Towards development of neurofilament light chain and glial fibrillary acidic protein as precision medicine biomarkers for multiple sclerosis

Inaugural dissertation

to

be awarded the degree of Dr. sc. med.
presented at
the Faculty of Medicine
of the University of Basel

by

Stephanie Victoria Meier

from Zurich, Switzerland

Basel, 2023

Approved by the Faculty of Medicine

On application of

Prof. Dr. med. Dr. phil. Jens Kuhle – primary advisor

Prof. Dr. med. Dr. h. c. mult. Ludwig Kappos – secondary advisor

Dr. med. Dr. phil. Sara Mariotto – external expert

Basel, 26.06.2023
(Date of the acceptance of the Faculty)

Prof. Dr. Primo Schär

Dean

Table of contents

Acknowledgements

Firstly, I would like to thank my PhD supervisor, Prof. Jens Kuhle, very much for his valuable guidance and strong involvement in all projects, from laboratory details to manuscript writing.

I would also like to thank Prof. Ludwig Kappos for his valuable input and the extensive expertise he shared.

A grateful thank you to my advisor, Prof. David Leppert, for his invaluable support and his knowledge and guidance throughout my PhD.

A big thank you to everyone from our group at the University Hospital Basel, especially my lab team, Aleksandra Maleska Maceski, Juan Gomez and Amar Zadic, for their support in the lab and the pleasant working atmosphere; the statistics team at the CTU for the great collaboration and their immense statistical knowledge, particularly Dr. Pascal Benkert, Dr. Eline Willemse and Sabine Schaedelin; the SMSC team, especially Dr. Anette Orleth and Marguerite Limberg for their helpfulness and for giving me an insight into their work; and Dr. Johanna Oechtering for sharing her knowledge.

Further I would like to thank our collaborators at the University Hospital Basel, especially the ThINk team around Prof. Cristina Granziera for their MRI expertise.

I would like to extend my gratitude to all the patients enrolled in the Swiss MS Cohort (SMSC), without whom this work would not be possible.

Thank you to the Prof. Derfuss group for being such good colleagues in the lab and in sharing the lab space.

A special thank you to Hailey Kim for her constant support and motivation, Mika Schneider for his passion and inspiration, Elisabeth Pössnecker for sharing the journey and excitement from the start and Asli Akin for her enthusiasm and reinforcement from across the lab.

Finally, a huge and heartfelt thank you to my family for their endless support, I am extremely grateful for their motivation and encouragement on each step of the way.

List of publications as part of the thesis main text

Asterisks (*) denote equal contributions.

Benkert P*, **Meier S***, Schaedelin S, Yaldizli Ö, Maceski A, Oechtering J, Achtnichts L, Aeschbacher S, Barakovic M, Buser A, Chan A, Conen D, Derfuss T, Disanto G, D'Souza M, Du Pasquier R, Findling O, Galbusera R, Hrusovsky K, Khalil M, Lalive PH, Lorscheider J, Mathias A, Mueller C, Müller S, Naegelin Y, Oksenberg JR, Orleth A, Pot C, Radue EW, Rahmanzadeh R, Salmen A, Sinnecker T, Subramaniam S, Vehoff J, Wellmann S, Wuerfel J, Zecca C, Willemse E, Blennow K, Zetterberg H, Gobbi C, Kappos L, Wiendl H, Berger K, Sormani MP, Granziera C, Leppert D, Kuhle J, for the Swiss Multiple Sclerosis Cohort Study. *Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study.* Lancet Neurol 2022; 21: 246–57.

Meier S*, Willemse EAJ*, Schaedelin S, Oechtering J, Lorscheider J, Melie-Garcia L, Cagol A, Barakovic M, Galbusera R, Subramaniam S, Barro C, Abdelhak A, Thebault S, Achtnichts L, Lalive P, Müller S, Pot C, Salmen A, Disanto G, Zecca C, D'Souza M, Orleth A, Khalil M, Buchmann A, Du Pasquier R, Yaldizli Ö, Derfuss T, Berger K, Hermesdorf M, Wiendl H, Piehl F, Battaglini M, Fischer U, Kappos L, Gobbi C, Granziera C, Bridel C, Leppert D, Maleska Maceski A*, Benkert P*, Kuhle J*. *Serum Glial Fibrillary Acidic Protein Compared With Neurofilament Light Chain as a Biomarker for Disease Progression in Multiple Sclerosis.* JAMA Neurol. 2023;80(3):287–297.

Abbreviations

ALS	Amyotrophic lateral sclerosis
BMI	Body mass index
CDW	Confirmed disability worsening
CRM	Certified Reference Material
CSF	Cerebrospinal fluid
DMT	Disease modifying treatment
EDSS	Expanded Disability Status Scale
eGFR	Estimated glomerular filtration rate
FDA	Food and Drug Administration
GAMLSS	Generalized additive model for location, scale and shape
GFAP	Glial fibrillary acidic protein
GM	Gray matter
HC	Healthy control
IgG	Immunoglobulin G
kDa	Kilodalton
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
NEDA	No Evidence of Disease Activity
NfL	Neurofilament light chain
NMOSD	Neuromyelitis optica spectrum disorder
PIRA	Progression independent of relapse activity
PMS	Progressive MS
RAW	Relapse-associated worsening
RDB	Reference Database
RRMS	Relapsing remitting multiple sclerosis
SIMOA	SIngle MOlecule Array

sGFAP	Serum glial fibrillary acidic protein
SMSC	Swiss Multiple Sclerosis Cohort
sNfL	Serum neurofilament light chain
SPMS	Secondary progressive MS
TBI	Traumatic brain injury
T2w	T2 weighted
WM	White matter

Abstract

Background: We have insufficient diagnostic tools to capture and anticipate the course of multiple sclerosis (MS) and to monitor treatment response. Blood-based biomarkers could provide a valuable measure to detect neurodegeneration and disease worsening in MS. Serum neurofilament light chain (sNfL) is a biomarker of neuro-axonal injury that has been investigated in its association with disease activity and disability accumulation in MS, but larger scale studies to determine the sNfL levels of healthy persons and MS patients are currently lacking. Furthermore, we lack biomarkers to discern the pathogenesis of 'pure progression' in MS from that due to focal inflammatory activity. Serum glial fibrillary acidic protein (sGFAP) is a marker for astrogliosis and a potential candidate biomarker that may be more strongly associated with disease progression than active inflammation in MS.

Objectives: We aimed to bring sNfL closer to clinical application by establishing a reference database of sNfL levels from control persons, in order to enable the determination of pathological sNfL levels by calculation of sNfL percentiles and Z scores of MS patients. Further, we used this reference database to analyze sNfL's ability to capture and prognosticate disease activity in patients followed in the Swiss MS Cohort (SMSC) and the Swedish MS Registry and explored the effectiveness of disease modifying therapies. Further, we assessed the value of sGFAP in addition to sNfL as a biomarker for disease progression and acute inflammation, as well as in patients under B-cell depleting therapy.

Methods: We used the Single Molecule Array (SIMOA) technology (Quanterix) for the measurements of sNfL and sGFAP. For the sNfL reference database, persons with no evidence of CNS disease were included from four cohort studies in Europe and North America. A generalized additive model for location, scale and shape (GAMLSS) was used to model the distribution of sNfL concentrations in function of age and body mass index (BMI). We tested the reference database by generating sNfL percentiles and Z scores in the SMSC, and as a validation in the Swedish MS Registry. In the second study, we measured sNfL and sGFAP in three different groups of patients in the SMSC: firstly, matched patients with MS who had either stable disease or disability progression with no relapses during the entire follow-up; secondly, patients with MRI or clinical signs of acute neuroinflammation or in remission; thirdly, patients who had initiated and continued B-cell–depleting treatment (ocrelizumab or rituximab).

Results: In the first study we measured sNfL concentrations in 10’133 serum samples from 5’390 control persons and found an age- and BMI-related sNfL increase. We also measured 7’769 serum samples from 1’313 MS patients from the SMSC. sNfL Z scores prognosticated an increased risk for future disease activity and normalized in patients under treatment with monoclonal antibodies compared to other treatments or untreated patients. These results were validated in 4’341 samples from the Swedish MS Registry.

In the second study we measured sNfL and sGFAP in 355 patients and 259 healthy controls. sGFAP concentrations in the controls increased with age and BMI and were higher in women than men. Patients with worsening progressive MS had higher levels of sGFAP than stable patients even after adjustment for sNfL. Furthermore, baseline sGFAP was associated with gray matter volume loss, but not white matter volume loss, and remained unchanged during relapses compared to remission phases. Additionally, the combination of sGFAP and sNfL Z scores could prognosticate future disability worsening and 'progression independent of relapse activity' (PIRA).

Conclusion: Our reference database and the therein derivable sNfL percentiles or Z scores enable the identification of individual persons with MS at risk for future disease worsening and treatment response also in otherwise seemingly stable disease stage. Furthermore, sGFAP may be a sensitive tool to capture and prognosticate future PIRA, especially in combination with sNfL.

Chapter 1: Introduction

1.1 Multiple Sclerosis Disease Patterns

Multiple Sclerosis (MS) is an inflammatory demyelinating and neurodegenerative disease of the central nervous system and one of the major causes of disability in young adults.[1,2] MS is characterized by a highly heterogeneous disease pattern, with two main types of disease courses: firstly, disease worsening due to acute neuroinflammation with new MRI lesions and/or clinical relapses, and secondly, due to insidious disease progression with increasing disability levels over time.[3,4] Acute disease activity appears to be driven by lymphocyte invasion into the CNS causing lesion formation as seen in MRI, and its clinical correlate acute attacks, which can lead to permanent functional deficits called 'relapse-associated worsening' (RAW).[4]

Progression instead, was defined as continuous and increasing neurologic impairment over time.[5,6] Although the majority of patients (85%) experience the onset of MS with relapsing-remitting disease (RRMS),[5,7] a study of RRMS patients showed that 24.2% of these cases transform into secondary progressive disease course (SPMS) 20 years after onset.[8] Especially in the later stages of disease the effect of relapses on disease worsening is strongly reduced,[9,10] hence at this point disease progression may play a larger role in accelerating disability worsening. In pure progression, it is assumed that brain-diffuse neurodegeneration resulting from inflammatory activity by brain-resident cells as an innate immune reaction leads to a smoldering loss of neurological functions.[11] This subclinical chronic inflammation, or so-called 'smoldering MS'[11] has also been defined as 'progression independent of relapse activity' (PIRA).[3] PIRA challenges the current practices of MS diagnosis and treatment monitoring, as this disease worsening is relapse-free and is difficult to measure in MRI or clinical assessment.[11] Instead, PIRA leads to confirmed disability progression (CDP) in terms of EDSS score progression, despite the absence of relapses.[3,12] Accordingly, alternative ways of measuring the underlying processes of disease worsening are required, especially when pure progression is involved. As the EDSS score increase is accelerated through earlier onset of PIRA in MS patients,[12] the urgency of detecting PIRA at an early and any stage of the disease becomes evident. The strong impact of PIRA on disability progression in MS patients shows the high importance of a better understanding of the factors associated with MS disease progression and finding superior ways to measure and eventually treat this disease course.

1.2 Disease Modifying Treatments in MS

The increased knowledge about the MS disease course and pathology in recent years has led to the development of a wide range of disease modifying treatments (DMT's).

Despite the almost complete suppression of acute disease activity with these 'high-efficacy therapies' (monoclonal antibodies targeting CD20, CD52 or VLA- 4), these therapies have little impact on progression.[8,9] Consequently, this chronic deterioration of neurologic functions is the largest unmet medical need in MS, both therapeutically and diagnostically.

The mechanism of current DMT's is the reduction of neuroinflammation by either depletion of lymphocytes or interference with their course of action, hence they are mainly active on acute inflammatory stages of disease, while the pathomechanims leading to disease progression are largely outside their pharmacological reach.[13] This also explains why, in a long-term cohort of MS patients followed for 10 years, 59% of patients experienced significant increase in disability despite being under treatment and undergoing treatment escalation.[8]

The current diagnosis and monitoring of treatment of MS are based on clinical criteria, analysis of cerebrospinal fluid (CSF) and MRI,[1] as described in the McDonald criteria.[6,14] CSF analysis is used to determine the intrathecal synthesis of immunoglobulin G (IgG) leading to the presence of increased IgG index and oligoclonal IgG bands,[1] the latter can be found in the CSF in 95% of MS patients.[15] The disadvantage of these measures is their lack of specificity, as they also occur in other inflammatory diseases than MS; and that they require lumbar puncture and hence cannot be measured longitudinally in routine clinical practice.[15]

MRI has become the gold-standard of paraclinical measures for diagnosis and treatment monitoring in MS, however, it is mainly a retrospective measure of neural inflammation and atrophy and its elaborate procedures impede frequent measurements.[16] Consequently, new methods to measure MS disease progression are required.

1.3 Blood-based Biomarkers

To fill in the gaps in the monitoring of MS, blood-based biomarkers may provide a minimally invasive alternative to measure real-time neuronal damage across the entire CNS. These are

aimed not to replace, but to complement the existing measures, in order to enhance and facilitate the regular examinations in a personalized medicine approach. Two potential candidate biomarkers will be discussed in more detail: Serum neurofilament light chain (sNfL) and serum glial acidic fibrillary protein (sGFAP).

1.3.1 Neurofilament Light Chain

Neurofilaments are structural proteins of neurons that have an important role in maintaining the neuronal shape.[17] These proteins take the form of heavy chain (190-210 kilodalton (kDa)), intermediate chain (150 kDa) and light chain neurofilaments (68 kDa), of which the light chain is the most abundant (NfL).[18] As neurofilaments are released during neuronal injury into the CSF and the blood, and are exclusively found in neurons, they qualify as specific markers for neuro-axonal injury in the central and peripheral nervous system.[18]

NfL has first been investigated in CSF, where levels were elevated in MS patients compared to healthy controls[19] and concentration increases occurred parallel to the onset and progression of brain lesions.[20,21] The emergence of novel assay platforms for the highly sensitive detection of proteins, such as the SIMOA (single molecule array) technology, enabled the measurement of NfL in serum samples.[22] NfL levels are 30-70x lower in serum than CSF,[18] but there is a high correlation between CSF and serum or plasma NfL levels, hence previous results in CSF of MS patients were highly congruent to those from serum.[23–27] This led to the rise of investigations on NfL as a biomarker in MS and numerous neurodegenerative diseases, where elevated levels of NfL compared to controls could be found in several diseases, such as amyotrophic lateral sclerosis (ALS)[28,29] and Guillain-Barré syndrome.[29] Furthermore, NfL correlated with brain atrophy rate and time to disease onset in Alzheimer's patients[30] and increased according to the severity of injury in patients with traumatic brain injury (TBI),[31] such as American football athletes,[32] boxers[33] and patients with spinal cord injury,[34] as well as ALS patients.[16,35] Despite the different processes of neuronal damage, the results in terms of sNfL fluctuations depending on disease severity can be seen in many of these diseases. In MS, sNfL has been widely investigated due to the need for additional diagnostic and monitoring methods. Previous work in our group has shown that sNfL levels can be used as a blood biomarker to predict MS disease worsening and to monitor treatment effects.[24,25,36,37]

As a highly sensitive real-time marker of neuro-axonal injury, sNfL levels can indicate presymptomatic stages of MS. In a study in the US army, persons who developed MS had higher sNfL levels as much as 6 years before disease onset compared to healthy controls.[38] Similarly, in a longitudinal study with RRMS patients, baseline sNfL was predictive of 4-year brain atrophy and the development of new T2 weighted (w) lesions.[39] This predictive value of sNfL for T2 lesion volume and brain parenchymal fraction could also be shown in a longitudinal study of MS patients followed over 10 years.[40] Subsequently the association of sNfL with T2w lesion volume was confirmed in several studies.[24,25,41–45] MS patients with increasing brain volume loss also had high baseline NfL,[24,40,46,47] which was also the case for a number of Gadolinium-enhancing lesions.[25,27,48]

These studies show the potential of sNfL as a biomarker for MS, as it fulfills the biological and technical criteria required for a biomarker to have potential as a clinical tool:[49] It is increased in MS patients compared to healthy controls and correlates with disease severity; As a blood-based biomarker it is relatively easy accessible and detectable with modern methods (SIMOA); Further, NfL in serum is not sensitive to repeated freeze/thawing cycles and can be stored using standard serum sample handling procedures.[50,51] Consequently, sNfL is not only in theory a promising biomarker for MS, previous studies have also shown the potential of sNfL as a prognostic biomarker and as a disease activity and treatment response biomarker in different MS patient cohorts. However, despite the large number of studies about sNfL in MS and neurodegenerative diseases, to date sNfL has not been approved as a biomarker for any disease. This could partly be due to the fact that these promising preliminary findings are mainly valid on a group level and potentially as an endpoint for clinical trials. In order to advance sNfL as a clinical biomarker for MS, these findings must be validated in a larger scale study, otherwise the reproducibility of previous study results cannot be guaranteed on an individual patient level.

An important issue in translating the results from the group level to the individual patient level is the impact of confounding factors on sNfL concentrations. Studies have shown that age and body mass index (BMI) have an effect on sNfL levels of healthy persons.[24,52,53] Therefore, applying fixed cut-offs ignoring the influence of age and BMI on measured concentrations is suboptimal, since the range of normal values differ between individuals depending on their age, BMI and potentially other factors. This would clearly hamper the clinical application of sNfL as a biomarker of disease activity in individual patients. A large reference database with sNfL levels of the healthy population is required to determine normal values of sNfL across age groups, before pathological levels in individual MS patients can be fully assessed.

The last decades have seen the development of a variety of new DMT's for MS patients. In order to better understand the efficacy of those therapies, stringent monitoring of disease activity in treated patients is required. Clinical trials have used sNfL as an endpoint to determine the effectiveness of the therapy in addition to clinical and MRI markers, noting a reduction in sNfL levels following therapy.[54,55] Further studies could show that sNfL levels were reduced in patients under DMT compared to the untreated patients.[25,42,48,56] However, even after treatment initiation sNfL levels remained elevated in progressive MS (PMS) compared RRMS patients and controls,[24,57,58] supporting the previous findings of a reduced treatment efficacy in PMS patients.[8,9]

1.3.2 Glial Fibrillary Acidic Protein

The second investigated biomarker is glial fibrillary acidic protein (GFAP), an intermediate filament of astrocytes with a fiber diameter of 8-12nm that is released into the CSF and blood following astrocytic damage or activation.[59,60] Different from NfL, it is assumed that GFAP levels can be increased both in phases of acute astrocyte damage and as a reflection of astrogliosis.[61–63] Therefore, GFAP has been explored as a biomarker for a variety of neurological conditions such as TBI, MS, neuromyelitis optica spectrum disorder (NMOSD) and neurodegenerative dementias, such as Alzheimer's disease and frontotemporal dementia.[64–67] In patients with NMOSD, GFAP levels were increased within one week of an NMOSD attack and the amount of increase correlated with the attack severity.[64,65] In neurodegenerative dementias, in particular frontotemporal dementia, GFAP was increased in comparison with controls[67] and correlated with age, NfL and brain volume.[66] The literature shows that although it is relatively new to the MS biomarker field, sGFAP has been investigated in depth in other neurodegenerative diseases, and has also already been authorized by the FDA in the form of a blood test to measure mild TBI.[68,69] Clearly the research on sGFAP is already advanced in other diseases, allowing its application as a biomarker for astrocytic damage and showing its potential value as a biomarker for disability progression in MS.

