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

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

Lancet Neurol 2022; 21: 246–57

See **Comment** page 207

*Contributed equally

Department of Clinical Research, Clinical Trial Unit (P Benkert PhD, S Schaedelin MSc), **Neurologic Clinic and Policlinic, MS Centre and Research Centre for Clinical Neuroimmunology and Neuroscience Basel** (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), **Department of Medicine and Biomedical Engineering, Translational Imaging in Neurology Basel** (C Granziera), **Department of Medicine, Cardiology Division, Cardiovascular Research Institute Basel** (C Mueller 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); **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, ON, Canada** (D Conen MD); **Department of Neurology, Multiple Sclerosis Center, Neurocenter of Southern Switzerland, Lugano,**

Summary

Background Serum neurofilament light chain (sNfL) is a biomarker of neuronal damage that is used not only to monitor disease activity and response to drugs and to prognosticate disease course in people with multiple sclerosis on the group level. The absence of representative reference values to correct for physiological age-dependent increases in sNfL has limited the diagnostic use of this biomarker at an individual level. We aimed to assess the applicability of sNfL for identification of people at risk for future disease activity by establishing a reference database to derive reference values corrected for age and body-mass index (BMI). Furthermore, we used the reference database to test the suitability of sNfL as an endpoint for group-level comparison of effectiveness across disease-modifying therapies.

Methods For derivation of a reference database of sNfL values, a control group was created, comprising participants with no evidence of CNS disease taking part in four cohort studies in Europe and North America. We modelled the distribution of sNfL concentrations in function of physiological age-related increase and BMI-dependent modulation, to derive percentile and Z score values from this reference database, via a generalised additive model for location, scale, and shape. We tested the reference database in participants with multiple sclerosis in the Swiss Multiple Sclerosis Cohort (SMSC). We compared the association of sNfL Z scores with clinical and MRI characteristics recorded longitudinally to ascertain their respective disease prognostic capacity. We validated these findings in an independent sample of individuals with multiple sclerosis who were followed up in the Swedish Multiple Sclerosis registry.

Findings We obtained 10 133 blood samples from 5390 people (median samples per patient 1 [IQR 1–2] in the control group). In the control group, sNfL concentrations rose exponentially with age and at a steeper increased rate after approximately 50 years of age. We obtained 7769 samples from 1313 people (median samples per person 6·0 [IQR 3·0–8·0]). In people with multiple sclerosis from the SMSC, sNfL percentiles and Z scores indicated a gradually increased risk for future acute (eg, relapse and lesion formation) and chronic (disability worsening) disease activity. A sNfL Z score above 1·5 was associated with an increased risk of future clinical or MRI disease activity in all people with multiple sclerosis (odds ratio 3·15, 95% CI 2·35–4·23; p<0·0001) and in people considered stable with no evidence of disease activity (2·66, 1·08–6·55; p=0·034). Increased Z scores outperformed absolute raw sNfL cutoff values for diagnostic accuracy. At the group level, the longitudinal course of sNfL Z score values in people with multiple sclerosis from the SMSC decreased to those seen in the control group with use of monoclonal antibodies (ie, alemtuzumab, natalizumab, ocrelizumab, and rituximab) and, to a lesser extent, oral therapies (ie, dimethyl fumarate, fingolimod, siponimod, and teriflunomide). However, longitudinal sNfL Z scores remained elevated with platform compounds (interferons and glatiramer acetate; p<0·0001 for the interaction term between treatment category and treatment duration). Results were fully supported in the validation cohort (n=4341) from the Swedish Multiple Sclerosis registry.

Interpretation The use of sNfL percentiles and Z scores allows for identification of individual people with multiple sclerosis at risk for a detrimental disease course and suboptimal therapy response beyond clinical and MRI measures, specifically in people with disease activity-free status. Additionally, sNfL might be used as an endpoint for comparing effectiveness across drug classes in pragmatic trials.

Funding Swiss National Science Foundation, Progressive Multiple Sclerosis Alliance, Biogen, Celgene, Novartis, Roche.

Research in Context

Evidence before this study

We identified existing evidence through author knowledge and PubMed searches from database inception, to Sept 30, 2021 using the search terms: "neurofilament" and "multiple sclerosis". In multiple sclerosis, serum neurofilament light chain (sNfL) has been established as a marker of acute disease activity (eg, formation lesions and relapses), treatment response, and as a predictor of the long-term course of disability. However, the application of sNfL as a biomarker is restricted to group-level analyses, and its routine use in personalised medicine has not yet been possible. Arbitrary cutoffs to define normal values yield misleading interpretation of values as normal or increased, specifically for individuals and in comparisons across groups of variable age and weight.

Added value of this study

We established a large and statistically robust reference database using data from four cohort studies in Europe and North America that included people without any documented CNS disease. We expressed these data as percentiles and Z scores, adjusted for age and BMI, to create a new method with which clinicians can identify and interpret elevated values of sNfL. We tested and validated this new method, which has been developed into an internet-based app, in two large and independent cohorts of people with multiple sclerosis. We showed that elevated NfL Z scores were associated with increased risk of future disease activity and demonstrated that sNFL Z scores in longitudinal samples can be used to compare the long-term effectiveness of disease-modifying therapies in a real-world setting.

Implications of all the available evidence

Current clinical measures and standard imaging techniques are inadequate for identification of subclinical disease activity, which is the main driver of the course of disability in people with multiple sclerosis. The internet-based app for reference values of sNfL, and the evidence for sNfL as a real-time therapy monitoring biomarker, allows clinicians to use sNfL as a biomarker in the diagnostic work-up of disease activity in individual people with multiple sclerosis. This ability closes the diagnostic gap in the detection of subclinical disease activity in people with multiple sclerosis with a timely choice between therapy options.

Introduction

Multiple sclerosis is a chronic inflammatory and neurodegenerative disease of the CNS characterised by acute deterioration of neurological function (relapse) and chronic accumulation of relapse-independent disability (progression). In the past three decades, increasingly effective disease-modifying therapies have led to ground-breaking success in suppressing relapses and its MRI correlate, focal brain lesion formation.[1] However, 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] ie, no relapses, no clinically significant increase in Expanded Disability Status Scale [EDSS], no new or enlarging T2-weighted lesions, and no T1-weighted contrast-enhancing lesions on brain MRI), has become a treatment goal for multiple sclerosis and a new outcome measure in clinical trials.[1–4] However, fewer than 8% of individuals keep NEDA-3 status. Moreover, this outcome was not associated with better EDSS outcomes 7–8 years later.[4,5] Cree and colleagues have also called into question the utility of annual MRI assessments as a treat-to-target approach for long-term multiple sclerosis care.[5] Furthermore, there is no biofluid marker available in clinical practice to monitor a patient's response to drugs or to predict the course of disease progression.[6] Accordingly, no common denominator endpoint has been established for objective evaluation of the relative effectiveness of disease-modifying therapies, and head-to-head comparisons of modern, high-efficacy, disease-modifying therapies are scarce.[2]

Neurofilament light chain (NfL) is a neuroaxonal cytoskeletal protein that is released into the CSF, and eventually into blood, on neuronal injury.[7] It was the first serum biomarker shown at the group level (eg, in clinical trials where relative changes between treatment arms are compared) to reflect acute disease activity (relapse and lesion formation) in people with multiple sclerosis, to correlate with therapy response, and to predict the course of disability worsening.[7–14] Serum NfL (sNfL) provides a rater-independent quantification of the intensity of ongoing neuronal damage based on a standardised assay platform.[7] Therefore, sNfL could serve as a common denominator for the objective comparative assessment of drug effectiveness across all disease-modifying therapies.[15] However, sNfL is not a stable measure because it increases physiologically with age[7] and decreases with body-mass index (BMI).[16,17] These physiological modulators hamper the validity of fixed cutoff values to define pathological levels for individuals, and they limit the use of sNfL as a biomarker for group-level comparisons, for which (through randomisation or other ways of adjustment) these confounding factors can be controlled. 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) comorbidities that affect sNfL concentrations.

