**Suppl Figure 3**).

At entry into the SMSC, 376 (29%) of the patients were untreated, 169 (13%) were on platform, 453 (35%) were on oral and 303 (23%) on HEmAb therapy (**Table 1**). Over the median follow-up time of 5·6 (IQR: 3·2-7·2) years, 121 patients (9·2%) remained untreated, 788 patients (60·0%) were treated with one, 404 patients (30·8%) were treated with more than one compound from these DMT classes. **Suppl Tables 6 and 7** describe the 2,348 treatment epochs based on all serum samples acquired during respective treatments: untreated: 535 epochs (1150 samples), platform: 262 epochs (720 samples), oral: 891 epochs

(3779 samples) and HEmAb: 660 epochs (2120 samples) summing up to a total of 7,769 samples.

In a multivariable mixed-effects model with sNfL Z-scores as dependent variable, clinical and MRI measures of disease worsening (either active disease (relapses, T1w contrast enhancing lesions)), or progression (EDSS score, hyperintense T2w lesion volume) were strongly and independently associated with higher sNfL Z-scores; further, the clinical experience of an effectiveness hierarchy of HEmAb over oral, and of the latter over platform therapies in reference to untreated patients was clearly confirmed (**Figure 2**, **Suppl Table 8**; congruent results in validation cohort: **Suppl Figure 9**).

3. Identification of pathological sNfL levels in individual patients

Figure 3 shows the correlation of absolute sNfL values in MS patients as a function of age and with colour code for sNfL Z-scores. As seen for controls, absolute sNfL values in MS increase with age. Increased sNfL concentrations measured by higher Z-scores were more frequent in younger versus older patients.

Various fixed cut-offs to define pathological sNfL levels have been used earlier by us[11] and others.[24,25] The horizontal line in **Figure 3** depicts a conservative cut-off of 10pg/ml for an arbitrary definition of a 'non-pathologic' sNfL level. With this approach 68·0% (70/103) and 4·3% (7/164) of patients in the age range of 20-30 years and having increased Z-scores of 1·5-2·0 or >2·0, respectively, would be declared as having sNfL levels within the 'normal' range (i.e. ≤10pg/ml). However, these 77 patients with increased sNfL Z-scores (>1·5) showed more recent clinical (relapses or EDSS worsening) disease activity (p=0·023) and fulfilled concurrent EDA-3 status more frequently (p=0·016; **Suppl Figure 10A**) compared with patients with sNfL Z-scores ≤1·5. Moreover, the patients with increased Z-scores (>1·5) showed a higher propensity for clinical disease activity (p=0·041) and numerically for

fulfilling EDA-3 status (p=0·22) in the following year (**Suppl Figure 10B**). Conversely, in the age range of 30-60 years 39·1% of patients with a 'normal' Z-score (0-1·5), would be labelled as having 'increased' (>10pg/ml) sNfL levels (**Suppl Table 5**). The mismatch between these two ways to define normal values becomes more pronounced in patients >60 years with Z-score ranges 0-1·5 and ≤0: 100% and 50·3% (156/310) are above the set cut-off. The clinical consequence of increased sNfL Z-scores or absolute sNfL levels is a higher likelihood for disease activity in the following year (EDA-3) comparing three equivalent threshold levels (p<0·0001 for all six estimates; see estimates and 95% CI in **Suppl Figure 11** and for validation cohort: **Suppl Figure 12**). However, Z-scores led consistently to higher odds ratios (OR) than absolute sNfL values using 3 different cut-offs ('high' defined as top 25, 10 or 5% of the samples): absolute sNfL levels vs sNfL Z-scores: OR top 25%: 2·09 vs 3·09; OR top 10%: 2·83 vs 3·84; OR top 5%: 2·53 vs 4·43, respectively, which corroborates the superior performance of sNfL Z-scores over fixed cut-off levels of absolute sNfL values, irrespective where cut-off values were set. Accordingly, the association between a recent relapse (<4 months) and sNfL Z-scores was considerably stronger vs absolute sNfL levels (**Suppl Figure 13**; validation cohort: **Suppl Figure 14**).

4. sNfL percentiles/Z-scores as measures and predictors of future disease activity in MS

a) sNfL Z-score alone

Patients with higher sNfL Z-scores showed a higher probability of relapses (OR 1·41; 95% CI 1·30-1·54; p<0·0001; i.e. 41% higher risk per 1 Z-score unit); EDSS worsening (OR 1·11; 95% CI 1·03-1·21; p=0·0093) and EDA-3 (OR 1·43; 95% CI 1·31-1·57; p<0·0001; **Figure 4A**; validation cohort: **Suppl Figure 15A**) in the following year, based on a model with Z-score as a continuous predictor. As compared to the continuous analysis the use of sNfL Z-score cut-offs led to a substantially higher probability of EDA-3 in the following year, in

function of incremental increase of cut-off levels (**Figure 4B**; validation cohort: **Suppl Figure 15B**).

b) sNfL Z-score in combination with other state of the art measures of disease activity

When sNfL Z-scores are combined with disease activity measures currently used in clinical practice (EDSS worsening and relapse rate in the last year, new/enlarging T2w lesions, contrast enhancing lesions) in a multivariable model, the risk of EDA-3 in the following year was increased independently by 23% per 1 step higher sNfL Z-score (95% CI 1·06-1·44; p=0·0072; **Figure 4C**; validation cohort: **Suppl Figure 15C**). Noteworthy, the model quality was improved when sNfL Z-scores were included together with all classical measures (Chi square test: p=0·0023) as compared to the same model without sNfL Z-scores.

c) sNfL Z-scores in NEDA-3 patients

The clinical consequence of increased sNfL Z-scores in NEDA-3 patients is a higher likelihood for EDA-3 status in the following year: sNfL levels were higher than the 89·4th percentile (Z-score >1·25) in 57 of 608 serum samples (9·3%) from patients being classified as NEDA-3 since the last year: these cases displayed a 2·28-fold (95% CI 1·11-4·68; p=0·025) higher risk of experiencing any sign of clinical or MRI disease activity over the following year. This risk increased to 3·85-fold (95% CI 1·27-11·63; p=0·017) in patients with sNfL concentrations exceeding the 96·0th percentile (Z-score >1·75) (**Figure 4D**; validation cohort: **Suppl Figure 15D).**

5. sNfL Z-scores as a continuous endpoint reflecting disease activity and long-term treatment effects of DMT categories

Finally, we modelled the evolution of sNfL Z-scores over time in the four treatment categories in a mixed-effects model. **Figure 5** shows that sNfL concentrations decreased rapidly in the first year after therapy initiation, whereas they decreased only marginally in untreated patients. The reduction of the sNfL Z-score was more rapid under HEmAb, as compared to oral and platform therapies, as reflected by the steeper slope ($p<0{\cdot}0001$ for the interaction term between treatment category and treatment duration). Over the following four years only HEmAb, but not oral therapy was associated with sNfL levels overlapping with those of the control population (sNfL Z-score 0), while with platform therapies they remained increased. Platform therapies were associated with the weakest sNfL reduction in the first year of treatment, and were followed by a new increase thereafter, coming close to levels measured in untreated patients. As sensitivity analysis, a model adjusted for demographic and clinical covariates confirmed the effectiveness hierarchy established in the unadjusted analysis (as well as in the multivariable analysis in **Figure 2**) with estimated marginal effects (remaining disease activity explained by sNfL Z-score) being numerically lower (**Suppl Figure 7**).

6. Clinical use cases and development of an internet-based tool to determine Z-scores

Suppl Figure 16 shows 7 clinical use cases from the SMSC for the application of sNfL percentiles/Z-scores as biomarker, covering therapy monitoring, and risk assessment for future acute and chronic disease activity.

To facilitate the use of sNfL Z-scores in clinical practice, we created an application based on the sNfL values from the RDB, to determine Z-scores and respective percentile values by entering patients' measured sNfL concentrations, height, weight (or BMI) and age.