Early studies investigating GFAP in CSF of MS patients have found a correlation of GFAP levels with higher disability levels as measured by EDSS score and clinical parameters.[61,70] Interestingly, CSF GFAP was not affected by the acute phases of relapses or lesional activity, while NfL was a sensitive indicator of acute disease activity.[21,71,72] On the contrary, CSF GFAP was even increased in progressive patients in comparison to relapsing patients.[19,70] This is an

interesting finding insofar that the association of NfL with focal inflammation has already been found in previous studies,[24,25,42,73] while the contrary effect of GFAP gives rise to new possibilities in combining biomarkers that reflect different disease states in MS.

For GFAP, likewise as NfL, the new, more sensitive measuring devices have rather recently allowed the quantitation of GFAP in serum, showing a similar correlation with measures of disability as has been described in CSF.[63,74] Apart from clinical measures of disease, sGFAP also showed correlation with MRI measures in MS patients. Accordingly, high sGFAP was associated with higher T1w hypointense and T2w hyperintense lesion load as well as with gray and white matter atrophy.[68,74,75] These results suggest a promising link of sGFAP with disability worsening in MS patients. Additional to the association of sGFAP with clinical and MRI characteristics, studies have also found moderately high correlation between sNfL and sGFAP levels in MS patients (rho=0.53, $p<0.001$, n=79;[74] rho=0.4, $p<0.001$, n=80;[63] rho=0.66, $p<0.001$, n=129[75]) bearing potentially added value next to sNfL measurements. Since sGFAP has been investigated intensively in TBI, its half-life is already known and estimated at 24-72h after injury.[68,76] This is contrary to sNfL, of which an official half-life could to date not be defined.[76] There is only one study in patients with TBI that noted a return to normal sNfL levels around 3 months after injury.[33] This knowledge may be helpful in understanding the metabolism of sGFAP and sNfL, which in turn would be useful in determining the state of neurodegeneration and the required frequency of measurement of sNfL and sGFAP as biomarkers. Furthermore, considering the variation of sNfL concentrations based on confounding factors such as age or renal function, the same may be the case for sGFAP. Indeed, an association of sGFAP with age has been found by some studies in MS patients[74,75] and controls[63]. This needs to be pursued further in healthy controls to ensure a clear understanding of the influencing factors on sGFAP concentrations to avoid any false conclusions on MS pathologies and disease worsening.

Despite these strong advances in blood biomarker research, to date no blood-based biomarkers are clinically used in diagnostics and treatment monitoring of MS on a routine basis. Evidently, there is large potential for sNfL to advance as an MS biomarker, with only a few major pieces of information missing to reach the next step in its development as a tool for application in clinical practice. Furthermore, sGFAP is an emerging biomarker for astrocytic damage and, according to preliminary studies, a potential candidate for measuring disease worsening in MS. However, both of these biomarkers require further in-depth investigations in specialized and well-characterized MS cohorts in order to advance in their development as biomarkers for MS.

Chapter 2: Research Objectives

This project includes three main sub-projects, of which the first two were focused on bringing sNfL closer to clinical application, while the third focused on the potential of combining sNfL with another biomarker, sGFAP, to explore its value as a biomarker of disease progression in MS

1.) The first objective was to derive percentiles and Z scores for sNfL from a large reference database from control persons, to define levels of pathological increase of sNfL independent of BMI and age, in the most efficient, sensitive and specific way. Our objective was to test, in two large and independent cohorts of people with MS, whether sNfL Z scores would predict the risk for future disease activity also in patients with 'no evidence of disease activity-3' NEDA-3. This can be found in chapter 3.1 in the first publication.

2.) The second objective was to investigate whether the sNfL percentiles and Z scores could be used to compare effects of disease-modifying therapies on longitudinal sNfL levels. This work can also be found in chapter 3.1 as a second part of publication one.

3.) The third objective was to directly compare sGFAP and sNfL levels: how they reflect acute disease activity vs how they identify and prognosticate future disease progression and whether their combination provides added value. In cohort 1 (SMSC patients with either worsening progressive/stable MS and relapsing MS), we (1) measured sNfL and sGFAP levels in patients who either remained clinically stable or continued to accumulate more disability over time and (2) compared how they are impacted by acute inflammation in a cohort of patients with relapsing forms of MS. In cohort 2 (SMSC patients under B-cell depleting therapy (BCDT)), we evaluated how sNfL and sGFAP levels, alone and in combination, are prognostic for future disability worsening and PIRA in patients with MS receiving BCDT as a model of optimal suppression of acute disease activity. This can be found in chapter 3.2, the second publication.

Chapter 3: Publications

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

Note: This publication was awarded with the Viollier Prize 2022 and the prize of the Mogens und Wilhelm Ellermann-Stiftung 2022.

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.

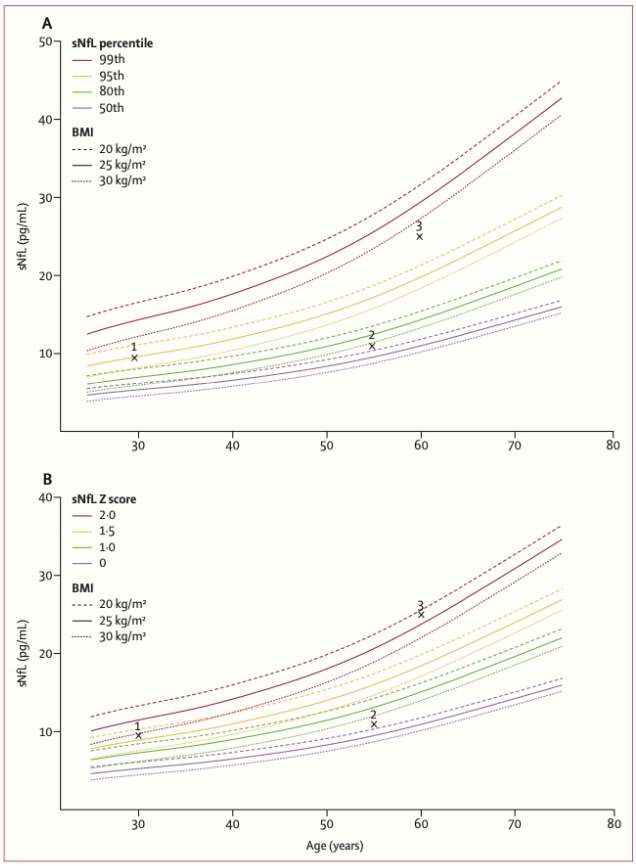


Figure 1: **sNfL percentiles (A) and Z scores (B) reference curves**
A generalised additive model for location, scale, and shape was used to model the association of sNfL concentration (pg/mL) in controls with data for BMI and age. Example 1, at 30 years and a BMI of 25 kg/m², shows sNfL of 9·5 pg/mL (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
10133 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

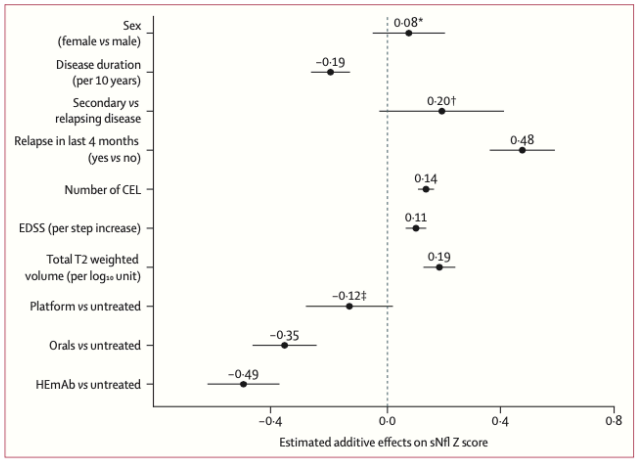


Figure 2: **Factors affecting sNfL Z scores in people with multiple sclerosis**
Model estimates including 95% CIs (see appendix [p 12] for numerical values). All values are p<0·0001, unless specified in the footnotes. CEL=contrast-enhancing T1-weighted lesions. EDSS=Expanded Disability Status Scale score.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

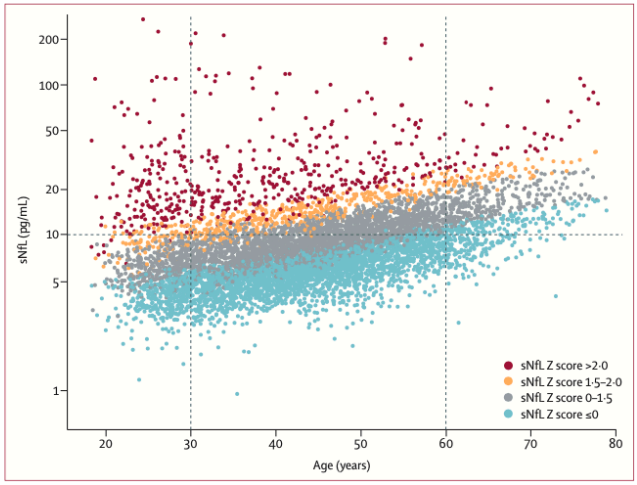


Figure 3: **sNfL Z scores according to age of people with multiple sclerosis participating in the Swiss Multiple Sclerosis Cohort**
Age-adjusted and BMI-adjusted sNfL Z scores are shown by colour gradient. The fixed sNfL cutoff is shown by the horizontal line at 10 pg/mL. Using a fixed cutoff in people 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.

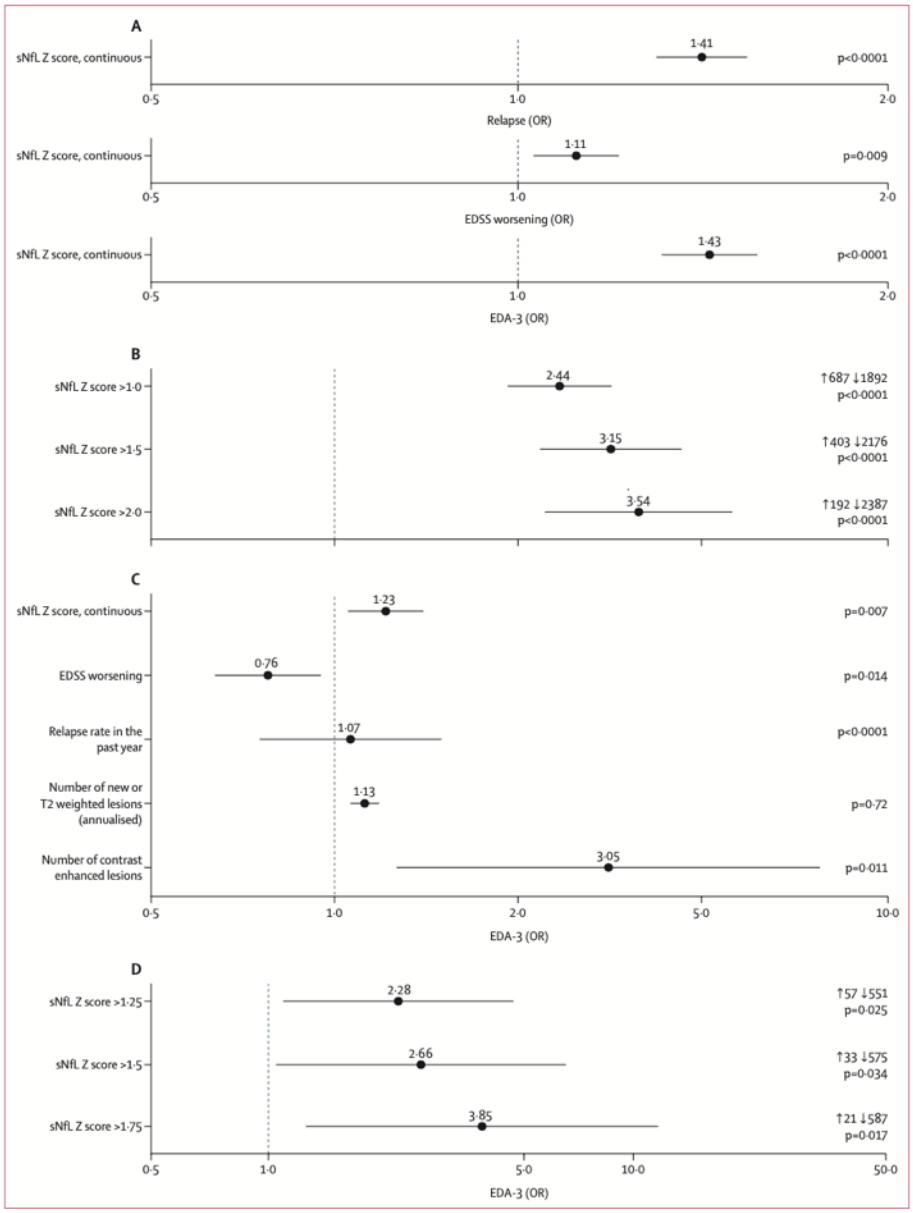


Figure 4: **sNfL Z scores predicting disease activity in the following year**
Estimates (ORs) and 95% CIs are shown. A, B, and D show three univariable models and C shows a multivariable model. Arrows display number of serum samples above or below the respective sNfL Z score cutoff. Probability of occurrence 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.

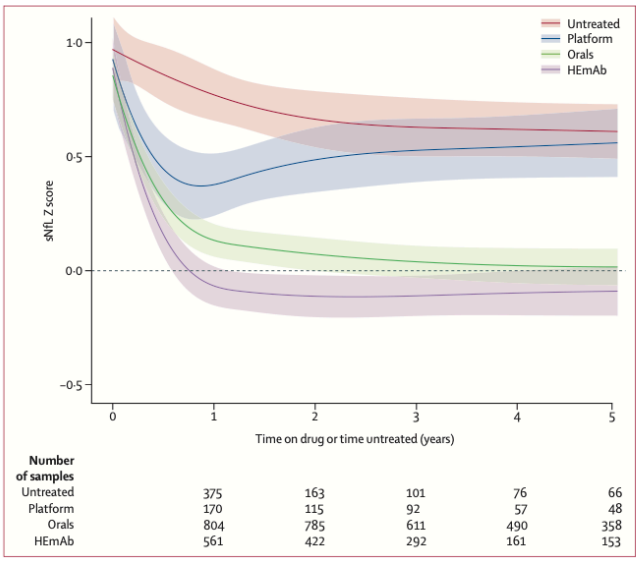


Figure 5: **Temporal evolution of sNfL Z scores under treatment**
Four treatment categories were included in a mixed-effects model, thereby using spline terms to model the non-linear temporal association and an interaction term between disease-modifying therapy category and treatment duration. The number 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.

THE LANCET
Neurology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Benkert P, Meier S, Schaedelin S, et al. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. *Lancet Neurol* 2022; **21:** 246–57.

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

Pascal Benkert*, Stephanie Meier*, et al., for the Swiss Multiple Sclerosis Cohort Study

Online only material

Index: Page

Supplementary Methods

1. Swiss Multiple Sclerosis Cohort (SMSC)

The SMSC (NCT02433028) is a prospective multicentre cohort study performed across eight Swiss academic medical centres: The Cantonal Hospital of Aarau, the University Hospitals of Basel, Berne, Geneva and Lausanne, the Regional Hospital of Lugano and the Cantonal Hospital of St. Gallen. Demographic, neuroimaging, and clinical data as well as serum samples are collected every 6 or 12 months. Standardized clinical assessments with functional system score and Expanded Disability Status Scale (EDSS) score calculations were performed by certified raters (http://www.neurostatus.net/).[1,2]

Relapses were defined as new, worsening or recurrent neurologic symptoms that lasted for at least 24 hours without fever, infection, or adverse reaction to a prescribed medication and that were preceded by a stable or improving neurologic status of at least 30 days. Recent relapses were defined as events with onset within 4 months before serum sampling. EDSS worsening (in the following year) was defined as an increase in EDSS from the current to the next visit of ≥1·5 points from an EDSS score of 0·0, ≥1·0 point from an EDSS score of 1·0–5·5 or ≥0·5 point from an EDSS score ≥6·0. NEDA-3 was defined as having had no EDSS worsening, no relapses, no new/enlarging T2 weighted and no T1 weighted contrast enhancing lesions in the last year, as opposed to patients with EDA-3 who fulfilled at least one of these criteria in the last year. Self-reported weight and height measures were collected at each visit and body mass index (BMI) calculated as weight in kg divided by squared height.

All samples are collected within 8 days from the clinical visit and stored at -80°C following standardised procedures.[2,3] 81 samples (0·9%) under immunosuppressive therapy (n=52: n=36 under mitoxantrone, n=16 azathioprine), randomised controlled phase 3 trial (n=25) or after hematopoietic stem cell transplantation (n=4) were excluded from the analysis (including 17 at baseline), but all patients remained in the study since additional samples were available from respective patients.

Treatment epochs were defined from the day of the first administration of the DMT until its discontinuation. As the pharmacodynamic effect of DMTs may last beyond their wash-out period, treatment effect durations after stopping administration were estimated for the DMTs and added to the time of administration if no other DMT was started (platform and oral: 2 months, except teriflunomide: 0·5 years; HEmAb: 1·5 years for rituximab and ocrelizumab, 5 years for alemtuzumab and 2 months for natalizumab).

2. Validation cohort: Epidemiological Investigation of Multiple Sclerosis (EIMS), Immunomodulation and Multiple Sclerosis Epidemiology (IMSE) and Comparison Between All immuno-Therapies for Multiple Sclerosis (COMBAT-MS)

We validated our findings in 4341 MS cases participating in the EIMS[4], IMSE[5] and/or Combat-MS cohorts.[6] In EIMS, individuals with newly diagnosed MS were identified at neurology clinics throughout Sweden and invited to participate by completing a questionnaire and donating a blood sample. All patients have been examined by a neurologist at the clinic where they were recruited. The IMSE cohorts are part of a nationwide phase 4 surveillance study aimed at investigating the long-term safety and efficacy of all more recent DMTs starting from natalizumab. The Combat-MS is an observational drug trial (clinicaltrials.gov, NCT03193866) in a contemporary relapsing MS cohort entailing a structured follow-up routine and validation of registered data compared to clinical routine.[6] Combined, these cohorts contribute additional data beyond what are collected in the Swedish MS registry (SMSreg), including ethylenediaminetetraacetic acid (EDTA)-treated plasma samples.[5–7] In EIMS, EDTA plasma samples were collected within 5 years of onset of MS and in the IMSE cohort EDTA plasma samples were collected at baseline before initiation of the DMTs (within a month prior to start of DMT), where patients were either treatment naive or were exposed to only interferons and/or glatiramer acetate and at follow-up (treatment duration > 4 months). In Combat-MS patients contribute a yearly plasma and serum sample from date of inclusion in the study. 3022 patients provided one, 1122 two, 141 three, 48 four and 8 patients five samples (overall number of samples: 5921). Baseline demographic and clinical characteristics are shown in **Suppl Table 1**.