We aimed to derive percentiles and Z scores for sNfL from a large reference database established from a general population, to define levels of pathological increase independent of BMI and age. Our objective was to test, in two large and independent cohorts of people with multiple sclerosis, whether these adjusted sNfL measures would predict the risk for future disease activity, both at the group level and in individuals in clinical practice. We also

Switzerland (C Gobbi MD); **Department of Clinical Neurosciences, Service of Neurology, Lausanne University Hospital, 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 H 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, CA, USA** (J R Oksenberg PhD); **Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland** (C Gobbi); **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, London, UK** (H Zetterberg); **UK Dementia Research Institute at UCL, London, UK** (H Zetterberg); **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); **Department of Health Sciences, University of Genoa, Genoa, Italy** (M P Sormani PhD)

Correspondence to:
Prof Jens Kuhle, Multiple Sclerosis Centre and Research Centre for Clinical Neuroimmunology and Neuroscience Basel, Neurology, Departments of Head, Spine and Neuromedicine, Biomedicine and Clinical Research, University Hospital Basel, University of Basel, Basel 4031, Switzerland
jens.kuhle@usb.ch

aimed to investigate whether the sNfL percentiles and Z scores could be used to quantify and compare the long-term effectiveness of disease-modifying therapies.

Methods

Study design and participants

For derivation of the reference database of sNfL values, we assembled a control group from participants in four European and US population-based studies and control groups of genetic multiple sclerosis studies spanning over six decades of life. People in these cohorts did not have documented CNS disease. The origins and characteristics of these four cohorts are described in the appendix (pp 4–5, 8).

See Online for appendix

For testing of the reference database, we used prospectively collected data from participants in the Swiss Multiple Sclerosis Cohort (SMSC),[18] which is a cohort study at eight academic medical centres in Switzerland (appendix p 4). All individuals in SMSC with a diagnosis of relapsing or secondary progressive multiple sclerosis, defined according to Lublin and colleagues,[19] were included in our analysis.

For validation of the findings, we included prospective collected data from people with multiple sclerosis in the Swedish Multiple Sclerosis registry, which comprises three partly overlapping large cohorts: the Epidemiological Investigation of Multiple Sclerosis (EIMS),[20] Immunomodulation and Multiple Sclerosis Epidemiology (IMSE),[21] and Comparison Between All immuno-Therapies for Multiple Sclerosis (COMBAT-MS; appendix p 4).[22]

Institutional review boards at the respective SMSC centres and the Stockholm regional ethics committee approved this study. Written informed consent was obtained from all participants.

Procedures

From each cohort, we obtained relevant data for our analysis, including participant's age, sex, BMI, and for the multiple sclerosis cohorts, clinical variables (eg, EDSS score, estimated glomerular filtration rate [eGFR], disease duration, and type of multiple sclerosis), current treatment, and MRI parameters. Treatments were categorised into high-efficacy monoclonal antibody therapies (alemtuzumab, natalizumab, ocrelizumab, and rituximab), oral therapies (dimethyl fumarate, fingolimod, siponimod, and teriflunomide), platform compounds (interferon beta and glatiramer acetate), and untreated (appendix p 4).

Blood samples were obtained from all controls and people with multiple sclerosis. In people with multiple sclerosis from SMSC and all controls, sNfL was measured in duplicate with the NF-light assay (Quanterix, Billerica, MA, USA) according to the protocol provided by the company. Intra-assay and inter-assay variability was evaluated with three native quality control serum samples during 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 of more than 20% were repeated. The mean coefficients of variation of duplicate determinations for concentration were 5·2% (6·2 pg/mL, sample 1), 3·1% (18·8 pg/mL, sample 2), and 3·0% (37·1 pg/mL, sample 3). The interassay coefficients of variation were 6·9% (sample 1), 5·5% (sample 2), and 5·8% (sample 3). In the validation cohort comprising people with multiple sclerosis from the Swedish Multiple Sclerosis registry, NfL was measured by NF-light assay[23,24] in duplicate in plasma samples (pNfL) treated with EDTA (edetic acid; appendix pp 5–6).

Statistical analysis

We used data from the control group to model the relation between sNfL, age, and BMI, to create the reference database. In the appendix (pp 6–7), we have explained our reasoning for inclusion of BMI and age, but not diabetes, and for excluding a few samples with an eGFR of less than 60 mL/min per 1·73 m². We have also described in detail the selection of one sample from each control person, the modelling procedures, how generalisability of the resulting reference database was tested, and how overtraining of the final reference database was ruled out (appendix pp 6–7, 13–18).

We used a generalised additive model for location, scale, and shape. From this model, percentiles and Z scores were calculated as two interchangeable measures that quantify the deviation of sNfL values from the control group.[25] Percentiles express the percentage of the general population expected to have an sNfL value (adjusted for age and BMI) 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.

Multivariable linear mixed-effects models with a random intercept for the patient were used to investigate associations between sex, clinical variables, and MRI parameters of disease worsening (either active disease [relapse, or T1-weighted contrast enhancing lesions] or progression [EDSS score, or hyperintense T2-weighted lesion volume]) and disease-modifying therapies, with longitudinal sNfL Z scores as a dependent variable. The estimates represent additive effects on the sNfL Z score.

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]; appendix pp 4–7) or recent disease activity (relapse ≤4 months). High sNfL was defined as the portion of samples with highest values (separately based on absolute sNfL concentration and based on sNfL Z score) using three different cutoffs—ie, the top 25% (ie, first quartile), top 10%, and top 5% of all samples. Generalised linear (logistic) mixed-effects models, with future and recent disease activity as dependent variables, were generated with the dichotomised variable (based on absolute sNfL concentration or sNfL Z scores) as the only predictor, and odds ratios (ORs) are presented. For comparison between the SMSC

and the validation cohort from the Swedish Multiple Sclerosis registry, identical absolute values of NfL for cutoffs were used.

We analysed the performance of sNfL Z scores to quantify the risk of future disease activity. In univariable generalised linear (logistic) mixed-effects models, sNfL Z scores were included as a continuous variable, and cutoffs were used (ie, sNfL Z scores above *vs* below 1·0, or 1·5, or 2·0) to predict future disease activity (ie, occurrence of relapse, EDSS worsening, or EDA-3) in the following year.