The adjusted sNfL measures (percentiles and Z-scores) can be retrieved in both numerical format and as a graphical illustration (**Suppl Figure 17**). The application is accessible under: under construction.

Discussion

Present results demonstrate that NfL can be used as a biomarker for monitoring of treatment efficacy and prognostication of disease course in individual patients, based on reference data from a general population and two large, independent real-world MS cohorts. The statistical transformation from absolute values into percentiles/Z-scores allows to reliably correct for confounding factors, such as age and BMI, and to discern pathological from physiological levels of sNfL. sNfL can be used as additional measure of disease activity (EDA-3) besides clinical assessments and MRI. It is specifically useful for stable patients, i.e. in NEDA-3 status, to identify ongoing disease activity that is below detection threshold of standard clinical and MRI markers. Based on these reference values, sNfL levels can be used as well for the quantitative comparison of long-term effectiveness across DMT groups (keeping in mind limitations based on design preventing proof of causation in real-world settings).

In 2018, Giovannoni coined the term of "NfL [as] the neurologist's C-reactive protein" (CRP) to measure neuroprotective effects of DMTs in the context of clinical trials.[26] Since then, numerous clinical studies have shown the utility of sNfL to quantitate disease activity in MS and other neurological disorders,[27] and recent phase 3 studies in MS used sNfL as an exploratory endpoint for treatment efficacy.[11,28–30] Despite these studies demonstrate that sNfL accurately reflects even subclinical disease activity[28,30] sNfL has not been generally accepted as clinical routine biomarker for individual MS patients, nor as primary or secondary trial endpoint. In contrast to CRP, sNfL lacked two essential premises for such a breakthrough: reference values from a general population of persons without clinically manifest diseases[7,27] and a way to interpret values without the interfering factor of age and BMI.

With the advent of high-efficacy MS therapies, relapses and high rates of lesion formation have been suppressed almost completely. In focus are now the questions how to control the

subclinical, diffuse brain damage that manifests clinically as continuously worsening disease ('progression'), and how to measure it. As sNfL values remain modestly increased in this disease state as compared to the more pronounced NfL level increases associated with relapses[31] the task to discern the disease signal from the age-related increase becomes more challenging. The earlier assumption of a constant increase of 2·2%/year of sNfL in controls[7] was based on cohorts[8,9,32] that were too small and insufficiently covered the age range specifically relevant for progressive MS. OurRDB data reveal that the evolution of sNfL with age follows a non-log-linear function and establish BMI as an important additional independent (not age-dependent) modulator of NfL levels in reference populations. In consequence, fixed cut-offs may lead to a misclassification, even if the cut-off is set at a lower level in the present analysis than in earlier ones.[25,33] Current results show that a significant proportion of young MS patients have ongoing disease activity that would remain unrecognized using such fixed-cut-off levels, and hence the purpose of measuring sNfL to guide therapeutic decisions might be missed. Additionally, the inclusion of BMI to define reference percentiles/Z-scores further increases the precision in determining pathological cut-off values. In general, Z-scores are more accurate versus absolute values of sNfL to reflect past and to predict future clinical disease activity. Conversely, a fixed cut-off may lead to a significant false-positive rate in patients above 40 years which is problematic for the interpretation of sNfL levels in patients with progressive MS, or primarily neurodegenerative diseases.[34]

Z-scores are a standard measure in other fields of medicine, e.g., echocardiographic measurement of aortic dilation or for determination of bone mineral density to separate pathology-indicating signals of biomarkers from physiological longitudinal changes.[21,35] Percentiles used e.g. in paediatric growth curves are like Z-scores a derivative of standard deviation calculations and are a very similar way to describe deviation from normality in

medicine.[21] However, they are less sensitive to longitudinal change, particularly for extreme values, due to their finite measuring range. Instead, Z-scores can quantify deviations from normal values beyond a percentile range.

On the group level, Z-scores allow to quantify the contribution of clinical and of MRI features to disease activity, as well as of effectiveness of therapy categories of DMT. Physicians have nowadays the choice between more than 10 registered DMT for MS therapy, but a quantitative assessment of their efficacy across the various clinical trials, specifically related to their effect on the long-term course of disease, is not possible due to methodological reasons. With the RDB and Z-scores we can now model the effectiveness of drugs and of residual disease activity over years of treatment. HEmAb and to a lesser extent oral therapy, coincides with a normalisation of sNfL levels over time. In contrast, the wearing-off of the treatment effect of platform therapies in presented models, as seen in earlier long-term extensions of two clinical studies with interferon beta, is mirrored by a continuous increase of sNfL.[36,37]

There are several limitations to this study. The RDB is based on a cohort of persons without clinical manifestation of somatic diseases. Many subclinical disease conditions may, however, go along with an increase of sNfL levels due to neuronal damage to the nervous system. For example, underlying primary neurodegenerative diseases, like Alzheimer's disease, can lead to NfL increase years before they clinically manifest.[34] On purpose we established our RDB not on a cohort of persons where subclinical laboratory aberrations have been excluded, i.e. whose serum samples were selected for absence of neurodegenerative or other diseases developing later in life. Such diseases may occur as well with similar incidence and prevalence in MS patients. Hence, in view of the use of the RDB percentiles/Z-scores in

clinical real-world practice for MS patients, we did not pursue the concept to correct for such comorbidities occurring at later stage in life.

While we have acquired limited data that mild renal insufficiency and diabetes have little impact on sNfL levels, we need to define how more severe stages of these diseases and possibly other confounding factors limit its interpretability in MS patients. Second, our results are largely based on relapsing MS patients of Caucasian origin; the generalisability in primary progressive MS, and in patients with different ethnical background needs to be validated in referring cohorts. Third, it is not known whether data acquired with the current standard assay system Simoa® is fully compatible with that of other analytical platforms for NfL, given they provide highly correlated, but different absolute values. Standardization efforts are now ongoing within the International Federation of Clinical Chemistry (IFCC), aiming for developing Certified Reference Materials (CRMs) for harmonization of readouts across platforms. In essence, the use of our internet-based percentile/Z-score tool requires that data is acquired with the standard kit and on the same hardware platform.

In conclusion, sNfL percentiles/Z-scores may become a clinical tool to identify subclinical disease activity in individual MS patients and to monitor drug response. It is now available for physicians by use of an internet-based application. This tool may be used as well in future trials where sNfL is an endpoint measure.

References

1 Hauser SL, Cree BAC. Treatment of Multiple Sclerosis: A Review. *Am J Med* 2020; **133**: 1380-1390.e2.

2 Tur C, Kalincik T, Oh J, *et al.* Head-to-head drug comparisons in multiple sclerosis: Urgent action needed. *Neurology* 2019; **93**: 793–809.

3 Gyllensten H, Kavaliunas A, Alexanderson K, Hillert J, Tinghög P, Friberg E. Costs and quality of life by disability among people with multiple sclerosis: a register-based study in Sweden. *Mult Scler J - Exp Transl Clin* 2018; **4**: 205521731878335.

4 Rotstein DL, Healy BC, Malik MT, Chitnis T, Weiner HL. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. *JAMA Neurol* 2015; **72**: 152–8.

5 University of California, San Francisco MS-EPIC Team: Cree BAC, Gourraud PA, Oksenberg JR, *et al.* Long-term evolution of multiple sclerosis disability in the treatment Era. *Ann Neurol* 2016; **80**: 499–510.

6 Ziemssen T, Akgün K, Brück W. Molecular biomarkers in multiple sclerosis. *J Neuroinflammation* 2019; **16**: 1–11.

7 Khalil M, Teunissen CE, Otto M, *et al.* Neurofilaments as biomarkers in neurological disorders. *Nat Rev Neurol* 2018; **14**: 577–89.

8 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.

9 Barro C, Benkert P, Disanto G, *et al.* Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis. *Brain* 2018; **141**: 2382–91.