In the SMSreg, data are recorded by neurologists or MS nurses through a web interface and include patient characteristics, MS course, DMT exposure, visits, clinical scales (including EDSS), relapses, MRI and laboratory tests. Most data are collected at routine clinical visits on annual or biannual basis.[6] Relapses, recent relapses, EDSS worsening, NEDA-3, EDA-3 and treatment epochs were defined as outlined above.

3. Origin and characteristics of the four cohorts of control persons included in the reference database (RDB)

a) Genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors study (GAPP)[8,9]

GAPP is a population-based prospective cohort study involving a representative sample of healthy adults in the Principality of Lichtenstein. Exclusion criteria were any cardiovascular disease, diabetes, obstructive sleep apnoea syndrome, daily intake of nonsteroidal anti-inflammatory drugs, a body mass index >35 kg/m² or any

other major illness. All inhabitants of the Principality of Liechtenstein aged 25–41 years were invited to participate in the study. The institutional review board of the University Hospital Zürich approved the study, and written informed consent was obtained from all participants.

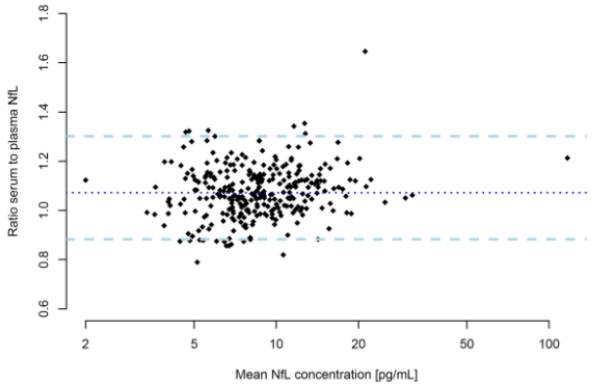
b) Multiple Sclerosis Expression, Proteomics, Imaging, Clinical Study (EPIC)[10] and Genetic MS Associations (Gene MSA)[2,11]
Healthy adults were ascertained through a prospective multicentre effort initiated in 2003. Three MS clinical centres were involved in patient enrolment and biological specimen collection using identical inclusion criteria, two in Europe (Vrije Universiteit Medical Center, Amsterdam; and University Hospital Basel) and one in the USA (University of California San Francisco). Serum samples from San Francisco and Basel were available for this study. The control group consisted of unrelated individuals, primarily spouses/partners, friends and other volunteers. A familial history or current diagnosis of MS as well as a relation to another case or control subject or other reported ongoing major illnesses were considered exclusionary for this group. The institutional review boards of University of California San Francisco and University Hospital Basel, respectively, approved the study, and written informed consent was obtained from all participants.

c) Establishing the links between subclinical arteriosclerosis and depression (BiDirect)[12–14]
The BiDirect study investigates the mutual relationship between depression and (subclinical) arteriosclerosis. It is a prospective observational study that integrates three different cohorts. Only two of the three cohorts were integrated in the generation of the RDB: Cohort 1 consisted of 899 patients, who suffered from an episode of depression at the time of recruitment. Recruitment took place at six different psychiatric and psychosomatic hospitals and departments located in and around the city of Münster, as well as two resident psychiatrists' practices located in Münster. The recruitment of outpatients was limited to those who had been hospitalized due to depression at least once during the 12 months period prior to inclusion into the study. Inclusion criteria were (i) age (≥35 and <66 years) and (ii) current in- or outpatient treatment due to acute depression. Exclusion criteria were (i) compulsory admission, (ii) comorbid dementia, and (iii) comorbid drug abuse (including alcohol). Cohort 2 included 813 community dwelling adults (age: ≥35 and <66 years). These participants had been randomly sampled from the population register of the city of Münster and were invited for BiDirect-Baseline via letter. Individuals in cohort 1 were slightly younger compared with cohort 2 (median (IQR): 52·0 (46·3-57·9) vs 56·6 (49·2-62·4) years) and sNfL levels were numerically slightly lower (8·7 (6·6-11·3) pg/ml vs 9·3 (7·1-12·3) pg/ml; p=0·45, after age correction). The joint ethics committee of the University of Münster and the Westphalian Chamber of Physicians approved the BiDirect-Study, and written informed consent was obtained from all participants.

Self-reported weight and height measures were collected at each visit and BMI calculated (coverage: GAPP: 100%; EPIC: 33·8%; Gene MSA: 100%; BiDirect: 100%). Information on an existing diagnosis of Diabetes mellitus was known in all SMSC patients and control persons and on estimated glomerular filtration rate (eGFR, as a measure of renal function, was calculated according the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine equation[15]) in 52% of SMSC patients and all participants from BiDirect and GAPP.
All serum samples were stored at -80C. Serum samples from EPIC had undergone one additional thaw cycle. sNfL has been shown to not be influenced by thawing.[16,17] We did not see an effect of storage time on sNfL concentrations.

4. Plasma Neurofilament light chain measurements
In EIMS, IMSE and Combat-MS, NfL was measured in duplicate in ethylenediaminetetraacetic acid (EDTA)-treated samples (pNfL) by NF-light® assay at the University Hospital Basel according to the manufacturer's instructions (Quanterix, Billerica, USA)[18,19]. Intra- and inter-assay coefficients of variation were ≤ 10%.
Two hundred ninety-nine paired serum and EDTA plasma samples from the first follow-up examination of BiDirect (**Suppl Table 2**) were selected to investigate the association between NfL concentrations in serum and EDTA plasma. The samples were selected in a way to represent a uniform distribution of age and sex.[20] NfL concentrations between both matrices were highly correlated (Pearson's r: 0·991, p<0·0001); concentrations were systematically higher in serum (**Suppl Figure 1**). The resulting formula for conversion from pNfL to sNfL was: serum NfL [pg/mL] = -0·33 + 1·11 × pNfL [pg/mL].[20]

Suppl Figure 1. Bland-Altman plot showing the relationship between serum and plasma concentrations of NfL



Legend: Ratio between NfL concentration in 299 paired serum and plasma samples plotted against the mean NfL concentration of the pairs. The y-axis represents the ratio of serum to plasma NfL. The blue dotted line represents the mean ratio difference (1·07); the light blue dashed lines represent the 95% limits of agreement (0·88-1·30).

5. MRI assessment methods

Brain MRI scans were performed annually in the SMSC. A standardised imaging protocol was applied across centres including a 3D Magnetization Prepared - RApid Gradient Echo (MPRAGE), a 3D Fluid Attenuated Inversion Recovery (FLAIR) sequence, and a post contrast T1 sequence acquired at a spatial resolution of $1mm^3$. T2w lesion volume was automatically assessed annually by using a deep-learning based approach[21] and a longitudinal evaluation method[22] respectively, followed by manual quality assessment and correction. The number of gadolinium-enhancing lesions was assessed manually. The MRI protocol in the validation cohort has been published previously.[23]

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

6.1. Generation of the final RDB and investigation of relevant comorbidities

Current knowledge indicates an increase of sNfL by comorbid renal insufficiency,[24,25] diabetes mellitus[26] and a decrease of sNfL concentrations in individuals with higher BMI.[26,27]

65 control persons showed at one or several time points an eGFR <60 mL/min/1.73 m^2 contributing 108 serum samples. In addition, 177 control persons with a diagnosis of diabetes mellitus contributed 439 serum samples.

In a mixed effects linear model with log(NfL) as dependent variable (coefficient of determination of the overall model: R^2=0·417; n=8379 serum samples with complete information), the following factors were independently associated with sNfL:

- age (per year; estimate: 1·025; 95% CI: 1·025-1·026; p<0·0001; i.e., 2·5% increase in sNfL concentration per year of age; R^2 remaining: 0·031, when excluding this variable)
- BMI (per 1 kg/m^2; 0·980; 0·978-0·982; p<0·0001; remaining R^2: 0·374),
- eGFR < 60 mL/min/1·73 m^2 (1·215; 1·139-1·295; p<0·0001; remaining R^2: 0·415),
- diagnosis of diabetes mellitus (1·119; 1·064-1·177; p<0·0001; remaining R^2: 0·413),
- but not sex (1·008; 0·988-1·028; p=0·43; remaining R^2: 0·417).

Based on these analyses (**Suppl Figure 2**), BMI and age were integrated in the RDB (and accounted for in the statistical model), whereas diabetes mellitus was not (for marginal contribution). Although eGFR explains overall only a very limited portion of the variability of sNfL in the general population as expressed by the coefficient of determination (due to the low prevalence), the change in sNfL Z-score associated with very low eGFR values is considerable (**Suppl Figure 2E**): As can be seen in **Suppl Figure 2C**, eGFR values ≥60ml/min/1·73 m^2 shows limited association with sNfL Z-score, while eGFR <60ml/min/1·73m^2 coincides with a significant increase. We therefore excluded the 108 samples with an eGFR <60 mL/min/1·73 m^2 from the final RDB.

6.2. Establishing the GAMLSS model

The relationship between sNfL, BMI and age in control persons was modelled using a generalized additive model for location, scale and shape (GAMLSS) based on a Box-Cox t distribution.[28,29] GAMLSS is an extension of generalized linear models (GLMs) and additive linear models (GAMs) suitable for large complex datasets. This method is appropriate for the right-skewed and heavy-tailed distribution of sNfL. It enables the estimation of the first four moments (mean, variance, skewness and kurtosis) as linear or smooth functions of exploratory variables, allowing an accurate estimation also of extreme percentiles and Z-scores. Several distributions were tested comparing goodness-of-fit statistics to find the most appropriate fit for the data. sNfL was the dependent variable and spline terms with 3 degrees of freedom for both explanatory variables (age, BMI) were used to model the non-linear association and no interaction term was used since the Akaike information criterion (AIC) was not considerably better than the more parsimonious model. The optimal number of degrees of freedom of the splines was chosen based on the model's AIC. The model fit was assessed using detrended QQ-plots.
To avoid repeated measures in a reference database, only one sample was selected from each control person by a heuristic approach optimizing towards an even sample distribution over the entire age range (n: 4532) (**Suppl Table 2**). **Suppl Figure 3** shows the number of samples per age group in participants of the reference database and of patients in the SMSC.

From this final GAMLSS model age- and BMI-specific reference values for various percentiles/Z-scores cut-offs were generated and made available as lookup tables (**Suppl Table 3**), reference curves (**Figure 1**) and as an internet-based App.

6.3. Derivation of percentiles and Z-scores from the GAMLSS model

Suppl Figure 4 shows that absolute sNfL levels in this RDB cohort almost double in persons over 60 years compared to those of 30-40 year; at the same time, the standard deviation increases in higher age groups with a considerable number of samples with higher sNfL values. The age- and BMI-adjusted Z-scores derived from the GAMLSS model represent the number of standard deviations the adjusted sNfL levels differ from respective values in the entire control population. Besides being an age- and BMI-corrected measure, Z-scores have the advantage over absolute values to be normally distributed, a mathematical property that is advantageous for statistical modelling. As an alternative measure of the deviation from "normal", percentiles are provided.[30] 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 as a given value or lower. The two measures are interchangeable (a Z-score of 1 and 1·5 represents the 84·1 and 93·3 percentile, respectively, according to the standard normal distribution). However, Z-scores are not bound between 0-100 enabling different applications.
Suppl Table 3 shows the comparison of absolute sNfL values in function of BMI and age and the respective Z-score/percentile levels.

6.4. Validating the GAMLSS model

The theoretical number and the actual number of samples above various cut-offs when sNfL Z-scores were applied on the RDB are in agreement, indicating that the distribution is adequately modelled (**Suppl Table 4**). The remainder of control samples was used to investigate the variability of model parameters and generalisability of the resulting reference curves:
First, we investigated whether the variation in sNfL levels observed within the same control person has an impact on the resulting reference curves by randomly selecting alternative samples per patient (repeated 100 times; **Suppl Figure 5**). Second, we applied a bootstrapping procedure to rule out that the GAMLSS model was overtrained: 100 bootstrap samples (n=4532, with replacement) were drawn from the RDB and the same model was fit on the individual bootstrap replicates. The locations of resulting reference curves were inspected graphically (**Suppl Figure 6**).

Suppl Table 1. Baseline demographic and clinical characteristics of included MS patients in the Epidemiological Investigation of Multiple Sclerosis (EIMS), Immunomodulation and Multiple Sclerosis Epidemiology (IMSE) cohorts (validation cohort) and Comparison Between All immuno-Therapies for Multiple Sclerosis (COMBAT-MS) (validation cohort).

Number of patients (n)	**4341**
Demographic data	
Female (n, %)	3073 (70·8)
Male (n, %)	1268 (29·2)
Age (Y)	38·1 (30·2, 45·9)
Clinical data, samples and follow-up	
RMS (n, %)	3746 (86·3)
SPMS (n, %)	595 (13·7)
Disease duration (Y)	3·0 (1·0, 9·0)
Nr. relapses in last year (mean, SD)	0·3 (0·5)
EDSS	1·5 (0·0, 2·5)
Nr. of serum samples per patient (n)	1·0 (1·0, 2·0)
Disease-modifying treatment at inclusion	
IIEmAb (n, %)	356 (8·2)
Orals	130 (3·0)
Platform	2047 (47·2)
Untreated	1808 (41·6)

Abbreviations: EDSS: Expanded Disability Status Scale; HEmAb: high efficiency monoclonal antibody therapies; MS: multiple sclerosis; Nr.: number; RMS: relapsing MS; SD: standard deviation; SPMS: secondary progressive MS; Y: years.
Numbers are reported as median and interquartile range if not mentioned differently.
"HEmAb" include: alemtuzumab (n=6), natalizumab (n=265), rituximab (n=85); "Orals" include: fingolimod (n=36), dimethyl fumarate (n=88), teriflunomide (n=6); "Platform" includes: all interferon beta (n=1700) and glatiramer acetate (n=347) preparations.

Suppl Table 2. Source of data of persons and samples included into the reference database.

Data Source	**Subjects (n)**	**Time points: Samples (n)**	**Age**		**Selected for RDB* (n)**
			median (IQR)	**range**	
Genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors study (GAPP)[8,9]	2,163	BL: 2,162 FU 1: 1,535	36·8 (31·2-40·2) 41·3 (36·2-44·5)	23·8-44·3 27·9-49·0	1652 511
Multiple Sclerosis Expression, Proteomics, Imaging, Clinical Study (EPIC)[10]	1,181	BL: 1,181 FU 1: 87	44·0 (34·0-53·0) 46·0 (41·0-55·5)	18·0-81·0	295 55
Genetic MS Associations (Gene MSA)[2,11]	259	BL: 259 FU 1: 226	44·3 (36·3-52·3) 45·4 (39·1-53·4)	18·1-70·6 19·2-71·6	118 141
Establishing the links between subclinical arteriosclerosis and depression (BiDirect)[12–14]	1,787	BL: 1,712 FU 1: 1,185 FU 2: 877 FU 3: 909	51·3 (45·1-57·8) 53·6 (47·4-59·8) 55·0 (48·7-61·0) 60·3 (53·9-66·3)	30·9-70·2 36·0-71·2 36·8-72·0 41·2-73·5	555 224 208 773
Total	5,390	10,133	44·7 (38·3-55·2)	18·0-81·0	4532

*A single sample per control person was selected from all with available BMI.
Abbreviations: BL: baseline; FU: follow up; IQR: interquartile range; RDB: reference database.
Age in years. GAPP, EPIC and Gene MSA had one follow up visit. BiDirect had baseline and 3 follow up visits.

Suppl Table 3. sNfL concentrations (pg/ml) corresponding to a certain Z-score (percentile) at a given age and BMI as derived from the statistical model applied on reference database (refer to http://shiny.dkfbasel.ch/baselnflreference to calculate sNfL Z-scores based on BMI and age as continuous units).

		Z-score *(Percentile)*							
Age	**BMI**	**-2**	**-1**	**0**	**1**	**1·5**	**2**	**2·5**	**3**
		(2·28)	***(15·87)***	***-50***	***(84·13)***	***(93·32)***	***(97·72)***	***(99·40)***	***(99·87)***
20	**20**	2·6	3·6	4·8	6·6	8·0	10·3	14·5	24·1
	25	2·1	2·9	3·9	5·4	6·6	8·5	12·0	19·9
	30	1·7	2·4	3·1	4·3	5·3	6·8	9·5	15·9
25	**20**	3·0	4·2	5·5	7·6	9·3	12·0	16·9	28·0
	25	2·6	3·5	4·7	6·5	7·9	10·2	14·3	23·8
	30	2·1	2·9	3·9	5·4	6·6	8·4	11·9	19·8
30	**20**	3·4	4·6	6·2	8·5	10·4	13·4	18·8	31·3
	25	2·9	4·0	5·3	7·4	9·0	11·5	16·3	27·0
	30	2·5	3·4	4·5	6·3	7·6	9·8	13·8	23·0
35	**20**	3·7	5·1	6·7	9·3	11·3	14·6	20·5	34·1
	25	3·2	4·4	5·9	8·1	9·9	12·7	18·0	29·8
	30	2·8	3·8	5·1	7·0	8·6	11·0	15·5	25·8
40	**20**	4·0	5·6	7·4	10·2	12·5	16·0	22·6	37·5
	25	3·6	4·9	6·6	9·1	11·1	14·2	20·0	33·3
	30	3·2	4·3	5·8	8·0	9·7	12·5	17·6	29·3
45	**20**	4·5	6·2	8·2	11·3	13·8	17·8	25·1	41·7
	25	4·0	5·6	7·4	10·2	12·4	16·0	22·5	37·4
	30	3·6	5·0	6·6	9·1	11·1	14·3	20·1	33·4
50	**20**	5·0	6·9	9·2	12·7	15·5	19·9	28·0	46·6
	25	4·6	6·3	8·4	11·5	14·1	18·1	25·5	42·3
	30	4·1	5·7	7·6	10·4	12·7	16·4	23·1	38·3
55	**20**	5·7	7·8	10·4	14·3	17·5	22·5	31·7	52·6
	25	5·2	7·2	9·6	13·2	16·1	20·7	29·1	48·4
	30	4·8	6·6	8·8	12·1	14·7	19·0	26·7	44·4
60	**20**	6·5	8·9	11·8	16·3	19·9	25·6	36·0	59·9
	25	6·0	8·3	11·0	15·1	18·5	23·8	33·5	55·7
	30	5·6	7·7	10·2	14·1	17·2	22·1	31·1	51·7
65	**20**	7·3	10·1	13·4	18·5	22·6	29·1	40·9	68·0
	25	6·9	9·5	12·6	17·4	21·2	27·3	38·4	63·8
	30	6·4	8·9	11·8	16·3	19·9	25·5	36·0	59·8
70	**20**	8·3	11·4	15·1	20·8	25·4	32·7	46·1	76·6
	25	7·8	10·7	14·3	19·7	24·0	30·9	43·5	72·3
	30	7·4	10·1	13·5	18·6	22·7	29·2	41·1	68·3
75	**20**	9·2	12·7	16·9	23·2	28·3	36·4	51·3	85·3
	25	8·7	12·0	16·0	22·0	26·9	34·6	48·8	81·0

	30	8·3	11·4	15·2	21·0	25·6	32·9	46·3	77·0

Abbreviations: BMI: Body mass index; sNfL: serum neurofilament light chain.
A Z-score of 0 (50th percentile) represents the mean=median sNfL in a normal population (after adjustment for BMI and age).
Example: a 40-year-old patient with a BMI of 25 and sNfL level of 9·1 pg/ml reflects an adjusted sNfL value above the 84th percentile or a Z-score of 1. Please note, the resource https://shiny.dkfbasel.ch/baselsnfreference will allow to enter the exact age, height and weight of the patient and will calculate the exact sNfL percentile and Z-score, i.e. not in age and BMI steps of 5 years/kg/m^2 or 0·5/1·0 Z-score increments.