To quantify the potentially added contribution of sNfL Z score to predict the risk of future (following year) EDA-3 status, we combined disease activity measures currently used in clinical practice (eg, EDSS worsening [appendix p 4], rate of relapse in the past year, new and enlarging T2-weighted lesions in the past year, and current contrast enhancing lesions), with sNfL Z scores in multivariable generalised linear (logistic) mixed-effects models. The fit of the two alternative multivariable models (including and excluding sNfL Z score) was compared with the χ^2 test.

Finally, to quantify the risk of future (following year) EDA-3, we analysed the performance of sNfL Z score cutoffs (dichotomising with the above *vs* below cutoffs) in people currently (past year and present) fulfilling NEDA-3 criteria (ie, without clinical or MRI evidence of disease activity; appendix p 4). We used univariable generalised linear (logistic) mixed-effects models in these stable patients (defined clinically and according to conventional MRI).

To model disease activity as expressed by sNfL Z scores under specific disease-modifying therapy categories, a multivariable model with sNfL Z score as dependent variable was built using treatment regimen (disease-modifying therapy categories or untreated) and time since the start of treatment (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 disease-modifying therapy groups. The non-linear dynamics in disease activity over time was modelled using spline terms for time under treatment and 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 disease-modifying therapy groups over time were extracted and plotted together with 95% CIs, using the R package sjPlot.[26] As a sensitivity analysis, a model adjusted for demographic and clinical covariates (ie, sex, age, disease duration, secondary progressive multiple sclerosis *vs* relapsing multiple sclerosis , presence of relapse in the past 4 months, EDSS) was built (appendix p 19). All analyses were done using the statistical software package R (version 4.0.4) using two-sided tests.

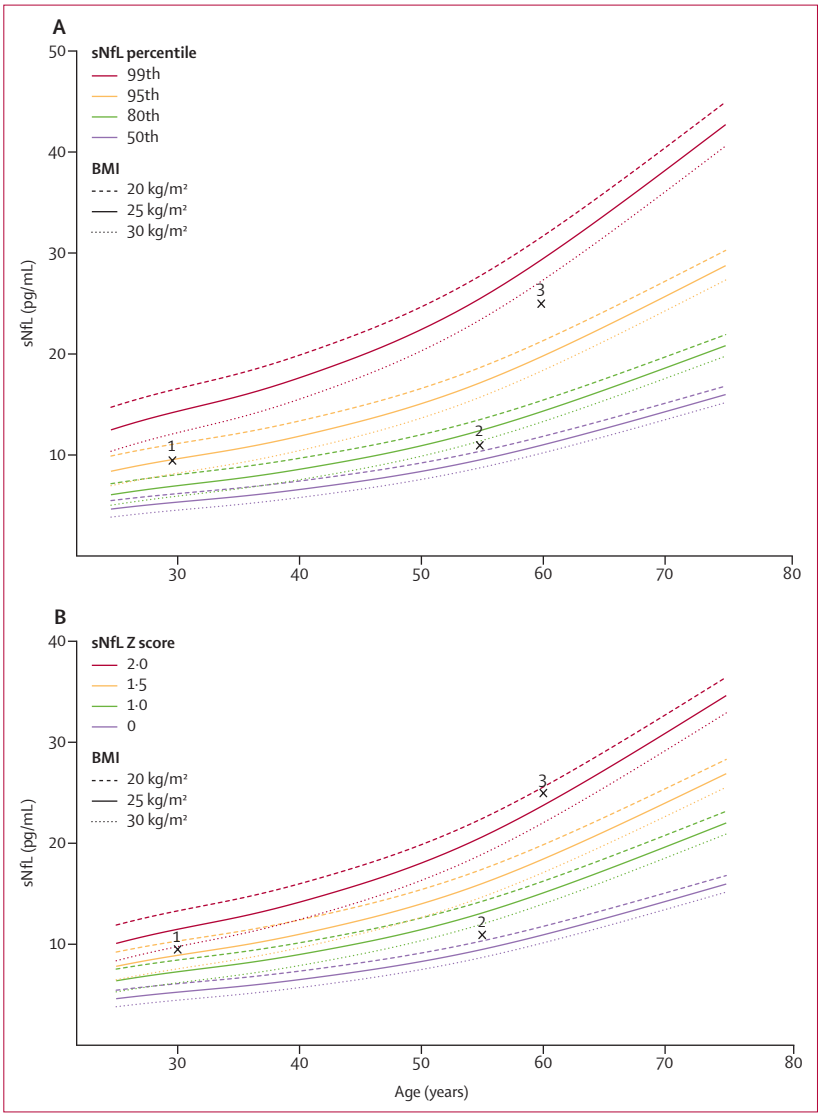


Figure 1: **sNfL percentiles (A) and Z scores (B) reference curves**
A generalised additive model for location, scale, and shape was used to model the association of sNfL concentration (pg/mL) in controls with data for BMI and age. Example 1, at 30 years and a BMI of 25 kg/m², shows sNfL of 9·5 pg/mL (95th percentile) and Z score of more than 1·5 (exact value 1·64, as calculated by the sNfL app), and the interpretation is elevated. Example 2, at 55 years with a BMI of 25 kg/m², shows sNfL of 11·0 pg/mL, below the 80th percentile (calculated as 68th percentile) and a Z score of less than 1·0 (calculated as 0·47), which is similar to levels seen in controls. Example 3, at 60 years and a BMI of 30 kg/m², shows sNfL of 25 pg/mL, close to the 99th percentile (calculated as 98·6th percentile) and a Z score of more than 2 (calculated as 2·2), and the interpretation is elevated. BMI=body-mass index. sNfL=serum neurofilament light chain.

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
10 133 serum samples (samples available per control person: median 1 [IQR 1–2]) from 5390 people without evidence of CNS disease were available for creation of the reference database of sNfL percentiles and Z scores values

	Number of participants (n=1313)
Demographic data	
Sex	
Female	883 (67·3%)
Male	430 (32·7%)
Age, years	40·5 (31·5–49·2)
Ethnicity	
White	1291 (98·3%)
Other	22 (1·7%)
Clinical data, samples, and follow-up	
Disease course	
Relapsing multiple sclerosis	1238 (94·3%)
Secondary progressive multiple sclerosis	75 (5·7%)
Disease duration, years	6·6 (1·9–13·8)
Relapses in past year, n	0·5 (0·70)
EDSS score	2·0 (1·5–3·0)
Serum samples per patient, n	6·0 (3·0–8·0)
Duration of follow-up, years	5·6 (3·2–7·2)
Disease-modifying treatment at inclusion	
High-efficiency monoclonal antibody therapies*	303 (23·1%)
Oral therapies†	453 (34·5%)
Platform compounds‡	169 (12·9%)
Other§	12 (0·9%)
Untreated	376 (28·6%)

Data are n (%), mean (SD), or median (IQR). EDSS=Expanded Disability Status Scale. *Alemtuzumab (n=10), natalizumab (n=244), ocrelizumab (n=35), rituximab (n=14). †Fingolimod (n=373), dimethyl fumarate (n=71), and teriflunomide (n=9). ‡Interferon beta (n=122) and glatiramer acetate (n=47) preparations. §Mitoxantrone (n=7), azathioprine (n=3), and participation in a randomised clinical trial (n=2).