10 Chitnis T, Gonzalez C, Healy BC, *et al.* Neurofilament light chain serum levels correlate with 10-year MRI outcomes in multiple sclerosis. *Ann Clin Transl Neurol*

2018; **5**: 1478–91.

11 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**: E1007–15.

12 Leppert D, Kuhle J. Blood neurofilament light chain at the doorstep of clinical application. Neurol. Neuroimmunol. NeuroInflammation. 2019; **6**: 4–5.

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

14 Bittner S, Steffen F, Uphaus T, *et al.* Clinical implications of serum neurofilament in newly diagnosed MS patients: A longitudinal multicentre cohort study. *EBioMedicine* 2020; **56**: 1–13.

15 Sormani MP, Haering DA, Kropshofer H, *et al.* Blood neurofilament light as a potential endpoint in Phase 2 studies in MS. *Ann Clin Transl Neurol* 2019; **6**: 1081–9.

16 Polymeris AA, Coslovksy M, Aeschbacher S, *et al.* Serum neurofilament light in atrial fibrillation: clinical, neuroimaging and cognitive correlates. *Brain Commun* 2020; **2**: fcaa166.

17 Manouchehrinia A, Piehl F, Hillert J, *et al.* Confounding effect of blood volume and body mass index on blood neurofilament light chain levels. *Ann Clin Transl Neurol* 2020; **7**: 139–43.

18 Delcoigne B, Manouchehrinia A, Barro C, *et al.* Blood neurofilament light levels segregate treatment effects in multiple sclerosis. *Neurology* 2020; **94**: e1201–12.

19 Manouchehrinia A, Stridh P, Khademi M, *et al.* Plasma neurofilament light levels are associated with risk of disability in multiple sclerosis. *Neurology* 2020; **94**: e2457–67.

20 Alping P, Piehl F, Langer-Gould A, Frisell T. Validation of the Swedish Multiple Sclerosis Register: Further Improving a Resource for Pharmacoepidemiologic

Evaluations. Epidemiology 2019; 30: 230–3.

21 Curtis A, Smith T, Ziganshin B, Elefteriades J. The Mystery of the Z-Score. *Aorta* 2016; **04**: 124–30.

22 Lüdecke D. sjPlot: Data Visualization for Statistics in Social Science. 2021. https://cran.r-project.org/package=sjPlot.

23 Lublin FD, Reingold SC, Cohen JA, *et al.* Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* 2014; **83**: 278–86.

24 Calabresi PA, Arnold DL, Sangurdekar D, *et al.* Temporal profile of serum neurofilament light in multiple sclerosis: Implications for patient monitoring. *Mult Scler J* 2020; : 11–3.

25 Håkansson I, Tisell A, Cassel P, *et al.* Neurofilament levels, disease activity and brain volume during follow-up in multiple sclerosis. *J Neuroinflammation* 2018; **15**: 1–10.

26 Giovannoni G. Peripheral blood neurofilament light chain levels: the neurologist's C-reactive protein? *Brain* 2018; **141**: 2235–7.

27 Lambertsen KL, Soares CB, Gaist D, Nielsen HH. Neurofilaments: The C-reactive protein of neurology. *Brain Sci* 2020; **10**: 1–29.

28 Kapoor R, Sellebjerg F, Hartung H-P, *et al.* Natalizumab reduced serum levels of neurofilament light chain in secondary progressive multiple sclerosis patients from the phase 3 ASCEND study. *ECTRIMS Online Libr* 2018. https://onlinelibrary.ectrims-congress.eu/ectrims/2018/ectrims-2018/228118/raj.kapoor.natalizumab.reduced.serum.levels.of.neurofilament.light.chain.in.html.

29 Hauser SL, Bar-Or A, Cohen JA, *et al.* Ofatumumab versus Teriflunomide in Multiple Sclerosis. *N Engl J Med* 2020; **383**: 546–57.

30 Kuhle J, Kropshofer H, Haering DA, *et al.* Neurofilament light levels in the blood of

patients with secondary progressive MS are higher than in primary progressive MS and may predict brain atrophy in both MS subtypes. In: ECTRIMS Online Library. 2018.

31 Akgün K, Kretschmann N, Haase R, *et al.* Profiling individual clinical responses by high-frequency serum neurofilament assessment in MS. *Neurol Neuroimmunol NeuroInflammation* 2019; **6**. DOI:10.1212/NXI.0000000000000555.

32 Khalil M, Pirpamer L, Hofer E, *et al.* Serum neurofilament light levels in normal aging and their association with morphologic brain changes. *Nat Commun* 2020; **11**: 1–9.

33 Calabrese P, Kobelt G, Berg J, Capsa D, Eriksson J. New insights into the burden and costs of multiple sclerosis in Europe: Results for Switzerland. *Mult Scler J* 2017; **23**: 192–203.

34 Preische O, Schultz SA, Apel A, *et al.* Serum neurofilament dynamics predicts neurodegeneration and clinical progression in presymptomatic Alzheimer's disease. *Nat Med* 2019; **25**: 277–83.

35 Cummings SR, Bates D, Black DM. Clinical use of bone densitometry: scientific review. *JAMA* 2002; **288**: 1889–97.

36 Kappos L, Freedman MS, Polman CH, *et al.* Long-term effect of early treatment with interferon beta-1b after a first clinical event suggestive of multiple sclerosis: 5-year active treatment extension of the phase 3 BENEFIT trial. *Lancet Neurol* 2009; **8**: 987–97.

37 Pittock SJ. Uncertain BENEFIT of early interferon beta-1b treatment. *Lancet Neurol* 2009; **8**: 970–1.

Acknowledgements

The authors express their deep thankfulness to patients and relatives for their participation and support, study nurses in participating centres for their motivated collaboration and recruitment efforts and the administrative personnel of the Swiss Multiple Sclerosis Cohort and the Swedish MS registry cohorts. We would like to thank Mathieu Canales, Lilian Demuth and Irmtraut Scheerer for expert technical support.

RDB in the Swiss Multiple Sclerosis Cohort Study Group

Stefanie Aeschbacher PhD, Muhamed Barakovic PhD, Andreas Buser MD, Marcus D'Souza MD, Riccardo Galbusera MD, Johannes Lorscheider MD, Annette Orleth PhD, Ernst-Wilhelm Radue MD, Reza Rahmanzadeh MD, Tim Sinnecker MD, Suvitha Subramaniam MSc, Jens Wuerfel MD (University Hospital Basel, University of Basel, Basel, Switzerland); Andrew Chan MD (Bern University Hospital, University of Bern, Bern, Switzerland); Giulio Disanto MD (Neurocenter of Southern Switzerland, Lugano, Switzerland); Renaud Du Pasquier MD (Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland); Oliver Findling MD (Cantonal Hospital Aarau, Aarau, Switzerland); Kevin Hrusovsky MBA (Quanterix Billerica, MA, United States of America); Michael Khalil PhD (Medical University of Graz, Graz, Austria); Amandine Mathias PhD (Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland); Jochen Vehoff MD (Cantonal Hospital St. Gallen, St. Gallen, Switzerland); Sven Wellmann MD (University Children's Hospital Regensburg, University of Regensburg, Regensburg, Germany); Chiara Zecca MD (Università della Svizzera Italiana (USI), Lugano, Switzerland).

Author contributions

Conceptualisation: PB, SS, CG, LK, HW, KBe, CGr, DL, JK.

Data curation: PB, SM, SS, AMan, ÖY, AM, JO, SA, EW, DL, JK.

Formal analysis: PB, SM, SS, AMan, AM, SSu, JK.

Funding acquisition: DL, JK.

Investigation: PB, SM, SS, AMan, ÖY, AM, JO, LA, SA, MB, AC, DC, TD, GD, MD, RDP, OF, RG, PL, JL, AMat, CM, SMü, YN, JOk, AO, CP, ER, RR, AS, TS, JV, CZ, IK, CG, LK, CG, FP, JK.

Methodology: PB, SM, SS, MPS, DL, JK.

Project administration: PB, SM, AM, SSu, JK.