Suppl Table 4. Number of samples above certain thresholds in the reference database and in the Swiss MS Cohort Study.

sNfL Z-Score (percentile)	**Reference Database (RDB)**			**Swiss MS Cohort (SMSC)**		**Quotient: SMSC (%)/RDB (%)**
	n	**%**	**% theor.***	**N**	**%**	
>1 (84·13th)	712	15·7	15·9	1994	25·7	1·6
>1·5 (93·32nd)	291	6·4	6·7	1124	14·5	2·2
>2 (97·72nd)	102	2·3	2·3	528	6·8	3·0
>2·5 (99·40th)	30	0·7	0·6	266	3·4	5·5
>3 (99·87th)	11	0·2	0·1	97	1·2	

Abbreviations: sNfL: serum neurofilament light chain.
*Theoretically (according to a standard normal distribution), e.g., 2·28% of samples are expected to have a Z-score above 2 (see also **Suppl Figure 4**). The observed number of samples in the reference database above the cut-off estimated by the model (2·3%) is very close to the theoretical proportion. Overall, in the Swiss MS Cohort this proportion is 3 times higher than in the reference database (see column 7).

Suppl Table 5. Sensitivity and specificity of a fixed sNfL cut-off in detecting disease activity as shown in Figure 3.

			Proportion of samples above/below cut-off			
			> 10pg/ml		**≤10pg/ml**	
Age range (years)	**sNfL Z-score**	**N**	**n**	**%**	**n**	**%**
20-30	>2	164	157	95·7	7	4·3
	1·5-2	103	33	32·0	70	68·0
	0-1·5	409	0	0·0	409	100·0
	≤0	386	0	0·0	386	100·0
30-60	>2	317	317	100·0	0	0·0
	1·5-2	419	381	90·9	38	9·1
	0-1·5	2517	983	39·1	1534	60·9
	≤0	2731	28	1·0	2703	99·0
>60	>2	47	47	100·0	0	0·0
	1·5-2	74	74	100·0	0	0·0
	0-1·5	292	292	100·0	0	0·0
	≤0	310	156	50·3	154	49·7
Total		7769	2468	31·8	5301	68·2

Abbreviations: sNfL: serum neurofilament light chain.

Suppl Table 6. Demographic and clinical disease characteristics at time of first sampling from 2348 treatment epochs in 1313 patients.

	Treatment category (treatment epochs)			
	Untreated (n=535)	**Platform (n=262)**	**Orals (n=891)**	**HEmAb (n=660)**
Demographic data of patients				
Sex (female, n, %)	365 (68·2)	186 (71)	570 (66·7)	431 (68·4)
Age (Y)	40·0 (31·3-51·5)	41·1 (32·4-48·6)	41·1 (32·6-49·4)	39·8 (31·1-48·6)
Clinical data				
Disease duration (Y)	6·1 (1·3-14·2)	5·1 (2·0-11·3)	7·6 (2·8-14·5)	9·7 (4·8-17·1)
EDSS	2·0 (1·5-3·5)	2·0 (1·0-2·5)	2·0 (1·5-3·0)	2·5 (2·0-4·0)
SPMS (n, %)	67 (12·5)	14 (5·4)	9 (1·1)	52 (8·3)
Treatment initiation or switch				
Time to treatment initiation or switch (Y)	0·2 (0·0-0·7)	0·7 (0·1-1·7)	2·0 (0·9-3·7)	1·7 (0·8-2·6)
Untreated	190 (35·5)	93 (35·5)	543 (60·9)	421 (63·8)
Platform	71 (13·3)	27 (10·3)	20 (2·2)	4 (0·6)
Orals	198 (37·0)	119 (45·4)	129 (14·5)	83 (12·6)
HEmAb	76 (14·2)	23 (8·8)	199 (22·3)	152 (23·0)
Samples				
Interval untreated/DMT start to first sample (Y)	0·6 (0·2-1·8)	0·9 (0·4-3·3)	0·5 (0·2-1·0)	0·5 (0·3-1·5)
Nr. of samples	1150	720	3779	2120
DMT				
Interferon beta 1a (Avonex, Plegridy, Rebif)		118 (45·0)		
Interferon beta 1b (Betaferon, Extavia)		51 (19·5)		
Glatiramer acetate		93 (35·5)		
Dimethyl fumarate			233 (26·2)	
Fingolimod			588 (66·0)	
Siponimod			5 (0·6)	
Teriflunomide			65 (7·3)	
Alemtuzumab				19 (2·9)
Natalizumab				302 (45·8)
Ocrelizumab				206 (31·2)
Rituximab				133 (20·2)

Abbreviations: DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale; HEmAb: high efficacy monoclonal antibody therapies; Nr.: number; SPMS: secondary progressive multiple sclerosis; Y: years.
Numbers are reported as median and interquartile range and as counts and percentages if not mentioned otherwise.

Suppl Table 7. Composition of the 2348 treatment epochs.

Treatment Category	DMT Name	Active compound	Samples (n)	Epochs (n)	Patients (n)
Platform	Avonex	Interferon beta-la	187	56	54
Platform	Plegridy	Peginterferon beta-la	25	12	11
Platform	Rebif	Interferon beta-la	111	50	47
Platform	Betaferon	Interferon beta-lb	164	49	49
Platform	Extavia	Interferon beta-lb	16	2	2

Platform	Copaxone	Glatiramer acetate	217	93	82
Oral	Tecfidera	Dimethyl fumarate	747	233	224
Oral	Gilenya	Fingolimod	2812	588	554
Oral	Mayzent	Siponimod	21	5	5
Oral	Aubagio	Teriflunomide	199	65	62
HEmAb	Lemtrada	Alemtuzumab	46	19	19
HEmAb	Tysabri	Natalizumab	1132	302	288
HEmAb	Ocrevus	Ocrelizumab	542	206	206
HEmAb	Mabthera	Rituximab	400	133	119
Untreated			1150	535	459
Total			7769	2348	

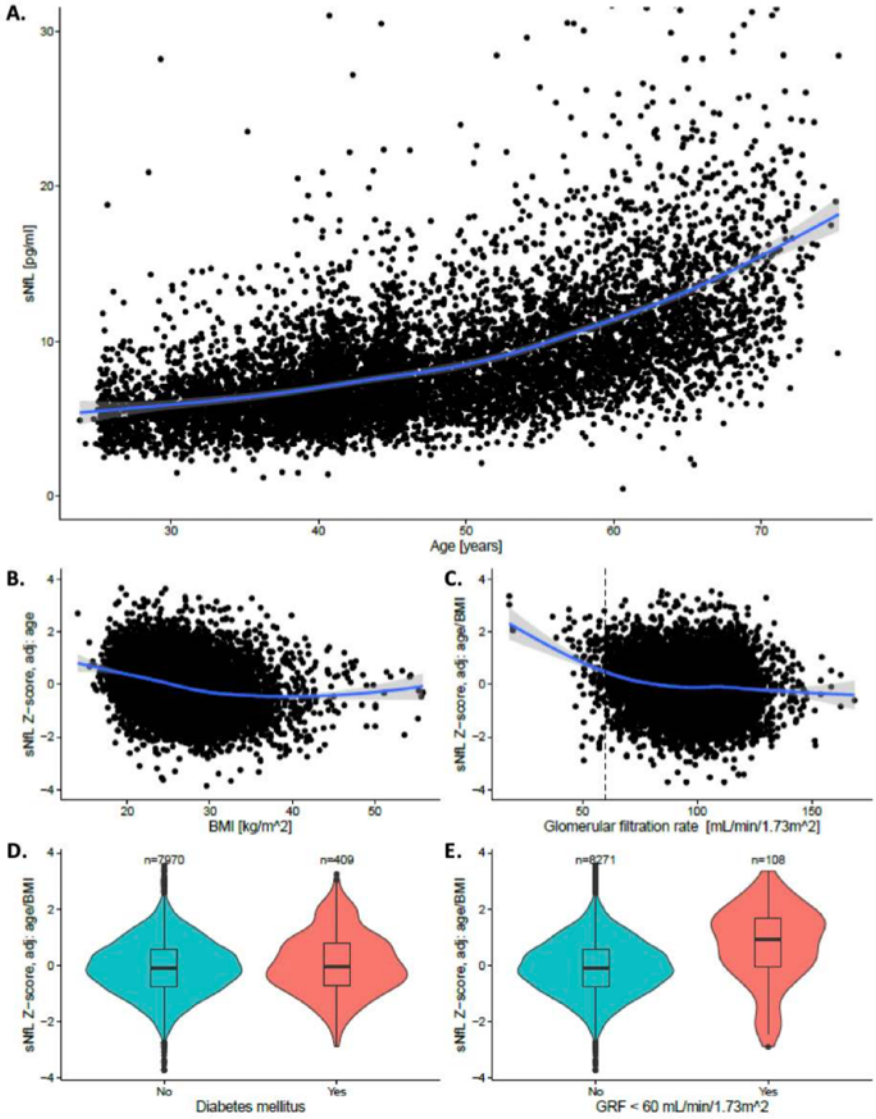
Abbreviation: DMT: disease modifying treatment; HEmAb: high efficacy monoclonal antibody therapies.

Suppl Table 8. Multivariable mixed-effects model with sNfL Z-score as dependent variable as shown in Figure 2.

Covariate (n=3868)	**Estimate**	**95% CI**	**P-value**
Sex (f vs. m)	0·08	-0·04;0·21	0·20
Disease duration (per 10 years)	-0·19	-0·26;-0·12	<0·0001
SPMS vs. RMS	0·20	-0·02;0·42	0·076
Relapse in last 4 months (yes vs. no)	0·48	0·37;0·59	<0·0001
Number of CEL	0·14	0·11;0·17	<0·0001
EDSS (per unit increase)	0·11	0·07;0·14	<0·0001
Total T2w volume (per log10 unit)	0·19	0·13;0·24	<0·0001
Platform vs untreated	-0·13	-0·28;0·03	0·10
Orals vs untreated	-0·35	-0·46;-0·24	<0·0001
HEmAb vs. untreated	-0·49	-0·62;-0·37	<0·0001

Abbreviations: CEL: contrast enhancing T1 weighted lesions; CI: confidence interval; EDSS: Expanded Disability Status Scale score; F: female; HEmAb: high efficacy monoclonal antibody therapies; M: male; RMS: relapsing remitting multiple sclerosis; SPMS: secondary progressive multiple sclerosis; T2w: T2 weighted.

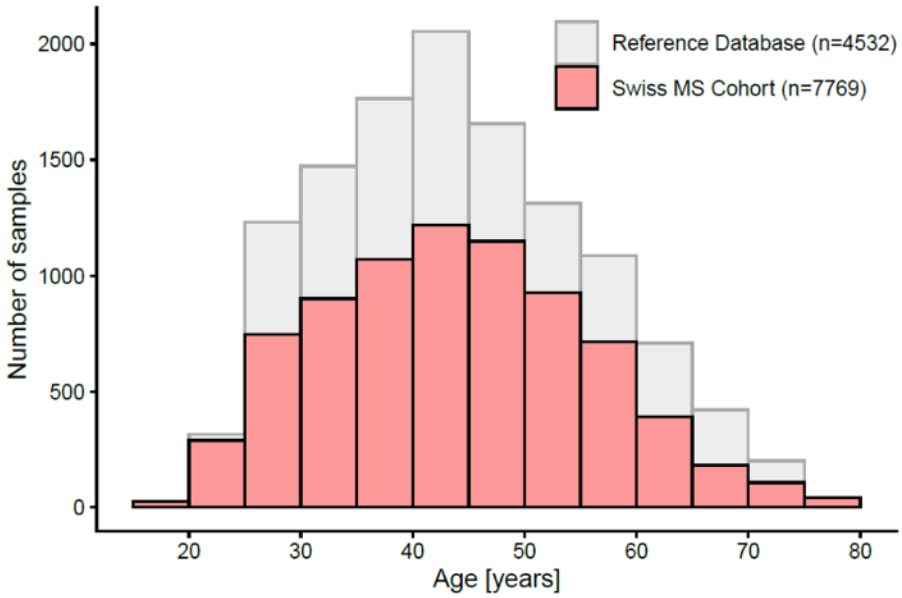
Suppl Figure 2. Associations between age, BMI and renal function and sNfL concentrations in 8379 serum samples from 3950 persons with complete information on relevant comorbidities.



Legend: In a mixed effects linear model with log(NfL) as dependent variable (coefficient of determination of the overall model: R^2: 0·417), the following factors were independently associated with sNfL: **A.**) age (per year; estimate: 1·025; 95% CI: 1·025-1·026; p<0·0001; i.e. 2·5% increase in sNfL concentration per year of age; R^2 remaining: 0·031, when excluding this variable; **B.**) BMI (per 1 kg/m^2; 0·980; 0·978-0·982; p<0·0001; remaining R^2: 0·374; **C.**), eGFR < 60 mL/min/1·73 m^2 (1·215; 1·139-1·295; p<0·0001; remaining R^2: 0·415; **D.** and **E.**), diagnosis of diabetes mellitus (1·119; 1·064-1·177; p<0·0001; remaining R^2: 0·413); and eGFR < 60 ml/min/1·73m^2 (1·215; 1·139-1·295; p<0·0001; remaining R^2: 0·415).

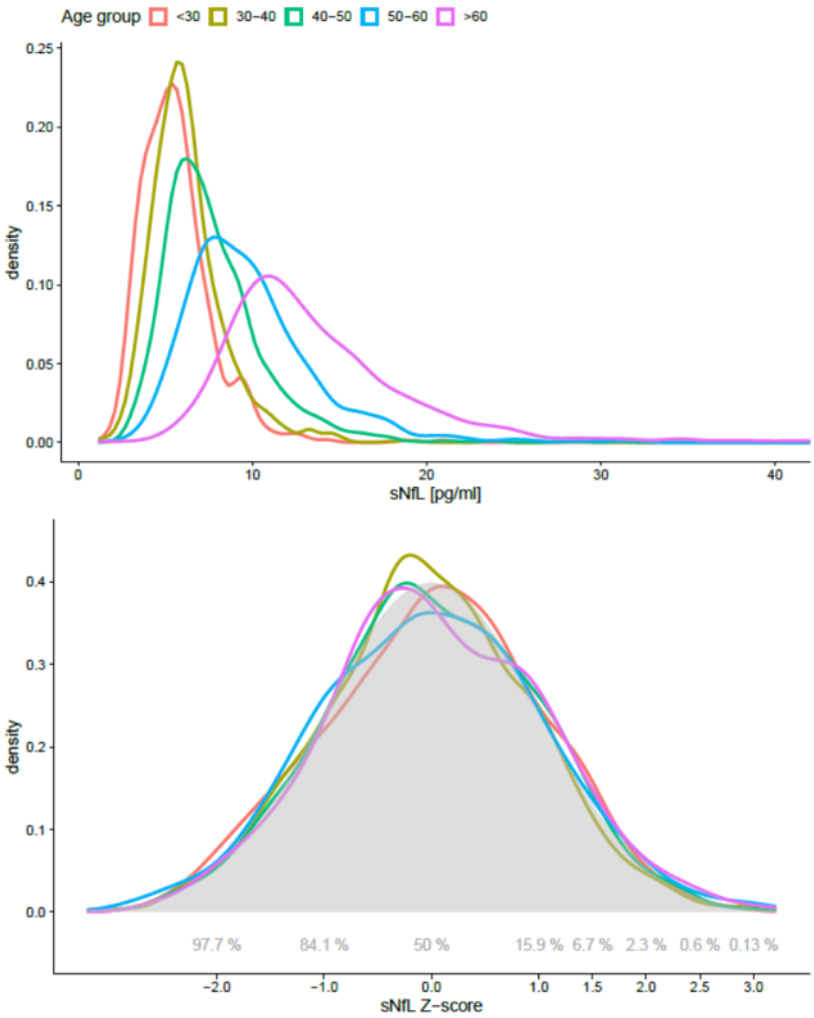
Generation of the reference database includes the effects of BMI (relevant amount of variance explained in the model), while renal function had an overall minor impact on the model due to the low prevalence, however eGFR < 60 ml/min/1·73m^2 can lead to high sNfL Z-scores (see **E.**). Control persons with eGFR < 60 ml/min/1·73m^2 (n: 28) were therefore excluded from the generation of the RDB and interpretation of sNfL levels in patients with an eGFR <60 ml/min/1·73m^2 are limited. Blue lines: Non-parametric smoothing lines with 95% confidence bands.