Table: **Baseline demographic and clinical characteristics of people with multiple sclerosis from the Swiss Multiple Sclerosis Cohort**

(appendix p 8). We have presented reference values in figure 1, in the appendix (p 9), and as an intenet-based app (appendix p 33).[27] The age-related increase of sNfL percentiles and Z scores in the control population was not linear (figure 1). Further analysis showed that the increase was exponential but with an inflection point around 50 years of age, with a steeper increase thereafter (appendix p 20). Lower levels of sNfL were seen with higher BMI. After age adjustment, BMI showed a constant but inverse correlation with sNfL (appendix p 13; figure 1).

3105 (58%) of 5390 people in the control population contributed several serum samples at different time-points. Whereas we only used one sample per patient in the final reference database (n=4532; appendix p 7), all available samples were used for sensitivity analyses. These samples confirmed that the shapes and positions of percentile and Z score reference curves were insensitive to alterations of the underlying reference dataset (ie, using alternative selections of samples per control person [appendix p 17] and using bootstrapping [appendix p 18]).

1313 people participating in the SMSC, with a disease course classified as relapsing or secondary progressive multiple sclerosis, were included in our analysis (table). The age distribution of people was congruent with that seen for the reference database population (appendix p 15). At entry into the SMSC, 376 (28·6%) people were untreated, 169 (12·9%) were on platform compounds, 453 (34·5%) were on oral therapies, and 303 (23·1%) were on high efficacy monoclonal antibody therapy (table). Over a median follow-up period of 5·6 (IQR 3·2–7·2) years, 121 (9·2%) of 1313 individuals remained untreated, 788 people (60·0%) were treated with one compound from these disease-modifying therapy classes, and 404 people (30·8%) were treated with more than one class of disease-modifying therapy. A total of 7769 serum samples were obtained from 1313 participants in the SMSC, with a median number of samples per person of 6·0 (IQR 3·0–8·0; table; appendix pp 11–12).

In the multivariable mixed-effects model with sNfL Z scores as a dependent variable, clinical and MRI measures of disease worsening or progression were strongly and independently associated with higher sNfL Z scores. Furthermore, a treatment effectiveness hierarchy was seen, compared with untreated people, of high efficacy monoclonal antibody therapies over oral therapies and of oral therapies over platform compounds (figure 2). This hierarchy was supported by results in the validation cohort (appendix p 21). The estimated additive effects on sNfL Z score were –0·14 (95% CI –0·23 to 0·05; p=0·0018) for high efficacy monoclonal antibody therapy versus oral therapy, and –0·23 (–0·36 to 0·10; p<0·0001) for oral versus platform therapy.

Similar to the results in the control group, absolute sNfL values in people with multiple sclerosis rose with age (figure 3). Increased sNfL concentrations measured by higher Z scores were more frequent in younger versus older individuals.

A conservative cutoff of 10 pg/mL was used as an arbitrary definition of a non-pathological sNfL concentration (figure 3). With this approach, in people aged 20–30 years, 70 (68%) of 103 with Z scores of 1·5–2·0 and seven (4%) of 164 with Z scores of more than 2·0 would be declared as having sNfl concentrations within normal range (≤10 pg/mL). However, compared with people with sNfL Z scores of 1·5 or less and sNfL below 10 pg/mL, these 77 people showed more recent clinical disease activity (p=0·023) and fulfilled concurrent EDA-3 status more frequently (p=0·016; appendix pp 10, 22). Moreover, the people with increased Z scores (>1·5) showed a higher propensity for clinical disease activity (p=0·041) and numerically fulfilling EDA-3 status (p=0·22) in the following year (appendix pp 10, 22). Conversely, in the age range of 30–60 years, 989 (39·1%) of 2517 people with a normal Z score (0–1·5) would be labelled as having elevated (>10 pg/mL) sNfL concentrations (appendix p 10). The mismatch between these two ways to define normal

values becomes more pronounced in individuals older than 60 years, since 292 (100%) of 292 with Z score ranges of 0–1·5 and 156 (50%) of 310 with Z scores of 0 or below are above the 10 pg/mL cutoff.

Using three different threshold levels for high and low samples, increased sNfL Z scores and absolute sNfL concentrations both showed a higher likelihood for disease activity in the following year (EDA-3; p<0·0001 for all six estimates [appendix p 23] and for the validation cohort [appendix p 24]). However, Z scores consistently led to higher ORs than did absolute sNfL values when using the three different cutoffs for a sample defined as high (ie, top 25%, top 10%, and top 5%). For ORs of absolute sNfL concentrations versus sNfL Z scores, the top 25% resulted in ORs of 2·09 *vs* 3·09; the top 10% in 2·83 *vs* 3·84; and the top 5% in 2·53 *vs* 4·43, which corroborates the superior performance of sNfL Z scores over fixed cutoff levels of absolute sNfL values, irrespective of where cutoff values were set. Accordingly, the association between a recent relapse (≤4 months) and sNfL Z scores was considerably stronger versus absolute sNfL concentrations in the validation cohort (appendix pp 25–26).

sNfL percentiles and Z scores were used as measures and predictors of future disease activity in multiple sclerosis. People with higher sNfL Z scores showed a greater probability of relapses (OR 1·41, 95% CI 1·30–1·54; p<0·0001), EDSS worsening (1·11, 1·03–1·21; p=0·0093), and EDA-3 (1·43, 1·31–1·57; p<0·0001; figure 4A; for the validation cohort, appendix p 27) in the following year, based on a model with Z score as a continuous predictor.

As compared with the continuous analysis, the use of sNfL Z score cutoffs led to a substantially higher probability of EDA-3 in the following year (figure 4B), by incremental increases of cutoff levels (validation cohort, appendix p 27). A sNfL Z score above 1·5 was associated with an increased risk of future clinical or MRI disease activity in all people with multiple sclerosis (OR 3·15, 95% CI 2·35–4·23; p<0·0001; figure 4B), and in people considered stable with no evidence of disease activity (2·66, 1·08–6·55; p=0·034; figure 4D).

When sNfL Z scores are combined with disease activity measures currently used in clinical practice in a multivariable model, the risk of EDA-3 in the following year was increased independently (OR 1·23, 95% CI 1·06–1·44; p=0·0072; figure 4C; validation cohort, appendix p 27). It is noteworthy that model quality was improved when sNfL Z scores were included together with all classic measures of disease activity shown in figure 4C (χ^2; p=0·0023) as compared with the same model without sNfL Z scores.