Resources: LA, SA, AB, AC, DC, GD, MD, RDP, OF, KH, PL, CM, SMü, JOk, CP, AS, JV, CZ, IK, CG, HW, KBe, FP, JK.

Software: PB, SS.

Supervision: DL, JK.

Validation: PB, SM, AMan, IK, FP, DL, JK.

Verification of data: PB, AMan, FP, JK.

Visualisation: PB, SM, SS, DL, JK.

Writing – original draft: PB, SM, DL, JK.

Writing – review & editing: all authors.

Access to raw data and final responsibility for the decision to submit for publication: PB, JK

Declaration of interests

P.B., S.M., S.S. and A.Man. report no conflicts of interest.

Ö.Y.'s received grants from ECTRIMS/MAGNIMS, University of Basel, Pro Patient Stiftung, University Hospital Basel, Free Academy Basel, Swiss Multiple Sclerosis Society and advisory board/lecture and consultancy fees from Roche, Sanofi Genzyme, Allmirall, Biogen and Novartis.

A.M. reports no conflicts of interest.

J.O. served on advisory boards for Roche and Merck.

L.A. served on scientific advisory boards for Celgene, Novartis Pharmaceuticals, Merck, Biogen, Sanofi Genzyme, Roche and Bayer; received funding for travel and/or speaker honoraria from Celgene, Biogen, Sanofi Genzyme, Novartis, Merck Serono, Roche, Teva and the Swiss MS Society; and research support from Biogen, Sanofi Genzyme, and Novartis.

S.A., M.B. and A.B. report no conflicts of interest.

A.C. received compensation for activities with Actelion, Almirall, Bayer, Biogen, Celgene, Sanofi-Genzyme, Merck, Novartis, Roche, Teva, all for hospital research funds. He receives research support from Biogen, Sanofi-Genzyme and UCB. He serves as associate editor for the European Journal of Neurology.

D.C. received speaker fees from BMS/Pfizer and consultation fees from Roche Diagnostics.

T.D. received speaker fees, research support, and/or served on Advisory Boards, data safety monitoring boards, or Steering Committees of Actelion, Alexion, Celgene, Polyneuron, Novartis, Merck, Biogen, GeNeuro, MedDay, Roche, and Genzyme. TD received research support from the Swiss National Science Foundation and the Swiss MS Society. TD is secretary and member of the executive board of ECTRIMS.

G.D. and M.D. report no conflicts of interest.

R.D.P. has received honoraria for advisory boards from Biogen, Celgene, Merck, Novartis, Roche and Sanofi-Genzyme.

O.F. and R.G. report no conflicts of interest.

K.H. is an employee and stockholder at Quanterix Corp.

M.K. has received funding for attending meetings or travel from Merck and Biogen, honoraria for lectures/presentations from Novartis and Biogen and speaker serves on

scientific advisory boards for Biogen, Merck, Roche, Novartis, Bristol-Myers Squibb and Gilead. He received research grants from Biogen and Novartis.

P.L. received honoraria for speaking from Biogen Idec, Genzyme, Merck Serono, Novartis, Sanofi Aventis, Teva, consulting fees from Biogen Idec, GeNeuro, Genzyme, Merck Serono, Novartis, Sanofi-Aventis, Teva, research grants from Biogen Idec, Merck Serono, Novartis.

J.L. has received research support from Innosuisse Innovation Agency, Biogen and Novartis and served on advisory boards for Roche and Teva.

A.Mat. reports no conflicts of interest.

C.M. has received research support from the Swiss National Science Foundation, the Swiss Heart Foundation, the KTI, the University of Basel; Abbott, Astra Zeneca, Beckman Coulter, Brahms, Idorsia, Novartis, Quidel, ortho clinical Diagnostics, Roche, Siemens, Singulex, Sphingotec, and the University Hospital Basel, as well as speaker honoraria/consulting honoraria from Amgen, Astra Zeneca, BMS, Bayer, Daiichi Sankyo, Osler, Novartis, Roche, Sanofi and Singulex, all outside the submitted work.

S.Mü. reports no conflicts of interest.

Y.N.'s institution (University Hospital Basel/Research Center for Clinical Neuroimmunology and Neuroscience Basel, Switzerland) has received financial support for lectures from Teva and Celgene, grant support from Innosuisse (Swiss Innovation Agency) and grant support from Novartis and Roche.

J.Ok. and A.O. report no conflicts of interest.

C.P. received consulting fees and/or travel compensation, used exclusively for research support, for activities with Biogen, Merck, Novartis, Roche and Sanofi Genzyme.

E.R. and R.R. report no conflicts of interest.

A.S. received speaker honoraria and/or travel compensation for activities with Almirall Hermal GmbH, Biogen, Merck, Novartis, Roche, and Sanofi Genzyme, and research support by the Swiss MS Society.

T.S. has received travel support from Actelion, Alkermes, and Roche. He is a part time employee of the MIAC AG in Basel.

S.Su. reports no conflicts of interest.

J.V. has received speaker honoraria from Allmiral Hermal GmbH and Roche.

S.W. is Chief Medical Officer and co-founder of Neopredix.

J.W. is an employee of MIAC AG, Basel, Switzerland; he received speaker or consulting honoraria or research grants from Actelion, Alexion, Biogen, Idorsia, ImmuneBio, Novartis, Roche, Sanofi, and is or was supported by the Eu (Horizon 2020), the SNCF, German Ministry of Science (BMBF) and the German Ministry of Economy (BMWi).

C.Z. received honoraria for speaking/consulting fees or grants from Abbvie, Almirall, Biogen Idec, Celgene, Genzyme, Lilly, Merck Serono, Novartis, Roche, Teva Pharma

E.W. and I.K. report no conflicts of interest.

K.B. has served as a consultant, at advisory boards, or at data monitoring committees for Abcam, Axon, Biogen, JOMDD/Shimadzu. Julius Clinical, Lilly, MagQu, Novartis, Roche Diagnostics, and Siemens Healthineers, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program.

H.Z. has served at scientific advisory boards for Alector, Denali, Roche Diagnostics, Wave, Samumed, Siemens Healthineers, Pinteon Therapeutics and CogRx, has given lectures in symposia sponsored by Cellectricon, Fujirebio, Alzecure and Biogen, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program (outside submitted work).

C.G. received honoraria for speaking/consulting fees or grants from Abbvie, Almirall, Biogen

Idec, Celgene, Genzyme, Merck Serono, Novartis, Roche, Teva Pharma.

L.K.'s employer (University Hospital Basel) has received and dedicated to research support steering committee, advisory board, and consultancy fees (Abbvie, Actelion, Almirall, Auriga Vison AG, Bayer HealthCare, Biogen, Eisai, EMD Derono Inc, Genzyme, Genentech Inc, F. Hoffmann-La Roche, Japan Tobacco, Janssen Pharmaceuticals Inc, Merck, Minoryx Therapeutics SL, Novartis, Sanofi, Santhera, Senda Biosciences, Shionogi BV, TG Therapeutics); speaker fees (Bayer HealthCare, Biogen, Celgene, Genzyme, Janssen Pharmaceuticals Inc, Merck, Novartis, Roche, and Sanofi); support of educational activities (Allergan, Bayer HealthCare, Biogen, CSL Behring, Genzyme, Merck, Novartis, Roche, Pfizer, Sanofi, Shire, and Teva); license fees for Neurostatus products; and grants (Bayer HealthCare, Biogen, European Union, Innosuisse, Merck, Novartis, Roche Research Foundation, Swiss MS Society, and Swiss National Research Foundation).

H.W. received honoraria and consultation fees from Bayer Healthcare, Biogen, Fresenius Medical Care, GlaxoSmithKline, GW Pharmaceuticals, Merck Serono, Novartis, Sanofi Genzyme, and Teva Pharma.

K.Be. received a grant from the BMBF (within the German Competence Net Multiple Sclerosis) plus additional funds from Biogen, all to the University of Münster, for an investigator-initiated AE registry for patients with multiple sclerosis.