Suppl Figure 3. Number of samples per age group in participants of the reference database and of patients in the Swiss Multiple Sclerosis Cohort Study (SMSC)



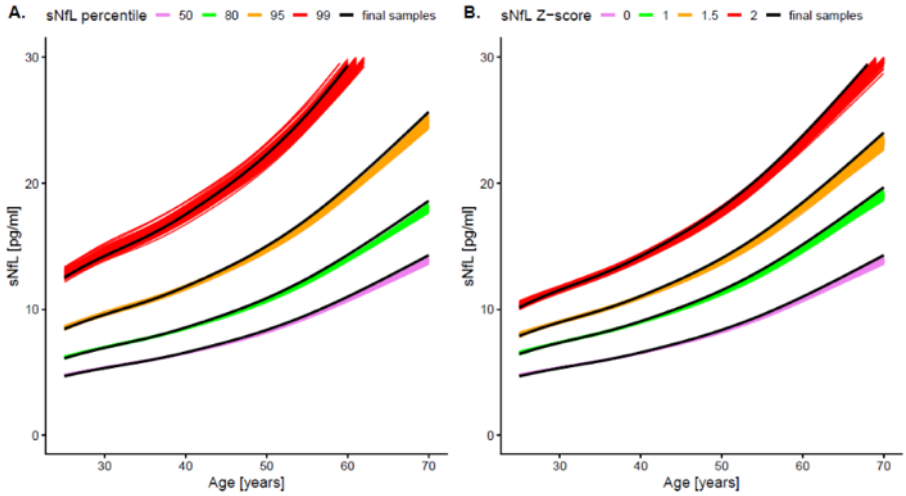
Legend: Age distribution at the time of sampling in samples from the reference database (n: 4,532) and in patients with relapsing or secondary progressive multiple sclerosis (MS) from the Swiss MS Cohort (n: 7,769 samples) covering six decades of life.

Suppl Figure 4. Density plot of absolute sNfL (top) and sNfL Z-scores (below) in 4,532 serum samples from the reference database stratified by age groups.



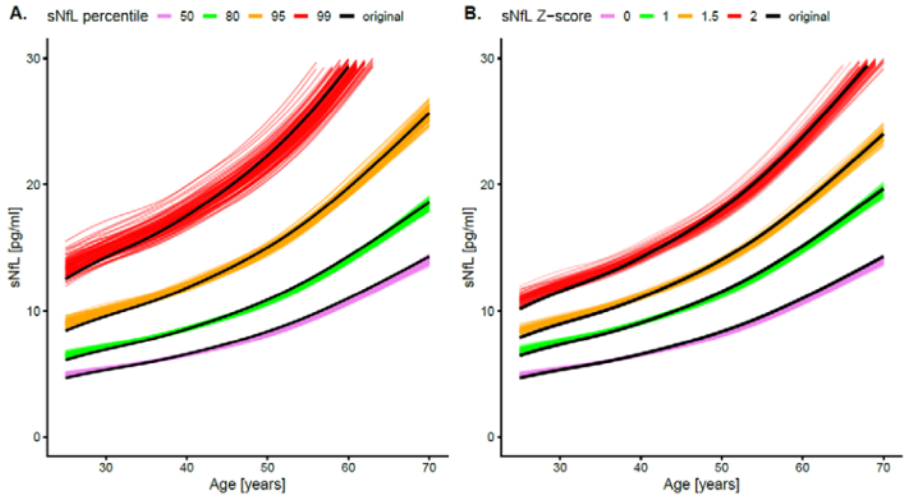
Legend: Top: Density plot of absolute sNfL and sNfL Z-scores in 4,532 serum samples from the reference database stratified by age groups.
Below: Standard normal distribution (in grey) for comparative reasons to visualize that sNfL Z-scores are transformed sNfL values independent of age and BMI which are normally distributed. The expected proportion of samples above a given sNfL Z-score cut-off is denoted (grey writing). The x-axis in the top figure was capped at 40 pg/ml.

Suppl Figure 5. Comparison between the final reference database (using single sample per control person) versus randomly selected sample. Percentiles (A.) and Z-scores (B.).



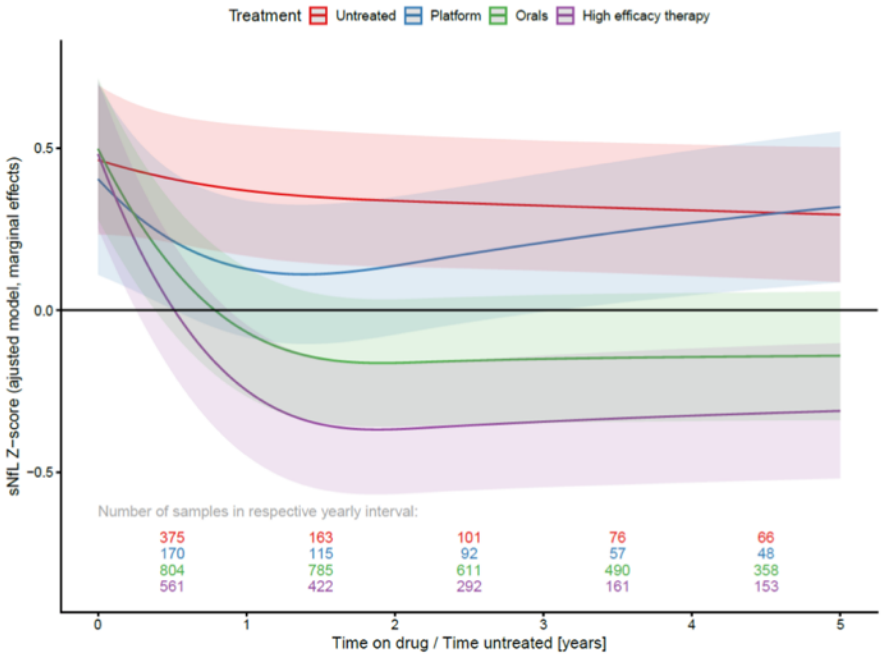
Legend: The black lines indicate (A.) percentiles and (B.) Z-score reference curves of the final RDB based on a single sample per control subject.
Coloured bands for respective percentiles (A.) and Z-sores (B.) using all samples were created by randomly selecting one sample from each person and by building the model 100 times on the resulting dataset (number of patients = number of samples). The reference curves of the final dataset/model (black lines) as compared with those of the 100 random selections from all control person samples (coloured bands) show minimal differences.
The curves are calculated assuming a body mass index (BMI) of 25.

Suppl Figure 6. Impact of alterations of the RDB on the location of the reference curves (bootstrapping): percentiles (A.) and Z-scores (B.)



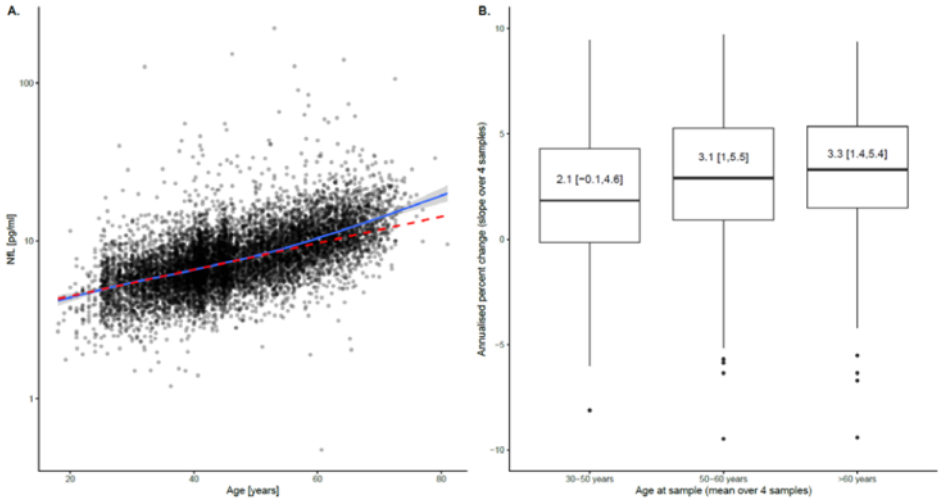
Legend: We investigated whether changes to the constitution of the RDB have impact on the location of reference curves (A: percentiles, B: Z-scores). We used a bootstrapping approach in which we selected 100 times a random sample of the original RDB based on which we built alternative, bootstrapped models. The individual lines represent the references curves derived from the bootstrapped GAMLSS models. The reference curves of the final dataset/model (black lines) as compared with those of the 100 bootstrap samples (coloured bands) show minimal differences.
The curves are calculated assuming a body mass index (BMI) of 25.

Suppl Figure 7. Modelling of the evolution of sNfL Z-scores in four treatment categories using mixed-effects models: adjusted model as sensitivity analysis.



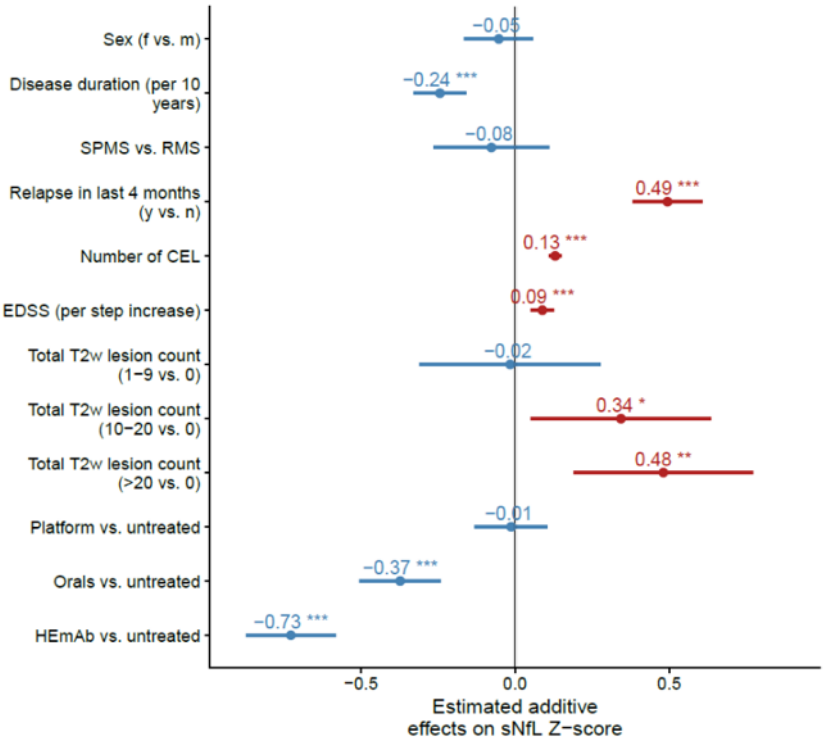
Legend: Temporal evolution of sNfL Z-scores (marginal effects) over time in four treatment categories using a mixed-effects model adjusted for all clinical covariates (sex, disease duration, SPMS vs. RMS, presence of relapse in the last 4 months, EDSS) as well as age was built. Estimated marginal effects (disease activity captured by sNfL Z-scores beside disease activity expressed by the other covariates) show the same hierarchy of DMT efficacy as compared with the model presented in **Figure 5** and the multivariable model in **Figure 2**. The number of samples in the respective yearly interval is shown in the different treatment groups.
Abbreviations: DMT: Disease modifying treatment; EDSS: Expanded Disability Status scale; sNfL: serum neurofilament light chain; RMS: Relapsing MS; SPMS: secondary progressive MS.

Suppl. Figure 8. sNfL values increase exponentially in function of age in control persons on the group level in all 10133 samples (A.) and on the individual level in samples of the same control persons (B.) at an increased rate over 50 years of age.



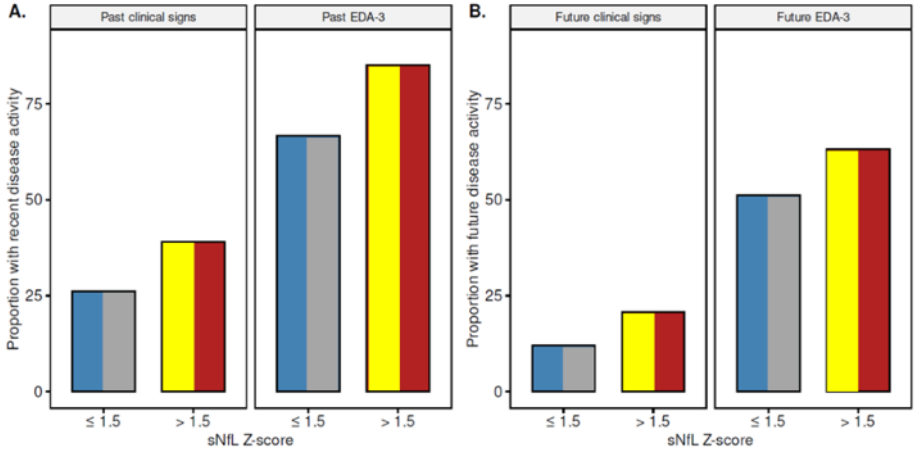
Legend: A. Scatter plot of sNfL on a log-scale and age in 10,133 control samples. A smoothing line (blue) and in red the log-linear model (log(NfL) explained by age) is depicted which was fitted on a subset of the data below age 50 and extrapolated on the entire age range to visualize the increase in slope after around age of 50. This representation shows that the association between sNfL and age is not constant on the log-scale (i.e., non log-linear), indicating that the percent change in sNfL per age increases after an age of approximately 50.
B. Annualised percent change within the same individual in 623 control persons from BiDirect (**Suppl Table 2**) with 4 available samples. The distribution of annualised percent changes (individual regression slopes over the 4 samples per patient) is shown along the mean age at follow-up. The figure indicates that the percent increase in sNfL is not constant but is higher in older subjects from a median increase of 2·1% in individuals 30-50 years up to 3·3% in individuals >60 years (i.e., is not linear and not log-linear).
Abbreviations: NfL: serum neurofilament light chain.

Suppl Figure 9. Factors influencing plasma NfL values converted into sNfL Z-scores in MS in the validation cohort.



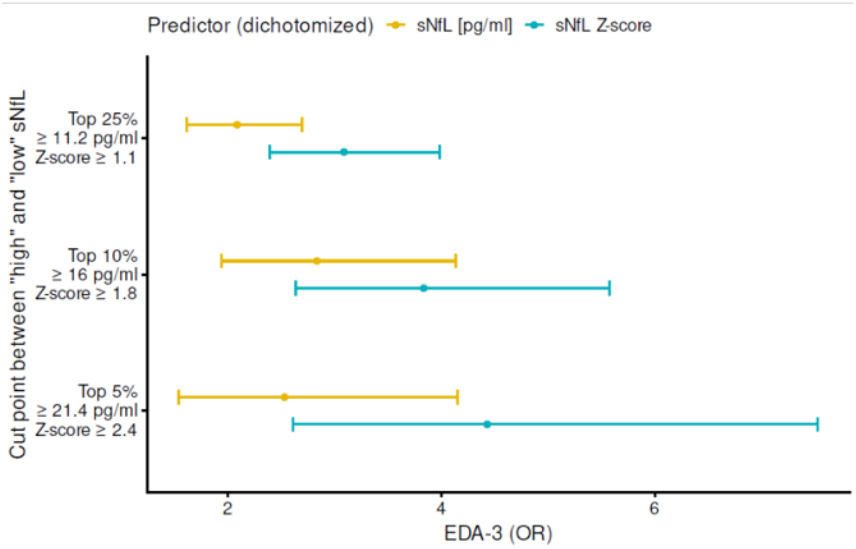
Legend:
Model estimates including 95% confidence intervals. Estimates for HEmAb vs oral therapy was -0·36, 95% CI: -0·51--0·20, p<0·0001; for oral vs. platform therapy the estimate was -0·36, 95% CI: -0·49--0·23, p<0·0001.
***: p<0·001; **: p<0·01; *: p<0·05.
Abbreviations: CEL: contrast enhancing T-weighted 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.

Suppl Figure 10. Proportion of 20-30 years old patients with sNfL below fixed cut-off of 10pg/ml but increased sNfL Z-scores and with clinical signs or EDA-3 in the last year (**A.**) or in the following year after sampling (**B.**).



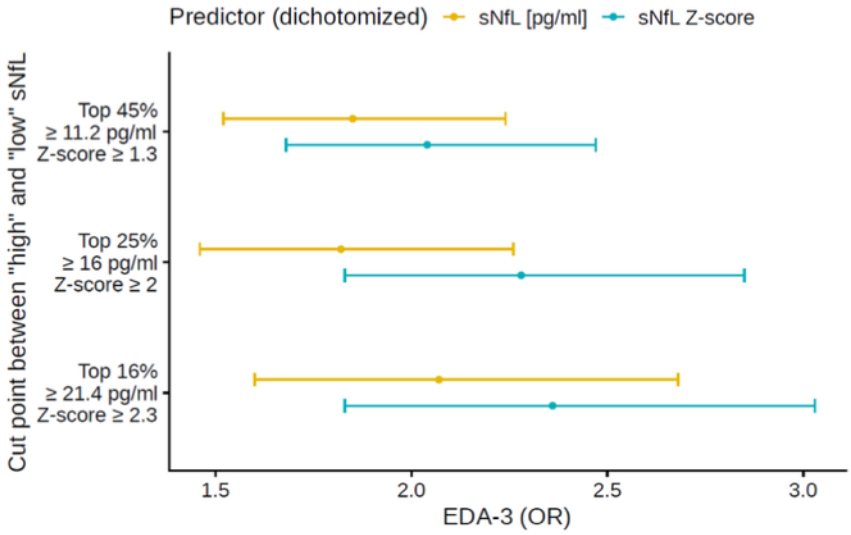
Legend: Patients under 30 years with increased sNfL Z-scores (>1.5; color code according **Figure 3**) but below a fixed cut-off of 10 pg/ml showed more recent clinical and MRI disease activity (clinical signs or EDA-3) compared with patients with sNfL Z-scores of ≤ 1.5 both the year before (**A.**) and after (**B.**) sampling.
Abbreviations: sNfL: serum neurofilament light chain; EDA-3: evidence of disease activity-3.

Suppl Figure 11. Comparison of sNfL (pg/ml) versus sNfL Z-scores predicting disease activity (EDA-3) in the following year.