The clinical consequence of increased sNfL Z scores in people with NEDA-3 was a higher likelihood for EDA-3 status in the following year. For example, sNfL concentrations were higher than the 89·4th percentile (ie, a Z score >1·25) in 57 (9%) of 608 serum samples from people being classified as NEDA-3 since the past

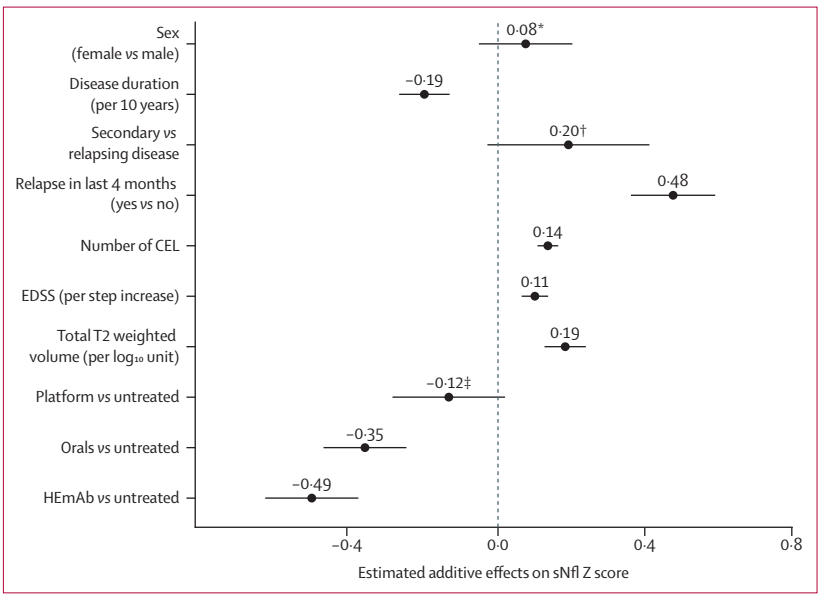


Figure 2: **Factors affecting sNfL Z scores in people with multiple sclerosis**
Model estimates including 95% CIs (see appendix [p 12] for numerical values). All values are p<0·0001, unless specified in the footnotes. CEL=contrast-enhancing T1-weighted lesions. EDSS=Expanded Disability Status Scale score. HEmAb=high efficacy monoclonal antibody therapies. sNfL=serum neurofilament light chain. *p=0·20. †p=0·076. ‡p=0·11.

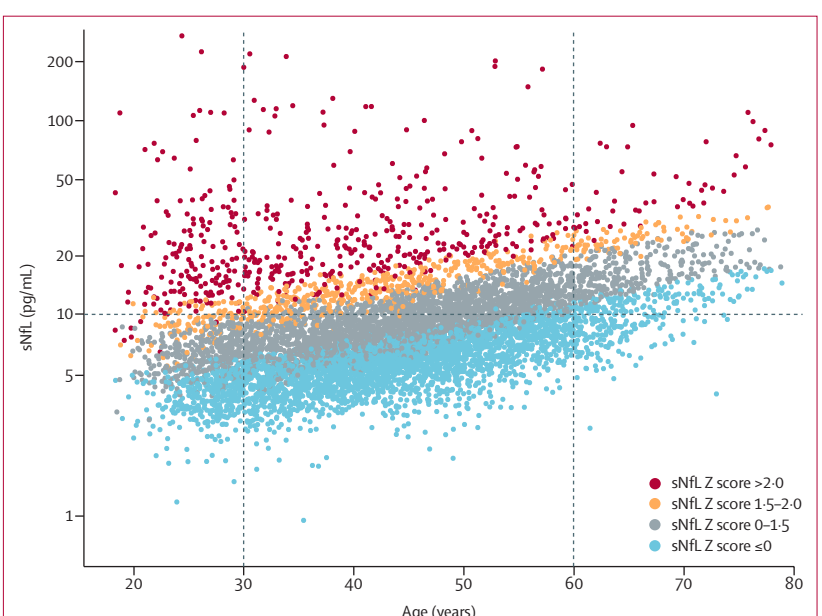


Figure 3: **sNfL Z scores according to age of people with multiple sclerosis participating in the Swiss Multiple Sclerosis Cohort**
Age-adjusted and BMI-adjusted sNfL Z scores are shown by colour gradient. The fixed sNfL cutoff is shown by the horizontal line at 10 pg/mL. Using a fixed cutoff in people with multiple sclerosis aged 20–30 years might miss people with increased sNfL Z scores (false negatives: yellow and red dots below horizontal 10 pg/mL cutoff). Conversely, in people older than 30 years, a large proportion of individuals with normal age-corrected sNfL (ie, sNfL Z scores 0–1·5 [grey], ≤0 [blue]) show values above the fixed threshold of pathology (false positives). Numerical values are provided in the appendix (p 10). Different Z scores can occur with similar sNfL concentrations and identical age, because of additional adjustment for BMI. sNfL=serum neurofilament light chain.

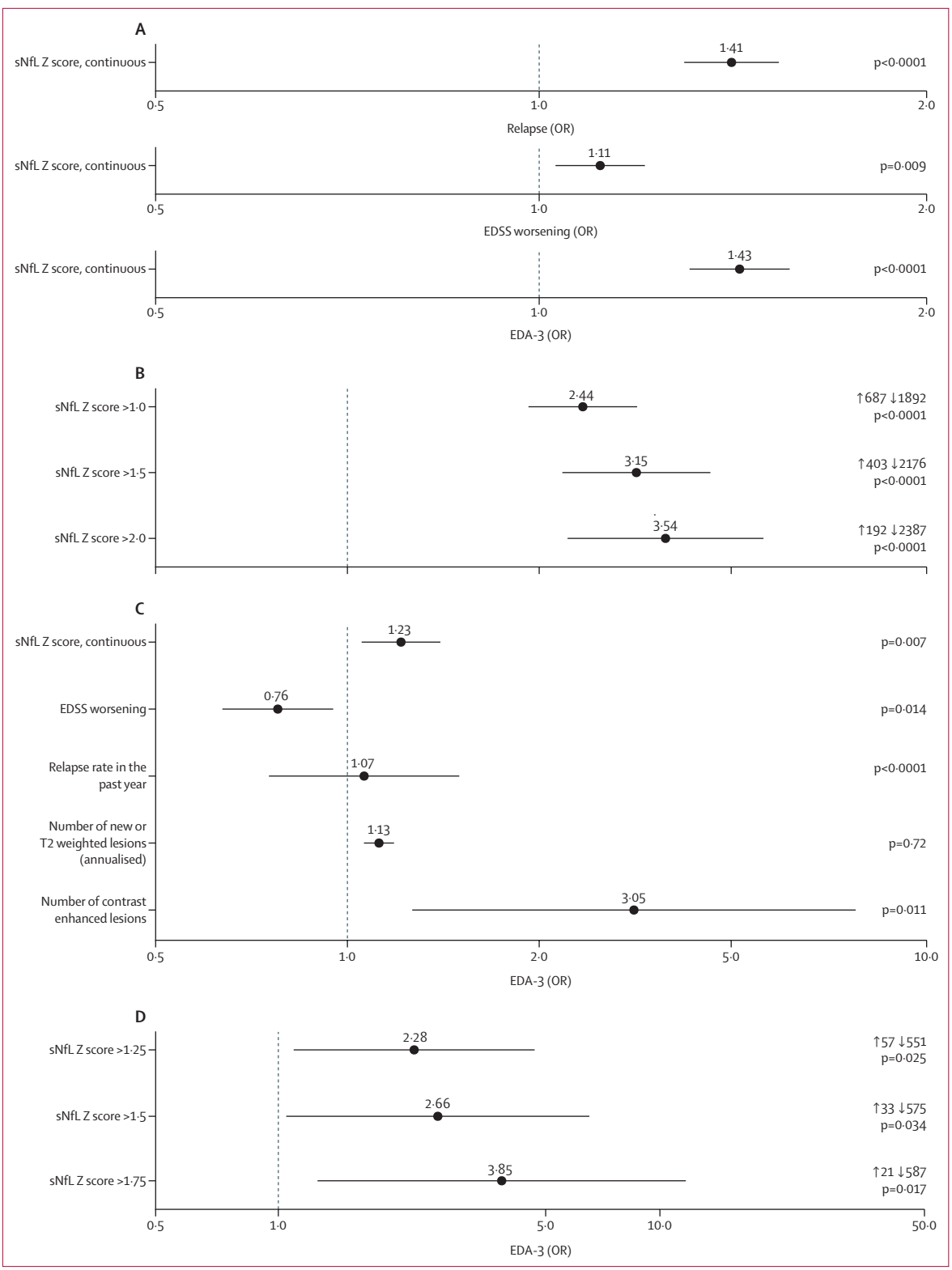
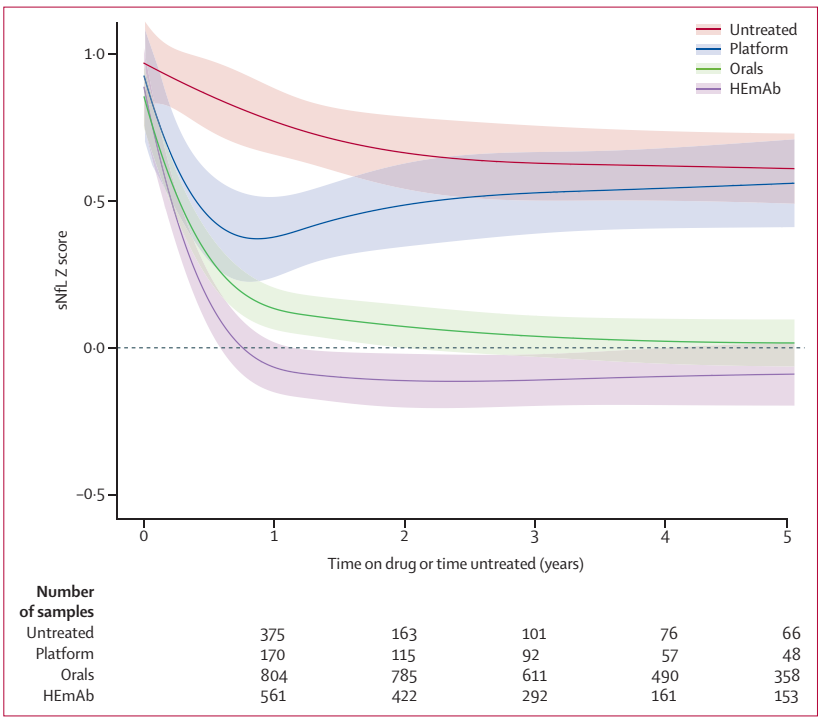