M.P.S. has received consulting fees from Biogen, Merck, Teva, Genzyme, Roche, Novartis, GeNeuro and MedDay.

C.Gr.: The University Hospital Basel (USB), as the employer of Cristina Granziera has received the following fees which were used exclusively for research support: (i) advisory board and consultancy fees from Actelion, Novartis, Genzyme and F. Hoffmann-La Roche; (ii) speaker fees from Biogen and Genzyme-Sanofi; (iii) research support by F. Hoffmann-La

Roche Ltd. Before my employment at USB, I have also received speaker honoraria and travel funding by Novartis.

F.P. has received research grants from Merck KGaA and UCB, and fees for serving on DMC in clinical trials with Chugai, Lundbeck and Roche.

D.L. is Chief Medical Officer of GeNeuro.

J.K. received speaker fees, research support, travel support, and/or served on advisory boards by the Progressive MS Alliance, Swiss MS Society, Swiss National Research Foundation (320030_189140 / 1), University of Basel, Biogen, Celgene, Merck, Novartis, Octave Bioscience, Roche, Sanofi.

Funding

This investigation was supported by Swiss National Science Foundation (grant 320030_189140 / 1), award from Progressive MS Alliance, award reference number PA-2007-36872, and grant funding from Biogen, Celgene, Novartis, Roche. The Swiss MS Cohort study received funding from the Swiss MS Society and grant funding from Biogen, Celgene, Merck, Novartis, Roche, and Sanofi. The Swedish MS registry is funded by Swedish Municipalities and County Councils and Swedish National Board of Health and Welfare. The IMSE cohorts received grant support from Biogen (natalizumab, peginterferon beta-1a and dimethyl fumarate), Genzyme (teriflunomide and alemtuzumab), and Novartis (fingolimod). The Combat-MS study is funded through a Patient-Centered Outcomes Research Institute (PCORI) Award (MS-1511–33196). The statements presented in this publication are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its Board of Governors or Methodology Committee. A. Man. was supported by grant from Magreta af Ugglas foundation. IK and the pNfL measurements in the validation cohort were supported by a Horizon 2020 Eu Grant (MultipleMS project number

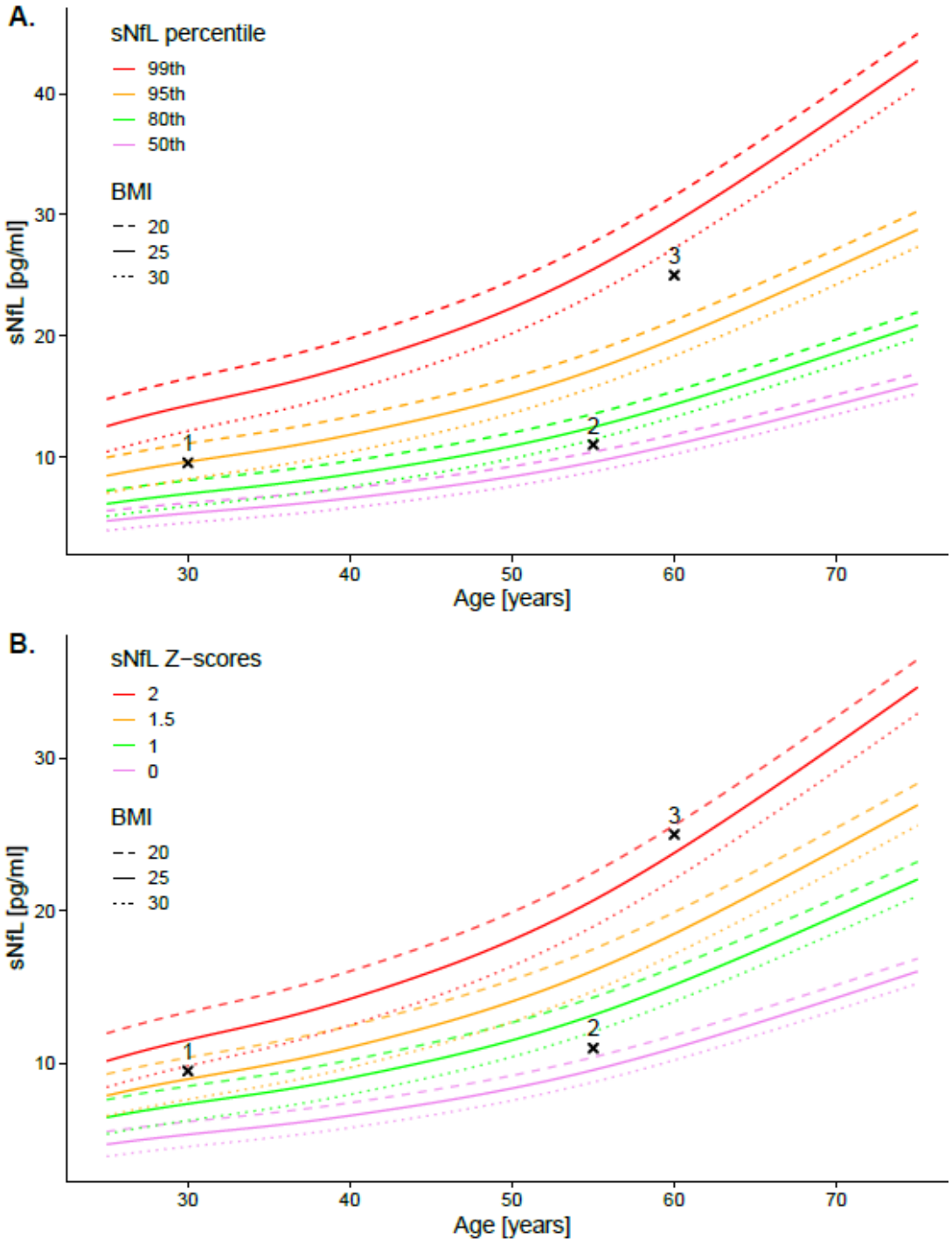
733161). The UCSF DNA biorepository is supported by grant Si-2001-35701 from the US National Multiple Sclerosis Society.

Data sharing

Written requests for access to the data reported in this paper will be considered by the corresponding author and a decision made about the appropriateness of the use of the data. If the use is appropriate, a data sharing agreement will be put in place before a fully de-identified version of the dataset used for the analysis with individual participant data is made available. The internet-based application for determination of sNfL Z-scores is available under: under construction.

Figures

Figure 1. sNfL percentiles (A.) and Z-scores (B.) reference curves based on controls.



Legend:

A generalised additive model for location, scale and shape (GAMLSS) was used to model the association of sNfL concentration (pg/ml) in controls with BMI and age. Lines indicate

percentiles (A.) or Z-scores (B.), standardised and age-independent measures of deviation from normal. sNfL values show a non-linear increase in function of age. Lower levels of sNfL are seen with higher body mass index (BMI; 30kg/m^2 versus 25kg/m^2 versus 20kg/m^2). The y-axis was capped at 40pg/ml.

Reading examples:

1.) 30-years; BMI: 25; sNfL: 9·5pg/ml: 95th percentile/ Z-score > 1·5 (exact values as calculated by the sNfL App: 95·0/ 1·64, resp.).

Interpretation: elevated.

2.) 55-years; BMI: 25; sNfL: 11·0pg/ml: below the 80th percentile/ Z-score < 1·0 (exact values: 68·0/ 0·47, resp.).