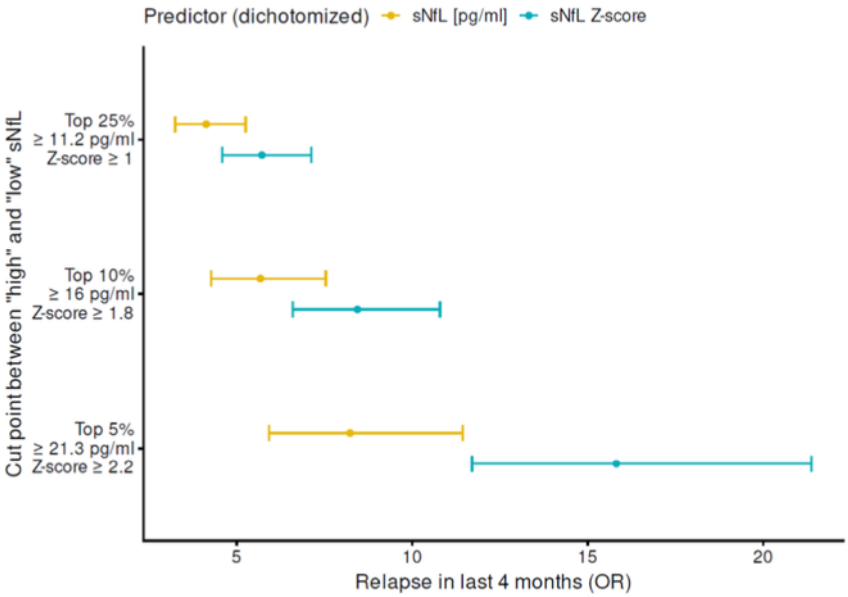
Legend:
Comparison of magnitude of association with future disease activity (EDA-3) between absolute sNfL concentration and sNfL Z-scores. sNfL concentrations and sNfL Z-scores were dichotomized in "high" (i.e., potentially pathological) and "low" values (*e.g.*, top 25% of sNfL concentration (≥11·2pg/ml)/sNfL Z-score (≥1·1) vs 75% remaining of sNfL concentrations/sNfL Z-scores etc). Z-scores led to consistently higher odds ratios (OR) 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, p<0·0001, respectively. Estimated odds ratios and 95% confidence intervals are shown.
Abbreviations: EDA-3: evidence of disease activity-3; OR: odds ratio; sNfL: serum neurofilament light chain.

Suppl Figure 12. Comparison of plasma NfL values converted into sNfL (pg/ml) versus sNfL Z-scores predicting disease activity (EDA-3) in the following year in the validation cohort.



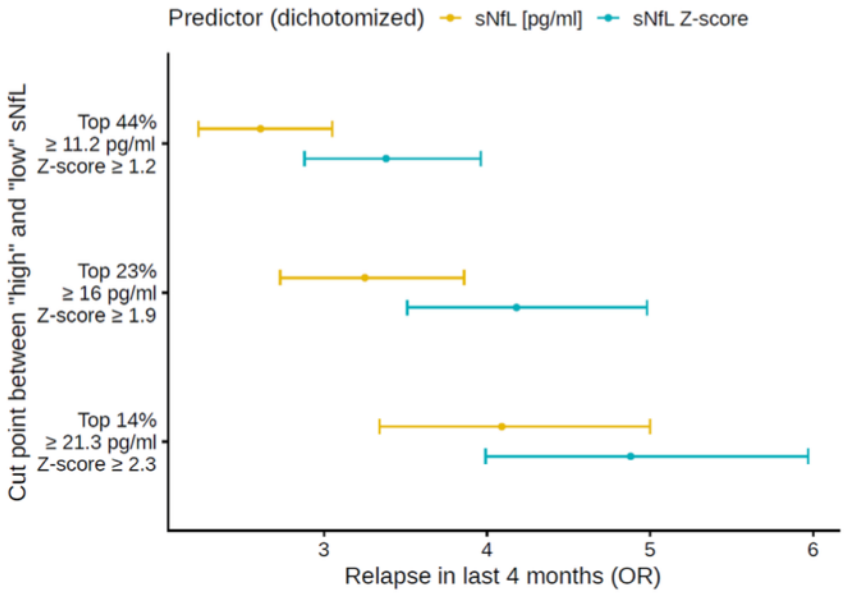
Legend:
Comparison of magnitude of association with future disease activity (EDA-3) between absolute sNfL concentration and sNfL Z-scores. sNfL concentrations and sNfL Z-scores were dichotomized in "high" (i.e., potentially pathological) and "low" values (*e.g.*, top 45% of sNfL concentration (≥11·2pg/ml)/sNfL Z-score (≥1·3) vs 55% remaining of sNfL concentrations/sNfL Z-scores etc). Z-scores led to consistently higher odds ratios (OR) using 3 different cut-offs ("high" defined as top 45, 25 or 16% of the samples): absolute sNfL levels vs sNfL Z-scores: $OR_{top\ 45\%}$: 1·85 vs 2·04; $OR_{top\ 25\%}$: 1·82 vs 2·28; $OR_{top\ 16\%}$: 2·07 vs 2·36, p<0·0001, respectively. Estimated odds ratios and 95% confidence intervals are shown.
Abbreviations: EDA-3: evidence of disease activity-3; OR: odds ratio; sNfL: serum neurofilament light chain.

Suppl Figure 13. Comparison of magnitude of association with recent relapse (≤4 months) between sNfL concentration and sNfL Z-scores.



Legend: The association between having experienced a relapse within the past 4 months before sampling and sNfL Z-scores was stronger vs absolute sNfL levels (odds ratios, $OR_{sNfL\ Z\text{-}scores}$: 5·72-15·82, p<0·001 vs $OR_{absolute\ sNfL\ levels}$: 4·13-8·24, p<0·0001). sNfL concentrations and sNfL Z-scores were dichotomized in 'high' and 'low' values (e.g., top 25% of sNfL concentration (≥11·2pg/ml)/sNfL Z-score (≥1·1) vs 75% remaining of sNfL concentrations/sNfL Z-scores etc). Z-scores led to consistently higher odds ratios using 3 different cut-offs ('high' defined as top 25, 10 or 5% of the samples).
Abbreviations: sNfL: serum neurofilament light chain; OR: odds ratio.

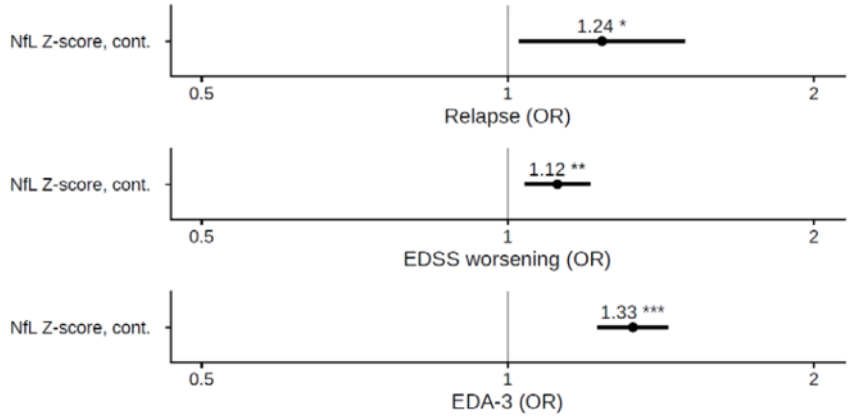
Suppl Figure 14. Comparison of magnitude of association with recent relapse (≤4 months) between plasma NfL concentration, converted into sNfL absolute values and sNfL Z-scores, in the validation cohort.



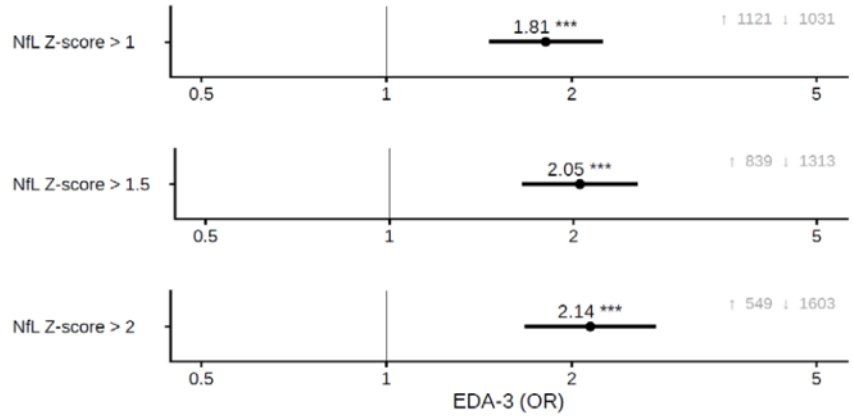
Legend: The association between having experienced a relapse within the past 4 months before sampling and sNfL Z-scores was stronger vs absolute sNfL levels (odds ratios, $OR_{sNfL\ Z\text{-}scores}$: 3·38-4·88, p<0·001 vs $OR_{absolute\ sNfL\ levels}$: 2·61-4·09, p<0·0001). sNfL concentrations and sNfL Z-scores were dichotomized in 'high' and 'low' values (e.g., top 44% of sNfL concentration (≥11·2pg/ml)/sNfL Z-score (≥1·2) vs 56% remaining of sNfL concentrations/sNfL Z-scores etc). Z-scores led to consistently higher odds ratios using 3 different cut-offs ('high' defined as top 44, 23 or 14% of the samples).
Abbreviations: sNfL: serum neurofilament light chain; OR: odds ratio.

Suppl Figure 15. sNfL Z-scores predicting disease activity in the following year in the validation cohort, based on converted plasma values: **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.

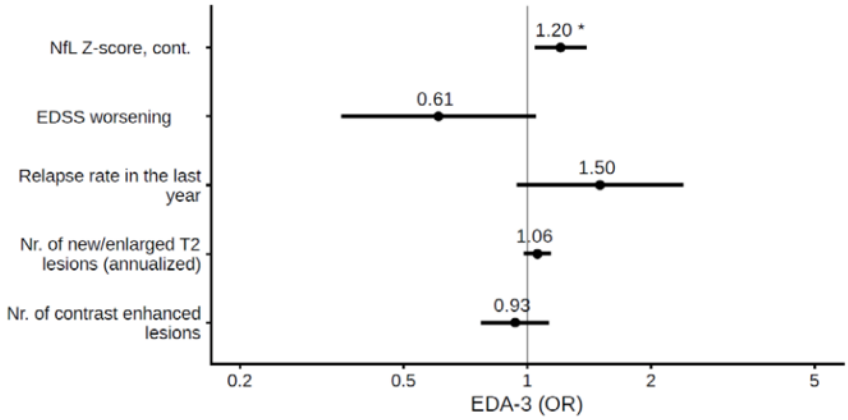
A.



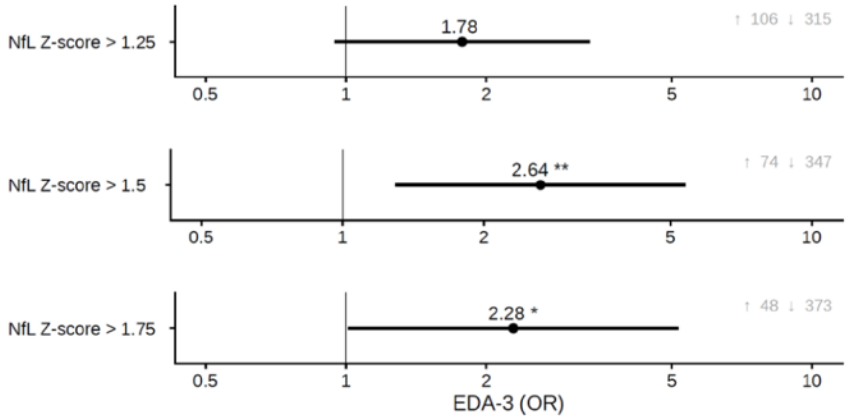
B.



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 2·1-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 20% per 1 step higher sNfL Z-score.

D.

NEDA-3 patients with sNfL levels above the 93·3rd percentile (Z-score >1·50) displayed a 2·64-fold (95% CI 1·30-5·37; p=0·0074) higher risk of experiencing EDA-3 in the following year.

***: p<0·001; **: p<0·01; *: p<0·05. Estimates and 95% confidence intervals are shown.

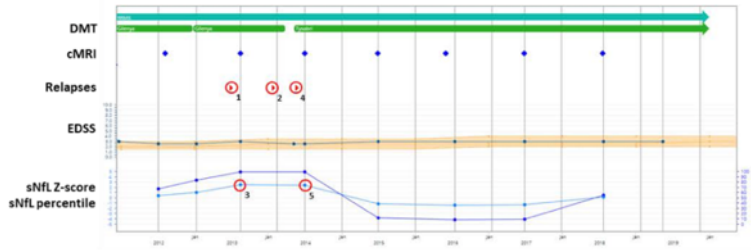
A., **B.** and **D.** show 3 univariable and **C.** a multivariable model.

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.

Suppl Figure 16. Use cases 1-6 illustrating the application of sNfL Z-scores/percentiles.

Use case 1A: Normalisation of sNfL with highly effective treatment 1

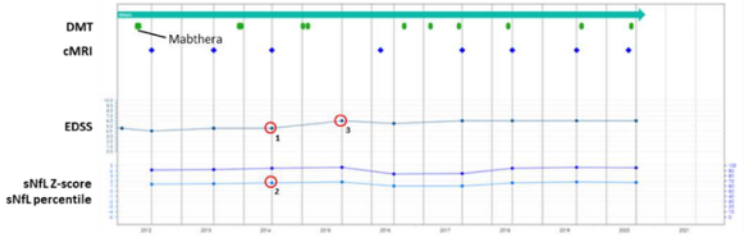


Patient with relapses under Fingolimod treatment (*red circles 1 and 2*) and increased sNfL levels on August 15, 2013 (99.4th percentile, Z-score: 2.5; *red circle 3*). Fingolimod was stopped due to persisting disease activity on March 24, 2014 and Natalizumab initiated on May 9, 2014 with additional relapse May 25, 2014 (*red circle 4*) and persistingly high sNfL levels on July 2, 2014 (99.3th percentile, Z-score: 2.5; *red circle 5*), which normalised after Natalizumab initiation: sNfL percentiles between 8.5-55.0/Z-scores -1.4-0.1 between July 1, 2015 and July 24, 2018.

Use case 1B: Normalisation of sNfL with highly effective treatment 2

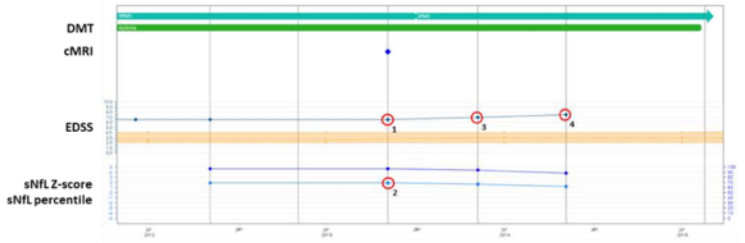
Persistently increased sNfL percentiles (94.5th-99.6th) and Z-Scores (1.6-2.7, respectively; *red circles 1 and 2*) under Glatiramer acetate and Fingolimod between November 20, 2014 and November 10, 2016 with a relapse on November 16, 2016 (*red circle 3*). Switch to Rituximab with gradual decrease in sNfL to levels between 40.0th - 64.0th percentile/-0.4 - -0.3 Z-score, respectively between November 26, 2019 (*red circle 4*) and January 5, 2021.

Use case 2: NEDA-3: Prediction of progression under B-cell depleting therapy



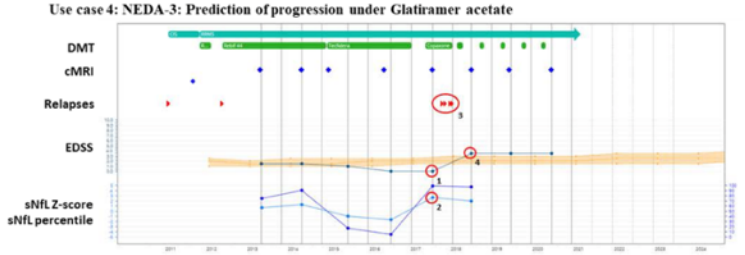
Patient fulfills NEDA-3 on August 12, 2014 with an EDSS score of 4.5 (*red circle 1*), treated with Rituximab and sNfL Z-score elevated at 1.6 (*red circle 2*) (95th percentile). EDSS score progression to 6.0 (*red circle 3*) on October 13, 2015 and 5.5 confirmed on August 25, 2016 (thereafter 6.0 at all subsequent visits). sNfL persistently high until end of follow up (sNfL Z-score: 1.0-1.9, 84.0 to 96.8th percentile) (not enough data points to illustrate median (IQR) EDSS range of similar patients with 31 years of disease duration).

Use case 3: NEDA-3: Prediction of progression under Fingolimod



Patient fulfills NEDA-3 on November 5, 2013 with an EDSS score of 6.5 (*red circle 1*), treated with Fingolimod and sNfL Z-score elevated at 1.9 (*red circle 2*) (97th percentile). EDSS score progression to 7.0 (*red circle 3*) on May 9, 2014 and 7.5 on November 6, 2014 (*red circle 4*).

Use case 4: NEDA-3: Prediction of progression under Glatiramer acetate



Patient fulfills NEDA-3 on December 7, 2017 with an EDSS score of 0 (*red circle 1*), treated with Glatiramer acetate and sNfL Z-score elevated at 2.7 (*red circle 2*) (99.6th percentile). Patient experiences 4 relapses between March 1, 2018 and June 15, 2018 (*red circle 3*) and EDSS increase to 3.5 on November 20, 2018 (*red circle 4*). Glatiramer acetate stopped June 1, 2018 and Ocrelizumab started July 11, 2018.

Use case 5: Disease activity during and after pregnancy

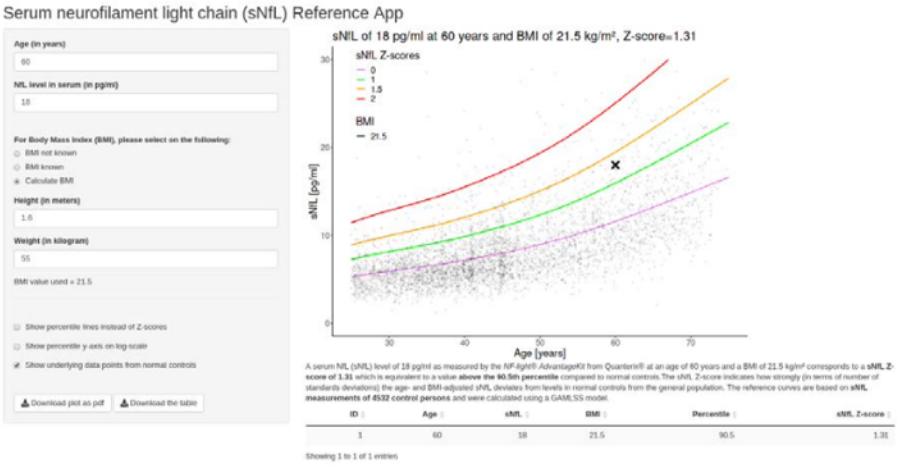
Increase in sNfL Z-score to 96.8th percentile (*red circle 1*) (Z-score: 1.9) and 78.0th percentile (*red circle 2*) (Z-score: 0.8) on January 31, 2017 and July 11, 2017, respectively under Fingolimod treatment. Fingolimod interrupted due to pregnancy wish November 20, 2017 and birth June 26, 2018 (*red circle 3*) with relapse during pregnancy February 26, 2018 (*red circle 4*). With restart of Fingolimod and eventually Ocrelizumab (due to reduction in walking distance) normalisation of sNfL levels between 6.6-21.0st percentile (-1.5 - -0.8 Z-score).