***Figure 4*: sNfL Z scores predicting disease activity in the following year**
Estimates (ORs) and 95% CIs are shown. A, B, and D show three univariable models and C shows a multivariable model. Arrows display number of serum samples above or below the respective sNfL Z score cutoff. Probability of occurrence of relapses or EDSS worsening or EDA-3 in the following year based on (continuous) sNfL Z score (A); using sNfL Z score cutoffs (B); in combination with other currently used measured of disease activity in clinical practice in a multivariable model (C); and in people with NEDA-3 (D). 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.

year. These cases displayed a higher risk (OR 2·28, 95% CI 1·11–4·68; p=0·025) of experiencing any sign of clinical or MRI disease activity over the following year (figure 4D). This risk increased in people with sNfL concentrations exceeding the 96·0th percentile (ie, a Z score >1·75 OR 3·85, 1·27–11·63; p=0·017; figure 4D; validation cohort, appendix, p 27).

In the mixed-effects model of disease activity and long-term treatment effects of disease-modifying therapy categories, the evolution of sNfL Z scores over time was assessed in the four treatment categories. In the first year after initiation of therapy, sNfL concentrations decreased rapidly in treated individuals, whereas they fell only marginally in untreated people (figure 5). The reduction of the sNfL Z score was more rapid with high efficacy monoclonal antibody therapies, compared with oral therapies and platform compounds, as reflected by the steeper slope of the line (p<0·0001 for the interaction term between treatment category and treatment duration). Over the following 4 years, high efficacy monoclonal antibody therapies and, to a lesser extent, oral therapies showed sNfL concentrations that overlapped with those of the control population (ie, sNfL Z score 0), whereas with platform compounds the sNfL concentrations remained increased. Platform compounds were associated with the weakest sNfL reduction in the first year of treatment, and were followed by a new increase thereafter, coming close to concentrations measured in untreated people. As a sensitivity analysis, a model adjusted for demographic and clinical covariates supported 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 (appendix p 19).

The appendix (p 30) shows seven clinical use cases from the SMSC for the application of sNfL percentiles and Z scores as a 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, an internet-based app was created based on sNfL values from the reference database, to determine Z scores and respective percentile values by entering individuals' 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 (appendix p 33) online.



Number of samples					
Untreated	375	163	101	76	66
Platform	170	115	92	57	48
Orals	804	785	611	490	358
HEmAb	561	422	292	161	153

Figure 5: **Temporal evolution of sNfL Z scores under treatment**
Four treatment categories were included in a mixed-effects model, thereby using spline terms to model the non-linear temporal association and an interaction term between disease-modifying therapy category and treatment duration. The number of samples in the respective yearly interval is shown in the different treatment groups. Shaded areas indicate 95% CI. HEmAb=high efficacy monoclonal antibody therapies.

Discussion

Our results show that NfL can be used as a biomarker for monitoring of treatment efficacy and prognostication of disease course in individual people with multiple sclerosis. A reference database with age-adjusted and BMI-adjusted sNfL concentrations was created using samples from a general population with no documented CNS disease. Statistical transformation of sNfL concentrations from absolute values into percentiles and Z scores allowed us to reliably correct for confounding factors to discern pathological from physiological levels of sNfL. This database and transformation was subsequently tested in two large, independent, real-world multiple sclerosis cohorts. Moreover, our results showed that sNfL can be used as an additional measure of disease activity (EDA-3) besides clinical assessments and MRI. It is specifically useful for stable people (ie, in NEDA-3 status) to identify ongoing disease activity that is below the detection threshold of standard clinical and MRI markers. Using the reference database, sNfL concentrations can also be applied for the quantitative comparison of long-term effectiveness across disease-modifying therapies (while considering limitations based on design preventing proof of causation in real-world settings).

In 2018, Giovannoni described NfL as "the neurologist's C-reactive protein"[28] for measurement of the neuroprotective effects of disease-modifying therapies in the context of clinical trials. Since then, clinical studies have shown how sNfL can quantify disease activity in multiple sclerosis and other neurological disorders.[29] Moreover, phase 3 studies in multiple sclerosis have used sNfL as an exploratory endpoint for treatment efficacy.[11,30–32] Despite these studies showing that sNfL accurately reflects even subclinical disease activity,[30,32] sNfL has not been generally accepted as a clinical routine biomarker for individual people with multiple sclerosis, nor as a

For the **online application** see https://shiny.dkfbasel.ch/baselnflreference

primary or secondary trial endpoint. By contrast to C-reactive protein, sNfL did not have two essential premises for such a breakthrough—namely, reference values from a general population who did not have clinically manifest diseases,[7,29] and a way to interpret values without interfering the factor of age and BMI.