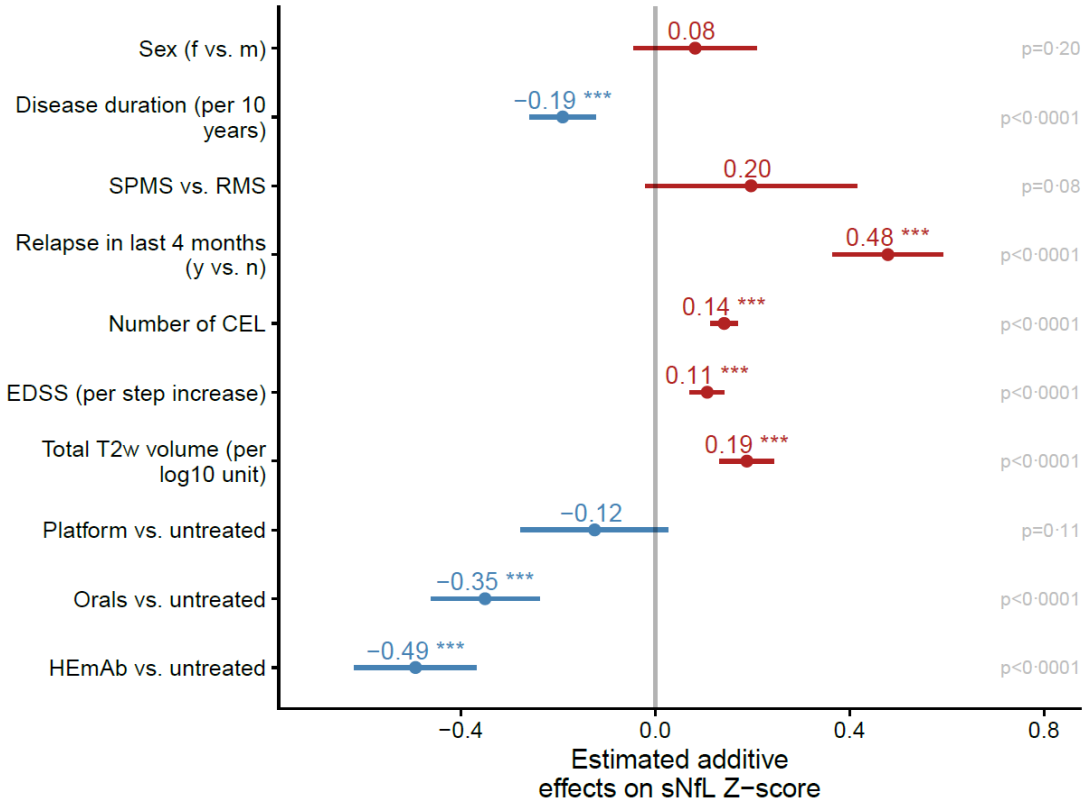
Interpretation: similar to levels seen in controls.

3.) 60-years; BMI: 30; sNfL: 25pg/ml: close to the 99th percentile/ Z-score > 2 (exact values: 98·6/ 2·2, resp.).

Interpretation: elevated.

Abbreviations: BMI: body-mass-index; sNfL: serum neurofilament light chain.

Figure 2. Factors influencing sNfL Z-scores in MS.

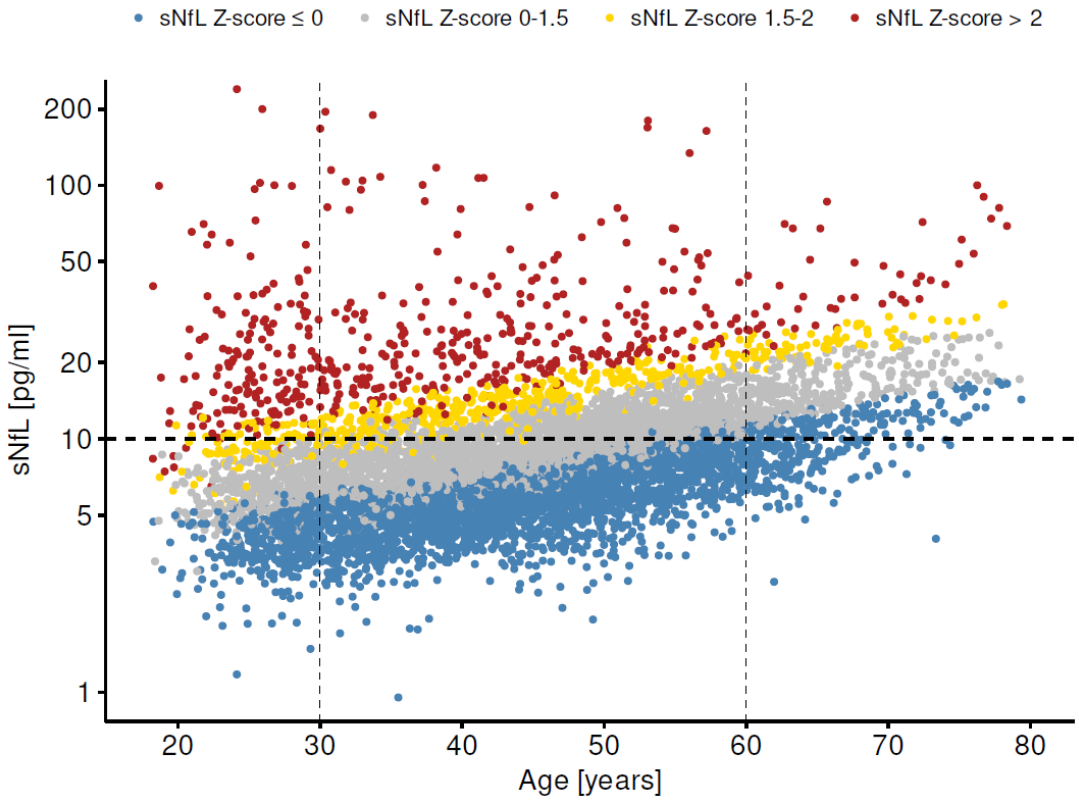


Legend:

Model estimates including 95% confidence intervals (see also **Suppl Table 8** for numeric values; e.g., sNfL Z-scores are on average 0·48 units higher within 4 months after a relapse). Estimates for HEmAb vs oral therapy was -0·14, 95% CI: -0·23--0·05, p=0·0018; for oral vs. platform therapy the estimate was -0·23, 95% CI: -0·36--0·10, p<0·0001. ***: p<0·001; **: p<0·01; *: p<0·05.

Abbreviations: CEL: contrast enhancing T-weighted lesions; EDSS: Expanded Disability Status Scale score; f: female; HEmAb: high efficacy monoclonal antibody therapies; sNfL: serum neurofilament light chain; m: male; n: no; PPMS: primary progressive multiple sclerosis; RMS: relapsing multiple sclerosis; SPMS: secondary progressive multiple sclerosis; T2w: T2-weighted; y: yes.

Figure 3. sNfL Z-scores in MS patients participating in the SMSC and comparison of the proportion of samples from MS patients with increased sNfL levels using either sNfL Z-scores (colour gradient) or a fixed sNfL cut-off (horizontal line at 10pg/ml).



Legend:

sNfL concentration in MS patients increases with age, which makes it difficult to distinguish between physiological age-related increases and disease activity. This is not the case when using sNfL Z-scores (shown as colour gradient).