Use case 6: Minimal disease activity under Tecfidera: potentially earlier escalation to HEmAb

Patient with one small new T2w periventricular lesion with faint contrast enhancement on June 2, 2017 (*see red circle 1*). sNfL Z-score increased to 2.2 (*red circle 2*) (98.6th percentile) on June 15, 2017. Dimethylfumarate (DMF) was continued. On January 12, 2018 patient experienced relapse (*red circle 3*) under DMF with slight atactic left sided hemiparesis (not EDSS score relevant). On January 18, DMF was stopped and Ocrelizumab started on February 7, 2018: sNfL Z-scores decreased to levels between -0.4 (36th percentile) and -1.9 (2.7th percentile).

Legend: Case studies from the Swiss MS Cohort Study (SMSC) to illustrate the clinical meaning and use of sNfL Z-scores/percentiles.
Symbols in the cartoons are used as per legend on the left of image. Brown band: median (IQR) EDSS range of patients with same disease duration based on all EDSS scores (>12 000 scores in 1516 individuals) and disease duration information in the SMSC as of 20 Apr 2021). For Ocrelizumab and Mabthera individual infusions (600 mg vor Ocrevus, except 2x300 mg for initial two doses and 1000 mg for Mabthera) are displayed.
Abbreviations: DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale; NEDA-3: no evidence of disease activity 3; sNfL: serum neurofilament light chain

Suppl Figure 17. Screenshot of the sNfL Shiny App for calculation of sNfL percentile and Z-score values.



Legend: Web application (http://shiny.dkfbasel.ch/baselnflreference) providing access to the GAMLSS model built from the RDB of 4532 serum sample from a general population. Given BMI, age and sNfL concentration of a specific patient, the App calculates the sNfL percentile and Z-score, and provides a graphical representation of the patients' sNfL level vis-à-vis values in the RDB.

Supplementary References

1 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**: 1–13.

2 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.

3 Teunissen CE, Petzold A, Bennett JL, *et al.* A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. Neurology. 2009; **73**: 1914–22.

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

5 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.

6 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.

7 Frisell T, Forsberg L, Nordin N, et al. Comparative analysis of first-year fingolimod and natalizumab drug discontinuation among Swedish patients with multiple sclerosis. Mult Scler 2016; 22: 85–93.

8 Conen D, Schön T, Aeschbacher S, *et al.* Genetic and phenotypic determinants of blood pressure and other cardiovascular risk factors: Methodology of a prospective, population-based cohort study. *Swiss Med Wkly* 2013; **143**: 1–9.

9 Krisai P, Aeschbacher S, Ruperti Repilado FJ, *et al.* Healthy lifestyle and glucagon-like peptide-1 in young and healthy adults: A population-based study. *Prev Med (Baltim)* 2017; **101**: 72–6.

10 Baranzini SE, Wang J, Gibson RA, *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 2009; **18**: 767–78.

11 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.

12 Teuber A, Sundermann B, Kugel H, *et al.* MR imaging of the brain in large cohort studies: feasibility report of the population- and patient-based BiDirect study. *Eur Radiol* 2017; **27**: 231–8.

13 Teismann H, Wersching H, Nagel M, *et al.* Establishing the bidirectional relationship between depression and subclinical arteriosclerosis - rationale, design, and characteristics of the BiDirect Study. *BMC Psychiatry* 2014; **14**: 1–9.

14 Wersching H, Berger K. Neue Kohorten: die BiDirect-Studie. *Bundesgesundheitsblatt - Gesundheitsforsch - Gesundheitsschutz* 2012; **55**: 822–3.

15 Inker LA, Schmid CH, Tighiouart H, *et al.* Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med* 2012; **367**: 20–9.

16 Gaiottino J, Norgren N, Dobson R, *et al.* Increased Neurofilament Light Chain Blood Levels in Neurodegenerative Neurological Diseases. *PLoS One* 2013; **8**: 1–9.

17 Altmann P, Leutmezer F, Zach H, *et al.* Serum neurofilament light chain withstands delayed freezing and repeated thawing. *Sci Rep* 2020; **10**: 1–8.

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 Rübsamen N, Willemse EAJ, Leppert D, *et al.* Neurofilament light as a blood-based biomarker for population-based epidemiological studies: handling of differences introduced by blood component and assay. *Submitted*.

21 Andermatt S, Pezold S, Carrin PC. Automated Segmentation of Multiple Sclerosis Lesions Using Multi-dimensional Gated Recurrent Units. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. 2017: 31–42.

22 Fartaria MJ, Kobera T, Granziera C, Cuadra MB. Longitudinal analysis of white matter and cortical lesions in multiple sclerosis. *NeuroImage Clin* 2019; **23**: 101938.

23 Vågberg M, Axelsson M, Birgander R, et al. Guidelines for the use of magnetic resonance imaging in diagnosing and monitoring the treatment of multiple sclerosis: recommendations of the Swedish Multiple Sclerosis Association and the Swedish Neuroradiological Society. Acta Neurol Scand 2017; 135: 17–24.

24 Akamine S, Marutani N, Kanayama D, *et al.* Renal function is associated with blood neurofilament light chain level in older adults. *Sci Rep* 2020; **10**: 1–7.

25 Korley FK, Goldstick J, Mastali M, *et al.* Serum NfL (Neurofilament Light Chain) Levels and Incident Stroke in Adults with Diabetes Mellitus. *Stroke* 2019; **50**: 1669–75.

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

27 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.
28 Rigby RA, Stasinopoulos DM. Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Stat Med* 2004; **23**: 3053–76.
29 Rigby RA, Stasinopoulos DM, Lane PW. Generalized additive models for location, scale and shape. *J R Stat Soc Ser C Appl Stat* 2005; **54**: 507–54.
30 Curtis A, Smith T, Ziganshin B, Elefteriades J. The Mystery of the Z-Score. *Aorta* 2016; **04**: 124–30.

3.2 Serum Glial Fibrillary Acidic Protein Compared With Neurofilament Light Chain as a Biomarker for Disease Progression in Multiple Sclerosis

JAMA Neurology | Original Investigation

Serum Glial Fibrillary Acidic Protein Compared With Neurofilament Light Chain as a Biomarker for Disease Progression in Multiple Sclerosis

Stephanie Meier, MSc; Eline A.J. Willemse, PhD; Sabine Schaedelin, MSc; Johanna Oechtering, MD; Johannes Lorscheider, MD; Lester Melie-Garcia, PhD; Alessandro Cagol, MD; Muhamed Barakovic, PhD; Riccardo Galbusera, MD; Suvitha Subramaniam, MSc; Christian Barro, MD; Ahmed Abdelhak, MD; Simon Thebault, MD; Lutz Achtnichts, MD; Patrice Lalive, MD; Stefanie Müller, MD; Caroline Pot, MD; Anke Salmen, MD; Giulio Disanto, MD; Chiara Zecca, MD; Marcus D'Souza, MD; Annette Orleth, PhD; Michael Khalil, MD; Arabella Buchmann, MSc; Renaud Du Pasquier, MD; Özgür Yaldizli, MD; Tobias Derfuss, MD; Klaus Berger, MD; Marco Hermesdorf, MD; Heinz Wiendl, MD; Fredrik Piehl, MD; Marco Battaglini, PhD; Urs Fischer, MD; Ludwig Kappos, MD; Claudio Gobbi, MD; Cristina Granziera, MD; Claire Bridel, MD; David Leppert, MD; Aleksandra Maleska Maceski, MSc; Pascal Benkert, PhD; Jens Kuhle, MD

IMPORTANCE There is a lack of validated biomarkers for disability progression independent of relapse activity (PIRA) in multiple sclerosis (MS).

OBJECTIVE To determine how serum glial fibrillary acidic protein (sGFAP) and serum neurofilament light chain (sNfL) correlate with features of disease progression vs acute focal inflammation in MS and how they can prognosticate disease progression.

DESIGN, SETTING, AND PARTICIPANTS Data were acquired in the longitudinal Swiss MS cohort (SMSC; a consortium of tertiary referral hospitals) from January 1, 2012, to October 20, 2022. The SMSC is a prospective, multicenter study performed in 8 centers in Switzerland. For this nested study, participants had to meet the following inclusion criteria: cohort 1, patients with MS and either stable or worsening disability and similar baseline Expanded Disability Status Scale scores with no relapses during the entire follow-up; and cohort 2, all SMSC study patients who had initiated and continued B-cell–depleting treatment (ie, ocrelizumab or rituximab).

EXPOSURES Patients received standard immunotherapies or were untreated.

MAIN OUTCOMES AND MEASURES In cohort 1, sGFAP and sNfL levels were measured longitudinally using Simoa assays. Healthy control samples served as the reference. In cohort 2, sGFAP and sNfL levels were determined cross-sectionally.

RESULTS This study included a total of 355 patients (103 [29.0%] in cohort 1: median [IQR] age, 42.1 [33.2-47.6] years; 73 female patients [70.9%]; and 252 [71.0%] in cohort 2: median [IQR] age, 44.3 [33.3-54.7] years; 156 female patients [61.9%]) and 259 healthy controls with a median [IQR] age of 44.3 [36.3-52.3] years and 177 female individuals (68.3%). sGFAP levels in controls increased as a function of age (1.5% per year; $P < .001$), were inversely correlated with BMI (−1.1% per BMI unit; $P = .01$), and were 14.9% higher in women than in men ($P = .004$). In cohort 1, patients with worsening progressive MS showed 50.9% higher sGFAP levels compared with those with stable MS after additional sNfL adjustment, whereas the 25% increase of sNfL disappeared after additional sGFAP adjustment. Higher sGFAP at baseline was associated with accelerated gray matter brain volume loss (per doubling: 0.24% per year; $P < .001$) but not white matter loss. sGFAP levels remained unchanged during disease exacerbations vs remission phases. In cohort 2, median (IQR) sGFAP *z* scores were higher in patients developing future confirmed disability worsening compared with those with stable disability (1.94 [0.36-2.23] vs 0.71 [−0.13 to 1.73]; $P = .002$); this was not significant for sNfL. However, the combined elevation of *z* scores of both biomarkers resulted in a 4- to 5-fold increased risk of confirmed disability worsening (hazard ratio [HR], 4.09; 95% CI, 2.04-8.18; $P < .001$) and PIRA (HR, 4.71; 95% CI, 2.05-9.77; $P < .001$).

CONCLUSIONS AND RELEVANCE Results of this cohort study suggest that sGFAP is a prognostic biomarker for future PIRA and revealed its complementary potential next to sNfL. sGFAP may serve as a useful biomarker for disease progression in MS in individual patient management and drug development.

JAMA Neurol. doi:10.1001/jamaneurol.2022.5250
Published online February 6, 2023.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Jens Kuhle, MD, PhD, Department of Neurology, University Hospital and University of Basel, Petersgraben 4, CH-4031, Basel, Switzerland (jens.kuhle@usb.ch).

The pathogenesis of multiple sclerosis (MS) involves both adaptive and innate immune disease mechanisms. The former is associated with recurring episodes of acute neurologic symptoms, relapses, and formation of localized lesions in the brain and spinal cord caused by invasion of blood-derived immune cells. In contrast, the latter has been suggested to drive more diffuse inflammation and neurodegeneration, also called smoldering MS,[1] that clinically presents as disease progression. Although high-efficacy therapies, such as B-cell–depleting treatment (BCDT), result in almost complete suppression of focal lesion formation, their effectiveness for preventing development of long-term disability is modest.[2,3] This therapeutic gap is mirrored by a diagnostic unmet need to assess progression. Serum neurofilament light chain (sNfL) is now well established as therapy response marker in active disease[4-6]; however, its capacity to reflect concurrent, or to predict progression, especially when acute inflammatory disease activity is suppressed by high efficacy therapies, is still under debate.[4,7-12]

Glial fibrillary acidic protein (GFAP) is an intermediate filament of astrocytes, equivalent to NfL in neurons, and has been proposed as a biomarker to identify present disease progression and to prognosticate future progression in MS.[13-18] Early studies measuring GFAP levels in the cerebrospinal fluid (CSF) of patients with MS found a correlation with neurologic disability in subsequent years; however, this was not the case for NfL levels.[14] Furthermore, high CSF GFAP levels were associated with faster progression to an Expanded Disability Status Scale (EDSS) score of 3 and 6,[19] and levels were higher in primary progressive MS than in relapsing-remitting MS (RRMS).[14,20,21] Moreover, there is also evidence of increased GFAP levels in the CSF of patients with progressive MS who had no recent relapses, showing the potential of GFAP levels for measuring pure progression.[13] In contrast, although NfL was a sensitive indicator of neuroaxonal injury during acute disease activity, ie, lesion formation and relapses, CSF levels of GFAP remained unaffected in this state.[20,22]

Based on different methodological approaches in 2 independent patient cohorts followed in the Swiss MS Cohort (SMSC), this study attempted a direct comparison of sGFAP and sNfL levels: how they reflect acute disease activity vs the identification and prognostication of future disease progression and whether their combination provides added value. In cohort 1, we (1) measured their levels in patients who either remained clinically stable or continued to accumulate more disability over time and (2) compared how they are impacted by acute inflammation in a cohort of patients with relapsing forms of MS (RMS). Cohort 2 comprised patients with MS receiving BCDT as a model of optimal suppression of acute disease activity to evaluate how sNfL and sGFAP levels, alone and in combination, are prognostic for future disability worsening and progression independent of relapse activity (PIRA).

Key Points

Question Are serum glial fibrillary acidic protein (sGFAP) and/or neurofilament light chain (sNfL) concentrations associated with and prognostic for disease progression in patients with multiple sclerosis?

Findings In this cohort study of 355 patients and 259 healthy controls (contributing 737 and 485 serum samples, respectively), elevated sGFAP *z* scores (corrected for confounding factors age, sex, and body mass index) identified current disease progression and were associated with future disease progression but not with acute inflammation. In addition, the association of sNfL levels with progression was less pronounced, whereas sNfL levels were strongly increased during relapse activity.

Meaning Results suggest that sGFAP is more strongly associated than sNfL with disease progression in MS, a finding that has clinical implications for patient management and development of novel drugs.

Methods

Study Design and Patients With MS

This cohort study, conducted from January 1, 2012, to October 22, 2022, was approved by the ethics committees of all participating centers. Patients in both cohorts provided written informed consent. A description of the SMSC and standard definitions are available in the eMethods in Supplement 1. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Cohort 1

Three groups of patients with MS with extreme phenotypes were compared: patients with either stable MS (stMS) or worsening disability[23] had similar baseline EDSS scores and no relapses during the entire follow-up; the focal inflammation group consisted of patients with relapsing MS from whom serum samples were acquired both during active disease phase (relapse and/or contrast-enhancing brain lesions) and remission. Patients with worsening progressive MS or stMS were matched for age, disease duration, EDSS scores, and T2-weighted lesion volume at baseline. Patients with worsening progressive MS presented with at least 1 PIRA event during follow-up. Further details are available in the eMethods in Supplement 1. Cohort 1 included patients of only White race and ethnicity. Other race and ethnic subgroups were too small for meaningful analysis.

Cohort 2

We included all SMSC patients who had initiated and continued BCDT (ocrelizumab or rituximab). sNfL and sGFAP levels were measured in the first sample available 8 months or more after treatment start (median [IQR], 12.2 [10.7-16.8] months). We included patients with RRMS and progressive MS. PIRA was defined by the occurrence of confirmed disability worsening (CDW) events in the absence of relapses between the visit defining baseline of the EDSS worsening event until its confirmation visit at least 6 months later. All other CDW events were defined to be relapse-associated worsening (RAW) events. Cohort 2 included patients of only White race and Hispanic ethnicity. Other race and ethnic subgroups were too small for meaningful analysis.

Healthy Controls

Blood samples from healthy controls (HCs) in the Genome-Wide Association Study of Multiple Sclerosis (GeneMSA[24,25]) were collected at the University Hospital Basel between July 7, 2004, and May 29, 2007. A family history or current diagnosis of MS, as well as other reported ongoing relevant illnesses (eg, diabetes, arterial hypertension), were considered exclusionary for this group.

sGFAP and sNfL Measurements

Blood samples were collected within 8 days from the clinical visit and stored at −80 °C following standardized procedures.[26] sGFAP and sNfL concentrations were measured in duplicate with the ultrasensitive single molecule array (Simoa) technology (Quanterix). In cohort 1, samples were measured using the singleplex Simoa GFAP Discovery Kit on the HD-X analyzer according to the manufacturer's instructions. sNfL levels had been measured in a previous study[4] using the Simoa Nf-Light kit. In cohort 2, samples were measured using the Neurology 2-plex B assay according to manufacturer's instructions (eFigure 1 in Supplement 1). Further details, including information on magnetic resonance imaging (MRI) assessment methods, are in the eMethods in Supplement 1.

Statistical Analysis

In HCs, the association between log-transformed biomarker concentrations as a dependent variable and age, sex, and body mass index (BMI; calculated as weight in kilograms divided by height in meters squared) as independent variables were analyzed using mixed models with a random intercept for person. In analogy with age- and BMI-adjusted sNfL reference values,[4] we calculated sGFAP *z* scores additionally adjusted for sex. A more detailed description of the statistical analysis is available in the eMethods in Supplement 1.

Cohort 1

Comparison of sGFAP and sNfL levels in stMS/worsening progressive MS and RMS cohorts vs HCs was performed using a linear mixed model with log-transformed sGFAP or sNfL levels as the dependent variable and age, BMI, sex, and phenotype group (stMS, worsening progressive MS; RMS in either remission or active disease state) as independent variables as well as a random intercept for the person to account for the repeated nature of the data. To assess the association of disease progression with sGFAP or sNfL levels (individual biomarkers as dependent variables), univariable and multivariable models with stMS vs worsening progressive MS status as well as age, sex, BMI, follow-up time, disease duration, disease-modifying treatment, and EDSS scores as independent variables were used. To evaluate the independent association between disease progression or active disease status and sGFAP or sNfL levels that is not explained by the other biomarker, the respective $\log_2$-transformed marker was additionally added to these models. The within-person variation of sGFAP or sNfL levels was assessed by the intraclass correlation coefficient (ICC) with 95% CI obtained by bootstrapping. Atrophy rates per year in the combined stMS and worsening progressive MS cohort were assessed with a linear mixed model. The associations between biomarker levels and gray matter volume and white matter volume loss were modeled using interaction terms between $\log_2$-transformed baseline sGFAP and sNfL levels, and follow-up time and estimates express the change in annualized atrophy rates per doubling in biomarker concentration. To compare the prognostic power of baseline sGFAP and sNfL levels for PIRA, univariable and multivariable Cox regression models were performed in the combined stMS and worsening progressive MS cohort.