With the advent of high-efficacy multiple sclerosis therapies, relapses and high rates of lesion formation have been suppressed almost completely. We now need to ask, how should we control the subclinical diffuse brain damage that manifests clinically as continuously worsening disease (progression), and how should we measure it? Since sNfL concentrations remain modestly raised in the progressive disease state, compared with the more pronounced NfL concentration increases that are associated with relapses,[33] the task to discern disease signal from age-related changes becomes more challenging. The earlier assumption of a constant increase of 2·2% per year of sNfL in controls[7] was based on cohorts[8,9,34] that were too small and insufficiently covered the age range specifically relevant for progressive multiple sclerosis. Data from our reference database show that the evolution of sNfL with age follows a non-log-linear function, and they establish BMI as an important additional modulator of NfL concentrations in reference populations. By consequence, fixed cutoffs to define pathological sNfL levels could lead to a misclassification, even if the cutoff is set at a lower level in the present analysis than in earlier ones.[35,36] Various fixed cutoffs to define pathological sNfL concentrations have been used previously.[11,35,37] We used a conservative cutoff of 10 pg/mL for an arbitrary definition of a non-pathological sNfL level. Current results show that a substantial proportion of young people (<30 years) with multiple sclerosis have ongoing disease activity that would remain unrecognised using such fixed cutoff levels and, hence, the purpose of measuring sNfL to guide therapeutic decisions might be missed. Additionally, the inclusion of BMI to define reference percentiles and Z scores further increases the precision in determining pathological cutoff 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 cutoff might lead to a significant false-positive rate in individuals older than 40 years, which is problematic for the interpretation of sNfL concentrations in people with progressive multiple sclerosis or primarily neurodegenerative diseases.[38]

Z scores are a standard measure in other fields of medicine—eg, echocardiographic measurement of aortic dilation, or determination of bone mineral density to separate pathology-indicating signals of biomarkers from physiological longitudinal changes.[25,39] Percentiles (which are used, for example, in paediatric growth curves) are akin to Z scores, a derivative of standard deviation calculations, and are a very similar way to describe deviation from normality in medicine.[25] 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 quantification not only of the contribution of clinical and MRI features to disease activity but also of effectiveness of therapy categories of disease-modifying therapy. Clinicians have the choice between more than ten registered disease-modifying therapies for multiple sclerosis. However, 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 for methodological reasons. With the reference database and Z scores, we can now model the effectiveness of drugs and of residual disease activity over years of treatment. High efficacy monoclonal antibody therapies, and to a lesser extent oral therapies, coincide with a normalisation of sNfL concentrations over time. By contrast, the diminishing treatment effect of platform compounds 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.[40,41]

Our study has several limitations. The reference database is based on a cohort of people without clinical manifestation of somatic disease. However, many subclinical disease conditions could be associated with an increase of sNfL concentration due to neuronal damage to the nervous system. For example, underlying primary neurodegenerative diseases (eg, Alzheimer's disease) can lead to an increase in NfL concentrations years before they clinically manifest.[38] On purpose, we did not establish our reference database on a cohort of people for whom subclinical laboratory aberrations have been excluded—ie, whose serum samples were selected for absence of neurodegenerative or other diseases developing later in life. Such diseases can occur as well with similar incidence and prevalence in people with multiple sclerosis. Hence, with a view to use the reference database percentiles and Z scores in clinical real-world practice for people with multiple sclerosis, we did not pursue the concept to correct for such comorbidities occurring at later stages in life.

Although we have acquired limited data that mild renal insufficiency and diabetes have little effect on sNfL concentrations, we need to define how more severe stages of these diseases, and possibly other confounding factors, limit the interpretability of findings in people with multiple sclerosis. Our results are largely based on people with relapsing multiple sclerosis who are White; therefore, the generalisability of our data for people with primary progressive multiple sclerosis and in people with different ethnic backgrounds needs to be validated in referring cohorts. It is not known whether data acquired with the current standard assay system (Simoa; Quanterix, Billerica, MA, USA) are fully compatible with those of other analytical platforms for NfL, given that they provide highly correlated but different absolute

values. Standardisation efforts are now ongoing within the International Federation of Clinical Chemistry, aiming for developing Certified Reference Materials for harmonisation of readouts across platforms. In essence, the use of our internet-based percentile and Z score tool requires that data are acquired with the standard kit and on the same hardware platform.

In conclusion, sNfL percentiles and Z scores could be used as a clinical methods to identify subclinical disease activity in individual people with multiple sclerosis and to monitor drug response. It is now available for clinicians by use of an internet-based app. This app can also be used in future trials, in which sNfL is an endpoint measure.

Contributors

PB, SSc, CG, LK, HW, KBe, CGr, DL, and JK conceptualised the study. PB, SMe, SS, AMan, ÖY, AM, JO, SA, EW, DL, and JK curated the data. PB, SMe, SS, AMan, AM, SSu, and JK analysed the data. DL and JK acquired funding. PB, SMe, SS, AMan, ÖY, AMac, JO, LA, SA, MB, AC, DC, TD, GD, MD, RDP, OF, RG, PHL, JL, AMat, CM, SMü, YN, JRO, AO, CP, E-WR, RR, AS, TS, JV, CZ, IK, CG, LK, CG, FP, and JK took part in the investigation. PB, SMe, SS, MPS, DL, and JK created the methodology. PB, SMe, AM, SSu, and JK took part in the project administration. LA, SA, AB, AC, DC, GD, MD, RDP, OF, KH, PHL, CM, SMü, JRO, CP, AS, JV, CZ, IK, CG, HW, KBe, FP, and JK provided resources. PB and SS provided the software. DL and JK gave supervision. PB, SM, AMan, IK, FP, DL, and JK provided validation. PB, AMan, FP, and JK verified the data. PB, SMe, SSc, DL, and JK created the figures visualisation. PB, SMe, DL, and JK wrote the original draft. All authors reviewed and edited the final manuscript. PB and JK had access to raw data and final responsibility for the decision to submit for publication.

NfL Reference Database in the Swiss Multiple Sclerosis Cohort Study Group

Austria—M Khalil (Medical University of Graz, Graz); *Germany*—S Wellmann (University Children's Hospital Regensburg, University of Regensburg, Regensburg); *Switzerland*—S Aeschbacher, M Barakovic, A Buser, M D'Souza, R Galbusera, J Lorscheider, A Orleth, E-W Radue, R Rahmanzadeh, T Sinnecker, S Subramaniam, J Wuerfel (University Hospital Basel, University of Basel, Basel); A Chan (Bern University Hospital, University of Bern, Bern); G Disanto (Neurocenter of Southern Switzerland, Lugano); R Du Pasquier (Lausanne University Hospital and University of Lausanne, Lausanne); O Findling (Cantonal Hospital Aarau, Aarau); A Mathias (Lausanne University Hospital and University of Lausanne, Lausanne); J Vehoff (Cantonal Hospital St Gallen, St Gallen); C Zecca (Università della Svizzera Italiana, Lugano); *USA*—K Hrusovsky (Quanterix Billerica, MA).