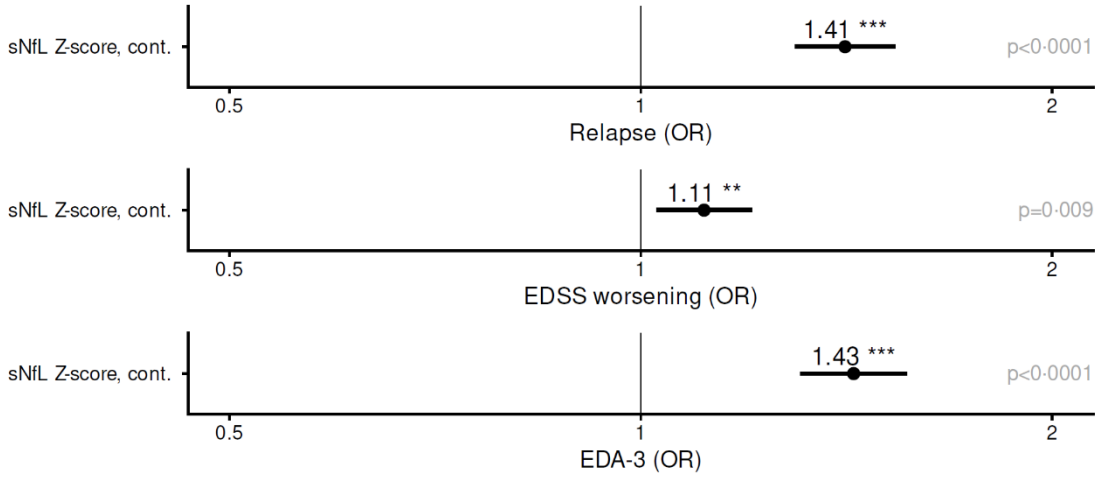
A fixed threshold (here 10pg/ml) for the detection of disease activity is compared to sNfL Z-scores considered increased (here Z-scores 1·5-2 (yellow) and >2 (red)), and Z-scores (≤1·5 (grey and blue)) not increased. Using a fixed cut-off in MS patients 20-30 years old may miss patients with increased sNfL Z-scores (false negatives: yellow and red dots below horizontal 10pg/ml cut-off). Conversely, in older patients a large portion of patients with normal age-corrected sNfL (*i.e.*, sNfL Z-scores 0-1·5 (grey), ≤0 (blue)) show values above the fixed threshold of pathology (false positives). Numerical values are provided in **Suppl Table 5**.

Age- and BMI-adjusted sNfL Z-scores are shown. Different Z-scores can occur with similar sNfL level and identical age as a result of their additional adjustment for BMI, beyond age.

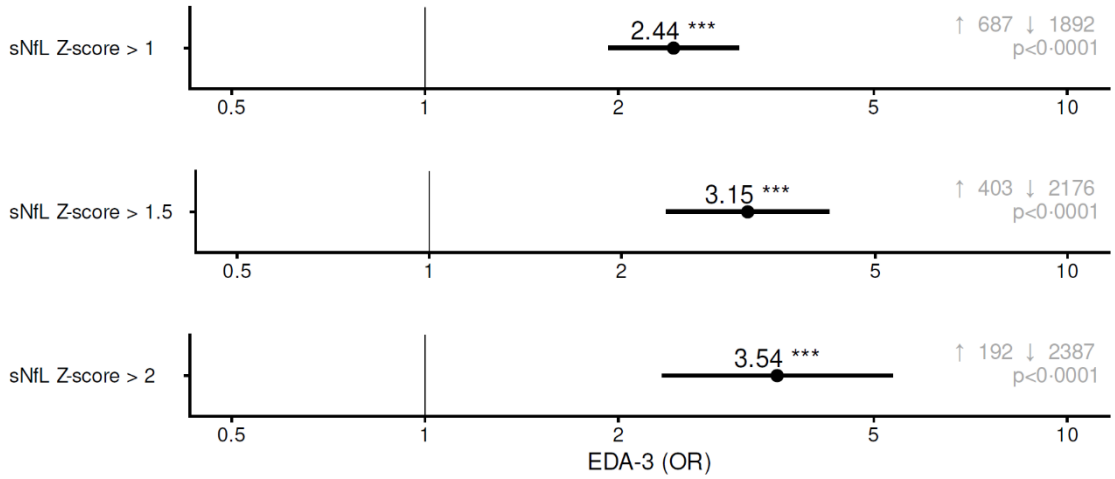
Abbreviations: sNfL: serum neurofilament light chain.

Figure 4. sNfL Z-scores predicting disease activity in the following year: **A.** Probability of occurrence of relapses or EDSS worsening or EDA-3 in the following year based on (continuous) sNfL Z-score; **B.** using sNfL Z-score cut-offs; **C.** in combination with other currently used measured of disease activity in clinical practice in a multivariable model; **D.** and in NEDA-3 patients.

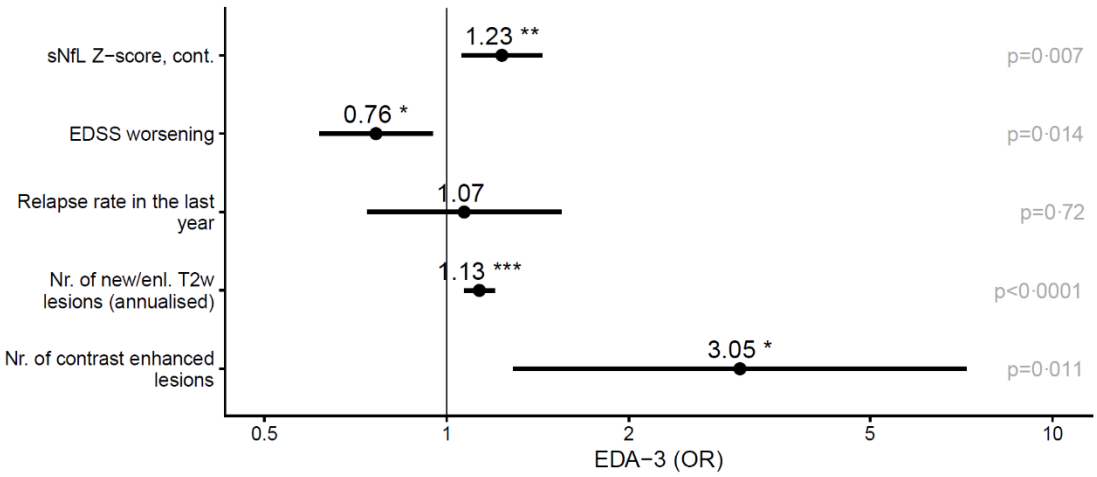
A.



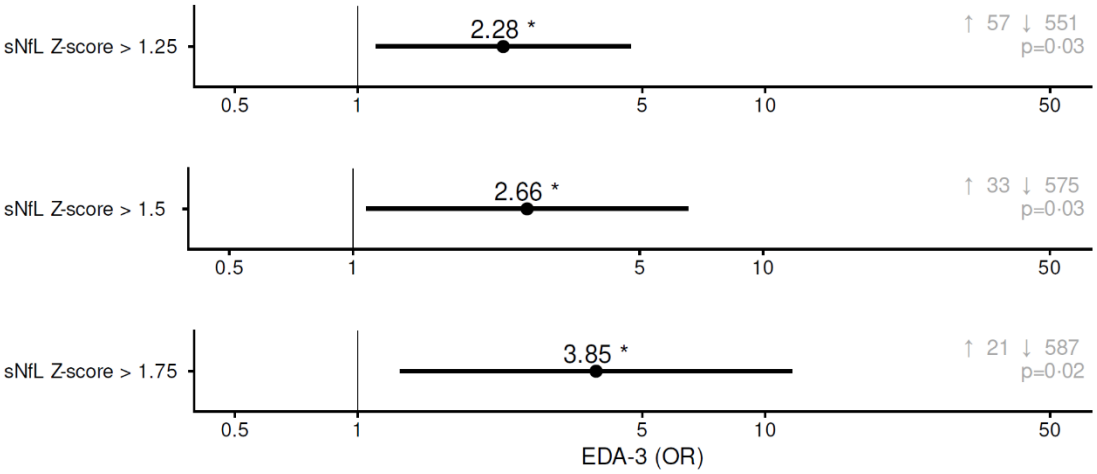
B.



C.



D.



Legend:

A.

Patients with higher sNfL Z-scores showed a higher probability of relapses, EDSS worsening, and EDA-3 in the following year.

B.

An incremental increase of risk of EDA-3 in the following year was observed with increasing sNfL Z-score cut-offs with an up to 3·5-fold risk in patients with sNfL above the 97·7th percentile (Z-score >2) as compared to below.

C.

When combined in a multivariable model with disease activity measures, the risk of EDA-3 in the following year was increased independently by 23% per 1 step higher sNfL Z-score.

D.

NEDA-3 patients with sNfL levels above the 89·4th percentile (Z-score >1·25) displayed a 2·28-fold (95% CI 1·11-4·68; p=0·025) higher risk of experiencing EDA-3 in the following

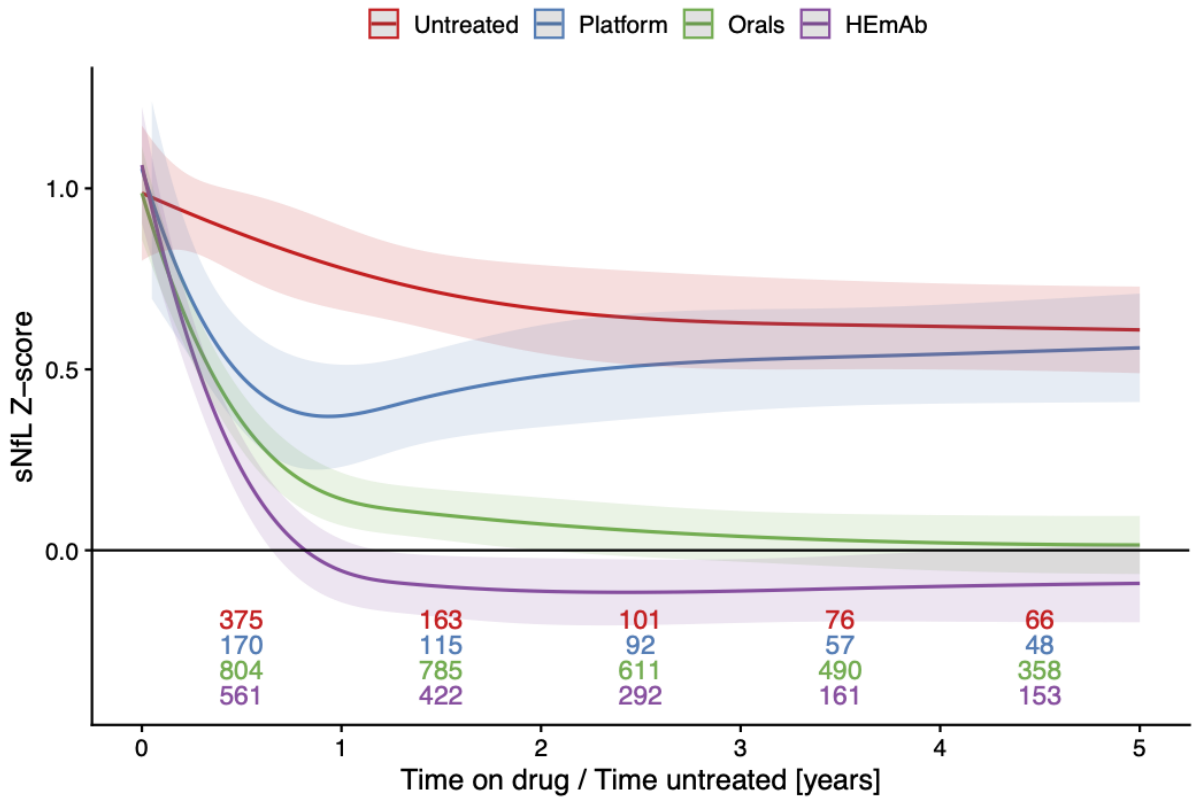
year (3·85-fold in those exceeding the 96·0th percentile (Z-score >1·75); 95% CI 1·27-11·6; p=0·017).

***: p<0·001; **: p<0·01; *: p<0·05. Estimates and 95% confidence intervals are shown. **A.**, **B.** and **D.** show 3 univariable and **C.** a multivariable model.

Grey arrows display number of serum samples above or below the respective sNfL Z-score cut-off.

Abbreviations: CI: confidence interval; EDA-3: evidence of disease activity-3; EDSS: Expanded Disability Status Scale score; NEDA-3: no evidence of disease activity-3; OR: odds ratio; sNfL: serum neurofilament light chain.

Figure 5. Temporal evolution of sNfL Z-scores over time under treatment as estimated by a mixed-effects model.



Legend:

Temporal evolution of sNfL Z-scores over time in four treatment categories using a mixed-effects model thereby using spline terms to model the non-linear temporal association and an interaction term between DMT category and treatment duration. The number of samples in the respective yearly interval is shown in the different treatment groups.

Table

Table 1. Baseline demographic and clinical characteristics of included MS patients (Swiss MS Cohort Study).

Number of patients (n)*	**1313**
Demographic data	
Female (n, %)	883 (67·3)
Male (n, %)	430 (32·7)
Age (Y)	40·5 (31·5, 49·2)
Clinical data, samples and follow-up	
RMS (n, %)	1238 (94·3)
SPMS (n, %)	75 (5·7)
Disease duration (Y)	6·6 (1·9, 13·8)
Nr. relapses in last year (mean, SD)	0·5 (0·70)
EDSS	2·0 (1·5, 3·0)
Nr. of serum samples per patient (n)	6·0 (3·0, 8·0)
Duration of follow-up	5·6 (3·2, 7·2)
Disease-modifying treatment at inclusion	
HEmAb (n, %)	303 (23·1)
Orals	453 (34·5)
Platform	169 (12·9)
Other	12 (0·9)
Untreated	376 (28·6)

* 98.3% of study participants are Caucasian.

Abbreviations: EDSS: Expanded Disability Status Scale; HEmAb: high efficiency monoclonal antibody therapies; MS: multiple sclerosis; Nr.: number; RMS: relapsing MS; SD: standard deviation; SMSC: Swiss MS Cohort; SPMS: secondary progressive MS; Y: years.

Numbers are reported as median and interquartile range if not mentioned differently. "HEmAb" include: alemtuzumab (n=10), natalizumab (n=244), ocrelizumab (n=35), rituximab (n=14); "Orals" include: fingolimod (n=373), dimethyl fumarate (n=71), teriflunomide (n=9); "Platform" includes: all interferon beta (n=122) and glatiramer acetate (n=47) preparations; "Other" includes mitoxantrone (n=7), azathioprine (n=3) and participation in a randomized clinical trial (n=2).

Research in Context

Evidence before this study:

Existing evidence was identified through author experience and PubMed searches through June 2021. Neurofilament light chain (NfL) is the first blood based biofluid marker to reflect neuronal damage in multiple sclerosis (MS), traumatic brain injury and primary neurodegenerative disease of the CNS. As component of the axonal cytoskeleton, the release of NfL into CSF and blood is highly specific for the process of neuronal damage, but not for a specific neurological disease. In MS, serum NfL (sNfL) has been established as marker of acute disease activity (formation of lesions, relapses), of treatment response, and as predictor of the long-term course of disability. However, the application as a biomarker is restricted to group-level analysis, e.g., in clinical trials where relative changes between treatment arms are compared. The routine use of sNfL for personalised medicine was not possible due to the lack of two essential premises a) reference values covering the life span in which MS occurs, and b) a way to correct for the physiologic increase of sNfL in function of age and body mass index (BMI), to derive its disease-specific signal. So far, the ways to correct for this deficit, arbitrary cut-offs to define normal values, yield misleading interpretation of values being 'normal' or 'increased', specifically in the analysis in individual patients, as well in comparisons across groups of variable age and weight.

Added value of this study:

The establishment of a large, statistically robust reference database and the expression of its data as percentiles or Z-scores, i.e., in a way that values are independent of age and BMI, provides a reliable tool for physicians to identify elevated values of sNfL as increased. This allows implementing the earlier prophecy of 'NfL becoming the C-reactive protein of neurologists'[26] for disease monitoring and as factor for therapeutic decision making.[12]

Implications of all the available evidence:

Current clinical measures and standard imaging techniques fall short in identifying subclinical disease activity that is the main driver of the course of disability. The internet-based tool for reference values of sNfL, and the evidence for sNfL as real-time therapy monitoring biomarker, allows physicians the use of sNfL as a biomarker in diagnostic work-up of disease activity in individual MS patients. This closes the diagnostic gap to detect subclinical disease activity in MS with the consequence of a timely choice between therapy options.