Declaration of interests

ÖY 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. JO served on advisory boards for Roche and Merck. LA served on scientific advisory boards for Celgene, Novartis Pharmaceuticals, Merck, Biogen, Sanofi Genzyme, Roche, and Bayer; received funding for travel or speaker honoraria, or both, from Celgene, Biogen, Sanofi Genzyme, Novartis, Merck Serono, Roche, Teva, and the Swiss MS Society; and research support from Biogen, Sanofi Genzyme, and Novartis. AC 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*. DC received speaker fees from BMS and Pfizer and consultation fees from Roche Diagnostics. TD received speaker fees, research support, 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. RDP has received honoraria for advisory boards from Biogen, Celgene, Merck, Novartis, Roche, and Sanofi-Genzyme. KH is an employee and stockholder at Quanterix Corp. MK has received funding for attending meetings or travel from Merck and Biogen, honoraria for lectures or 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. PHL received honoraria for speaking from Biogen Idec, Genzyme, Merck Serono, Novartis, Sanofi Aventis, and Teva; consulting fees from Biogen Idec, GeNeuro, Genzyme, Merck Serono, Novartis, Sanofi-Aventis, and Teva; and research grants from Biogen Idec, Merck Serono, and Novartis. JL has received research support from Innosuisse Innovation Agency, Biogen, and Novartis and served on advisory boards for Roche and Teva. CM has received research support from the Swiss National Science Foundation, the Swiss Heart Foundation, the KTI, and University of Basel; Abbott, Astra Zeneca, Beckman Coulter, Brahms, Idorsia, Novartis, Quidel, ortho clinical Diagnostics, Roche, Siemens, Singulex, Sphingotec, and 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. YN'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. CP received consulting fees or travel compensation, used exclusively for research support, for activities with Biogen, Merck, Novartis, Roche, and Sanofi Genzyme. AS received speaker honoraria or travel compensation for activities with Almirall Hermal GmbH, Biogen, Merck, Novartis, Roche, and Sanofi Genzyme, and research support by the Swiss MS Society. TS has received travel support from Actelion, Alkermes, and Roche. He is a part-time employee of the MIAC AG in Basel. JV has received speaker honoraria from Allmiral Hermal GmbH and Roche. SW is Chief Medical Officer and cofounder of Neopredix. JW 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, and the German Ministry of Economy. CZ received honoraria for speaking/consulting fees or grants from Abbvie, Almirall, Biogen Idec, Celgene, Genzyme, Lilly, Merck Serono, Novartis, Roche, and Teva Pharma. KB has served as a consultant, at advisory boards, or at data monitoring committees for Abcam, Axon, Biogen, and JOMDD/Shimadzu. Julius Clinical, Lilly, MagQu, Novartis, Roche Diagnostics, and Siemens Healthineers, and is a cofounder of Brain Biomarker Solutions in Gothenburg AB (BBS), which is a part of the GU Ventures Incubator Program. HZ 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 cofounder of Brain Biomarker Solutions in Gothenburg AB, which is a part of the GU Ventures Incubator Program, outside of the submitted work. CG received honoraria for speaking/consulting fees or grants from Abbvie, Almirall, Biogen Idec, Celgene, Genzyme, Merck Serono, Novartis, Roche, Teva Pharma. LK'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, Genzyme, Genentech, F Hoffmann-La Roche, Japan Tobacco, Janssen Pharmaceuticals, 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). HW received honoraria and consultation fees from Bayer Healthcare, Biogen, Fresenius Medical Care, GlaxoSmithKline, GW Pharmaceuticals, Merck Serono, Novartis, Sanofi Genzyme, and Teva Pharma. KBe 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 adverse event registry for people with multiple sclerosis. MP has received consulting fees from Biogen, Merck, Teva, Genzyme, Roche, Novartis, GeNeuro, and MedDay. CGr's employer, The University Hospital Basel, received the following fees: advisory board and consultancy fees from Actelion, Novartis, Genzyme and F Hoffmann-La Roche; speaker fees from Biogen and Genzyme-Sanofi; research support by F Hoffmann-La Roche Ltd. She has also received speaker honoraria and travel funding by Novartis. FP has received research grants from Merck KGaA and UCB, and fees for serving on DMC in clinical trials with Chugai, Lundbeck, and Roche. DL is Chief Medical Officer of GeNeuro. JK received speaker fees, research support, travel support, and 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. All other authors declare no competing interests.

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 app for determination of sNfL Z scores is available at https://shiny.dkfbasel.ch/baselnflreference.

Acknowledgments

We thank patients and relatives for their participation and support; study nurses at participating centres for their collaboration and recruitment of participants to the cohorts; and administrative personnel of the SMSC and the Swedish MS registry. We would like to thank Mathieu Canales, Lilian Demuth, and Irmtraut Scheerer for expert technical support. 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. AMan 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.

References

1 Hauser SL, Cree BAC. Treatment of multiple sclerosis: a review. *Am J Med* 2020; **133:** 1380–90.

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:** 2055217318783352.

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:** 152–58.

5 Cree BA, 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:** 272.

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 NeuroInflamm* 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:** 2230–37.

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:** 102807.

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:** 1081–89.

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 Disanto G, Benkert P, Lorscheider J, et al. The Swiss Multiple Sclerosis Cohort-Study (SMSC): a prospective Swiss wide investigation of key phases in disease evolution and new treatment options. *PLoS One* 2016; 11: e0152347.

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

20 Hedström AK, Hillert J, Olsson T, Alfredsson L. Smoking and multiple sclerosis susceptibility. *Eur J Epidemiol* 2013; **28:** 867–74.

21 Holmén C, Piehl F, Hillert J, et al. A Swedish national post-marketing surveillance study of natalizumab treatment in multiple sclerosis. *Mult Scler J* 2011; **17:** 708–19.

22 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:** 230–33.

23 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.

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

25 Curtis AE, Smith TA, Ziganshin BA, Elefteriades JA. The mystery of the Z-score. *Aorta* 2016; **4:** 124–30.

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

27 Jens K. Serum neurofilament light chain reference app. https://shiny.dkfbasel.ch/baselnflreference (accessed Jan 28, 2021).

28 Giovannoni G. Peripheral blood neurofilament light chain levels: the neurologist's C-reactive protein? *Brain* 2018; **141:** 2235–37.

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

30 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 Library, *Neurology* 2019; **92** (suppl 15)**:** S12.008.

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

32 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. ECTRIMS Online Library, 2018. https://onlinelibrary.ectrims-congress.eu/ectrims/2018/ectrims-2018/232039/ludwig.kappos.neurofilament.light.levels.in.the.blood.of.patients.with.html (accessed Jan 5, 2021).

33 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:** e555.

34 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:** 812.

35 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:** 209.

36 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* 2017; **23:** 192–203.

37 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; **27:** 1497–505.

38 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.

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

40 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.

41 Pittock SJ. Uncertain BENEFIT of early interferon beta-1b treatment. *Lancet Neurol* 2009; **8:** 970–71.