Cohort 2

Biomarker levels in patients with and without later CDW were visualized using box plots and were considered increased compared with HC when being significantly above $z = 0$ in the univariate Wilcoxon signed rank tests (a *z* score of 0, corresponding to the 50th percentile, indicates the physiologic mean level of HC[4]). A cross-sectional analysis was performed using linear models with individual biomarker *z* score as the dependent variable and demographic and clinical variables as predictors. The association between biomarker levels and time to CDW was investigated using Kaplan-Meier curves and Cox regression models. Receiver operating characteristics (ROC) analyses were performed to identify optimal cut points for sGFAP and sNfL *z* score values to dichotomize the respective biomarker levels in high and low groups to prognosticate CDW. The performance of a composite of both biomarkers in prognosticating CDW was investigated by categorizing patients into 4 groups according to high and low levels for each biomarker, using the constellation of low sGFAP/low sNfL as a reference.

Sensitivity analyses were performed using only CDW due to PIRA (ie, excluding CDW due to RAW). A 2-sided *P* value ≤ .05 was considered statistically significant. Analyses were performed in R, version 4.2.0 (R Project for Statistical Computing).

Results

Serum GFAP and sNfL Concentrations in HCs

This study included a total of 355 patients (103 [29.0%] in cohort 1: median [IQR] age, 42.1 [33.2-47.6] years; 30 male individuals [29.1%]; 73 female individuals [70.9%] and 252 [71.0%] in cohort 2: median [IQR] age, 44.3 [33.3-54.7] years; 96 male individuals [38.1%]; 156 female individuals [61.9%]). The cohort of 259 HCs (485 samples) included 177 female individuals (68.3%) and had a median (IQR) age at baseline of 44.3 (36.3-52.3) years. sGFAP levels increased with age (1.5% per year; $P < .001$) (eFigures 2 and 3 in Supplement 1) and were inversely correlated with BMI (1.1% decrease per BMI unit, estimate 0.989; 95% CI, 0.979-0.998; $P = .01$). Across all ages, levels were 14.9% higher in women than in men ($P = .004$). sNfL levels increased by 2.5% per year of age and decreased by 2.2% per unit BMI (estimate 0.978; 95% CI, 0.969-0.986; $P < .001$) in both sexes. sGFAP and sNfL levels were moderately correlated at baseline (Spearman $\rho = 0.47$; $P < .001$).

Cohort 1

At baseline, patients with stMS and worsening progressive MS showed little difference in demographic, clinical, or MRI data,

Table 1. Patient Characteristics of Stable, Worsening Progressive Multiple Sclerosis (MS) and Relapsing MS Sampled During Remission and Active Disease

	MS, No. (%)			No. (%)		
Variable	Stable	Worsening progressive	*P* value	Remission	Active	*P* value
No. of patients	19	18	NA	66		NA
Samples, No.	169	184		66	66	
No. of samples per patient	9 (8-10)	10 (9-12.5)	.10	NA	NA	NA
Follow-up time, median (IQR) [range], y	7.1 (5.7-8.0) [4.1-9.0]	6.5 (5.2-7.7) [2.7-8.5]	.40	NA	NA	NA
Sex						
Female	12 (63.2)	11 (61.1)	<.99	50 (75.8)		NA
Male	7 (36.8)	7 (38.9)		16 (24.2)		
Age, median (IQR), y	44.2 (39.5-49.2)	43.8 (40.9-53.8)	.78	40.6 (30.2-46.4)	39.9 (29.2-45.4)	.62
Disease category at study entry						
RRMS	18 (94.7)	10 (55.6)	.02	62 (93.9)	62 (93.9)	.80
Progressive MS	1 (5.3)	8 (44.4)		4 (6.1)	4 (6.1)	
EDSS score, median (IQR)	3.0 (2.5-3.8)	4.0 (3.1-4.4)	.07	2.0 (1.5-3.0)	2.0 (2.0-3.0)	.25
Disease duration, median (IQR), y	9.4 (6.3-20.1)	13.70 (7.8-18.7)	.43	7.8 (3.8-14.7)	7.5 (3.4-14.1)	.50
DMT			.09			.001
Untreated	3 (15.8)	7 (38.9)		8 (12.1)	23 (34.8)	
Platform	5 (26.3)	0 (0)		4 (7.6)	9 (13.6)	
Oral	6 (31.6)	6 (33.3)		40 (60.6)	31 (47.0)	
Monoclonal antibody therapies	5 (26.3)	5 (27.8)		13 (19.7)	3 (4.5)	
Relapse[a]	NA	NA	NA	0 (0)	36 (54.5)	NA
Time since last relapse, median (IQR), d	NA	NA	NA	NA	16.0 (4.8-22.5)	NA
T2w lesion volume, median (IQR), mL	10.9 (2.7-19.7)	16.3 (12.8-44.7)	.21	5.2 (2.0-14.6)	5.9 (2.6-17.9)	0.48
EDSS score at last sampling, median (IQR)	2.5 (2.0-3.8)	6.0 (5.6-6.9)	<.001	NA	NA	NA
No. of PIRA events						
0	19 (100)	0 (0)	<.001	NA	NA	NA
1	0 (0)	6 (33.3)				
2	0 (0)	8 (44.4)				
3	0 (0)	4 (22.2)				
DMT at last visit						
Untreated	1 (5.3)	4 (22.2)	<.001	NA	NA	NA
Platform	4 (21.1)	0 (0)				
Orals	11 (57.9)	0 (0)				
mAB	3 (15.8)	14 (77.8)				
CEL at sample	1 (0.8)	2 (1.9)	.83	NA	NA	NA
New/enlarging T2w lesion at sample	13 (7.7)	20 (10.9)	.41			
Presence of CEL	NA	NA	NA	0 (0)	30 (45.5)	NA
Relapse and CEL				0 (0)	9 (13.6)	NA
T2w lesion volume, median (IQR), mL				5.2 (2.0-14.6)	5.9 (2.6-17.9)	.48

Abbreviations: CEL, contrast-enhancing lesion; DMT, disease-modifying treatment; EDSS, Expanded Disability Status Scale; mAB, monoclonal antibody therapies; MS, multiple sclerosis; NA, not applicable; PIRA, progression independent of relapse activity; RRMS, relapsing-remitting MS; w, weighted.

[a] Within 30 days.

except that treatment with monoclonal antibodies at last follow-up was more frequent in worsening progressive MS; the EDSS score remained stable in stMS (decreased from 3.0 to 2.5 at 7.1 years median follow-up), whereas in worsening progressive MS, it increased from a score of 4.0 to 6.0 with a median follow-up of 6.5 years (**Table 1**; eFigure 4 in Supplement 1). Worsening progressive MS showed more total brain volume loss (0.28% per year) vs stMS (estimate 0.997; 95% CI, 0.996-0.998; $P < .001$) (eFigure 5 in Supplement 1). Patients with RMS were more frequently untreated in active vs remission state (Table 1).

Comparison of sGFAP and sNfL Concentrations Between Patients and HCs

sGFAP levels were highest in worsening progressive MS (103.0 pg/mL with a 77% increase vs 51.8 pg/mL in HCs; $P < .001$), followed by RMS in active disease (59.1 pg/mL; $P < .001$), RMS during remission (52.9 pg/mL; $P = .01$), and

patients with stMS (63.2 pg/mL; $P = .12$) (eTable 1, eFigure 6 in Supplement 1). Conversely, sNfL levels were highest in active RMS (10.2 pg/mL, namely 98.6% as per adjusted estimate higher than in HCs, 6.3 pg/mL; $P < .001$), followed by worsening progressive MS (10.9 pg/mL; $P < .001$), stMS (7.2 pg/mL; $P = .03$), and RMS in remission (6.7 pg/mL; $P < .001$).

Serum GFAP and sNfL Levels in Worsening Progressive MS vs stMS

sGFAP and sNfL concentrations were increased by 64.2% and 42.2%, respectively, in worsening progressive MS vs stMS (**Table 2**, model 1, **Figure 1**A). After multivariable adjustment, these differences were 57.5% and 24.8%, respectively (Table 2, model 2, Figure 1B), also after additional correction for sNfL (50.9% increase in worsening progressive MS vs stMS), whereas the 25% increase of sNfL levels disappeared after additional sGFAP-level adjustment (Table 2, model 3, Figure 1C). Additional sensitivity analyses adjusting for T2-weighted lesion volume, and number of new and enlarged and contrast-enhancing brain lesions confirmed these results and showed comparably increased sGFAP levels in worsening progressive MS vs stMS (eTable 2 in Supplement 1). sGFAP levels in the worsening progressive MS cohort showed less within-person variability over time (ICC: estimate, 0.91; 95% CI, 0.83-0.94, ie, 91% of the variation in sGFAP levels is explained by variation between patients), whereas for sNfL ICC was 0.80 (95% CI, 0.72-0.85; difference, 11%; 95% CI, 2%-19%; $P = .02$).

sGFAP and sNfL Levels in RMS During Active Disease and Remission

sNfL concentrations were 58.4% increased in active disease vs remission, whereas this difference was 7.3% for sGFAP levels (eTable 3 in Supplement 1, model 1). After adjustment for potential confounders, these differences were 53.2% and 4.8%, respectively (eTable 3 in Supplement 1, model 2). Additional correction for sGFAP levels did not influence the association of focal inflammation status with sNfL levels (50.6% increase in active vs remission state), whereas association with sGFAP levels remained insignificant (eTable 3 in Supplement 1, model 3).

Association of Baseline sGFAP and sNfL Levels With Brain Volume Loss and PIRA

Each doubling of baseline sGFAP levels was associated with an additional loss of gray matter volume (−0.24% per year; 95% CI, −0.35% to −0.12%; $P < .001$) but not white matter volume (0.05%; 95% CI, −0.09% to 0.18%; $P = .48$), whereas doubling of baseline sNfL levels was associated with an additional loss of white matter volume (−0.26%; 95% CI, −0.38% to −0.15%; $P < .001$) but not gray matter volume (−0.01%; 95% CI, −0.11 to 0.09; $P = .78$) (eTable 4, eFigure 7 in Supplement 1). Baseline values of sGFAP levels had a better prognostic capacity for future PIRA (HR per doubling, 3.88; 95% CI, 1.69-8.86; $P = .001$; ie, an almost 4-fold risk of PIRA by doubling of baseline sGFAP concentration) than sNfL levels (HR, 1.77; 95% CI, 1.11-2.83; $P = .02$). In a combined model, with additional adjustment for age, sex, BMI, and disease duration, these findings were confirmed: sGFAP levels (HR, 3.63; 95% CI, 1.46-9.04; $P = .006$) and sNfL levels (HR, 1.90; 95% CI, 0.86-4.19; $P = .11$).

Cohort 2

Cohort Characteristics

We included 252 patients receiving BCDT who were relapse-free in the 6 months prior to sampling (ie, baseline). The majority of patients presented with RRMS (181 of 252 [71.8%]), whereas the remaining had progressive MS (34 [13.5%] secondary progressive MS; 37 [14.7%] primary progressive MS). A total of 43 of 252 (17.1%) experienced CDW during follow-up, of which 39 (90.7%) were due to PIRA and 4 (9.3%) due to RAW (eTable 5 in Supplement 1).

sGFAP and sNfL Levels and Development of Future CDW

In patients with MS overall, sGFAP levels were strongly increased compared with those of HCs (z score = 0) with a median (IQR) of 0.82 (−0.05 to 1.95) z score units above normal ($P < .001$), whereas the increase of sNfL levels was less pronounced (0.50; IQR, −0.25 to 1.32; $P < .001$). Development of CDW was associated with a higher sGFAP z score 12.2 (IQR, 10.7-16.8) months after BCDT start than in patients without future CDW (1.94; IQR, 0.36-2.23 vs 0.71; IQR, −0.13 to 1.73) (**Figure 2**A). Although sNfL z score were less but still significantly increased vs that in HCs, the difference between patients with vs those without CDW development was not significant (Figure 2B). This pattern was similar when RAW events were excluded (sGFAP levels: PIRA, 1.98; IQR, 0.33-2.27 vs no PIRA, 0.71; IQR, −0.11 to 1.73; $P = .003$; sNfL levels: PIRA, 1.09; IQR, 0.14-1.49 vs no PIRA, 0.44; IQR, −0.25 to 1.23; $P = .04$).

Next, we explored which demographic and disease-related variables were associated with increased biomarker levels in patients receiving BCDT compared with HCs using multivariable models (ie, using biomarker z scores as dependent variable (eTable 6 in Supplement 1). Models on the absolute sGFAP and sNfL concentrations are included in eTable 7 in Supplement 1. The model for sGFAP z score explained 13.3% of the variance and was driven by female sex, younger age, higher EDSS, and whether the patient developed CDW while receiving BCDT (CDW status in eTable 6 in Supplement 1). The same model with sNfL z score as the outcome explained 1.8% of its variance. Specifically, only sGFAP z scores, but not those of sNfL, were linked to the EDSS score and future CDW. Again, findings were similar in the PIRA only set (not shown).

Prognostic Value of sGFAP and sNfL Levels for Future CDW

Time-to-event analyses showed that 1 sGFAP z-score unit increase led to a 1.36-fold (HR, 1.36; 95% CI, 1.09-1.69; $P = .006$) increased risk of CDW (after correction for covariates: HR, 1.32; 95% CI, 1.06-1.66; $P = .01$). For sNfL z score, a numerically higher risk was found (HR, 1.25; 95% CI, 0.95-1.65; $P = .11$; after correction: HR, 1.27; 95% CI, 0.95-1.71; $P = .11$). When combining both sGFAP and sNfL z scores in 1 model, sGFAP was associated with disease worsening: HR, 1.34 (95% CI, 1.03-1.73; $P = .03$), but not sNfL (HR, 1.04; 95% CI, 0.75-1.43; $P = .82$).

Next, we used different z score cut points to see whether their increase was associated quantitatively to the risk of CDW. sGFAP z score cut points of 1, 1.5, and 2 led to gradually increasing CDW hazard ratios ranging from 2.1 to 3.4

Table 2. Multivariable Mixed Linear Models Investigating the Association of Worsening Status (Stable Multiple Sclerosis [MS] vs Worsening Progressive MS) With Log-Transformed Serum Glial Fibrillary Acidic Protein (sGFAP) and Serum Neurofilament Light Chain (sNfL) Levels

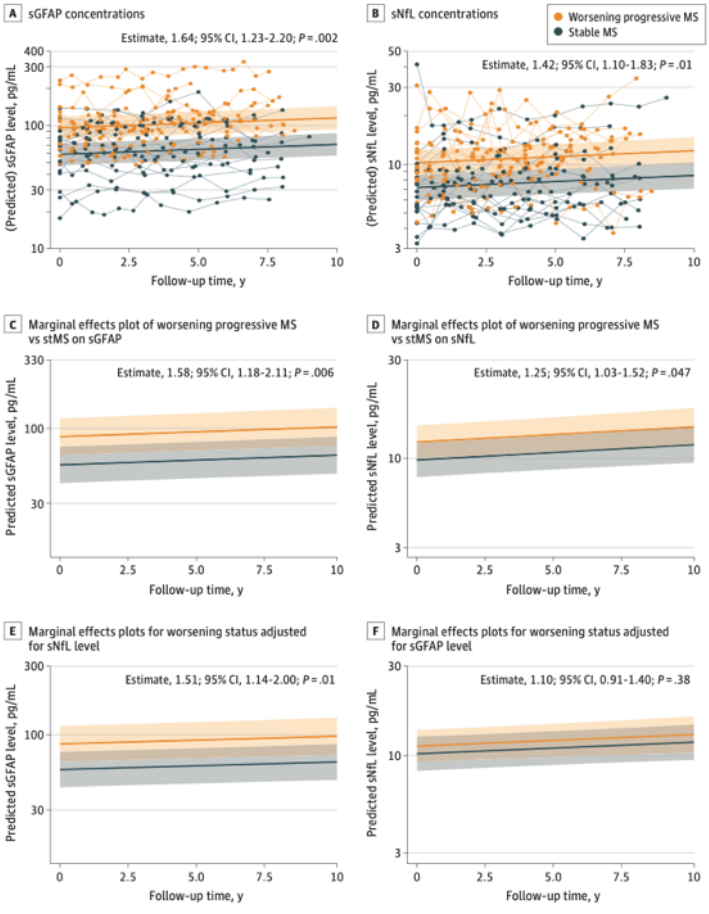
Model	Sample, No.	sGFAP, median (IQR), pg/mL	Estimate (95% CI)[a]	*P* value	sNfL, median (IQR), pg/mL	Estimate (95% CI)	*P* value
Model 1: simple							
Follow-up time	NA	NA	1.019 (1.011-1.026)	<.001	NA	1.017 (1.008-1.027)	<.001
Progression							
Stable MS	169	63.2 (43.4-90.7)	NA	NA	7.1 (5.4-9.4)	NA	NA
Worsening progressive MS	184	103.0 (81.3-132.5)	1.642 (1.226-2.199)	.002	10.9 (8.2-13.9)	1.422 (1.104-1.831)	.01
Model 2: multivariable							
Age at baseline	NA	NA	1.008 (0.993-1.023)	.35	NA	1.019 (1.009-1.029)	.002
Follow-up time	NA	NA	1.016 (1.007-1.025)	<.001	NA	1.019 (1.008-1.030)	.001
Sex							
Female	224	87.7 (57.3-109.7)	1.026 (0.764-1.378)	.87	8.4 (6.3-10.9)	0.875 (0.725-1.058)	.21
Male	129	84.3 (57.7-121.1)	NA		11.8 (5.8-16.7)	NA	
BMI[b]	NA	NA	0.991 (0.973-1.008)	.32	NA	0.969 (0.953-0.985)	<.001
Disease duration at baseline	NA	NA	1.002 (0.985-1.018)	.86	NA	1.005 (0.994-1.016)	.40
DMT							
Untreated	48	97.4 (63.8-112.9)	NA	NA	11.7 (8.7-16.4)	NA	NA
Platform	40	68.6 (57.7-90.4)	1.191 (1.048-1.356)	.009	9.7 (6.3-17.4)	0.956 (0.821-1.142)	.59
Orals	118	74.7 (39.4-97.4)	1.032 (0.933-1.139)	.54	7.7 (5.3-9.5)	0.921 (0.811-1.039)	.20
mAB	147	103.6 (68.4-136.5)	1.080 (0.997-1.171)	.06	9.4 (6.8-12.8)	0.938 (0.842-1.035)	.22
EDSS score	NA	NA	1.011 (0.982-1.041)	.46	NA	1.002 (0.969-1.039)	.92
Progression							
Stable MS	169	63.2 (43.4-90.7)	NA	NA	7.1 (5.4-9.4)	NA	NA
Worsening progressive MS	184	103.0 (81.3-132.5)	1.575 (1.178-2.106	.006	10.9 (8.2-13.9)	1.248 (1.024-1.521)	.05
Model 3: plus sNfL/sGFAP							
Age at baseline	NA	NA	1.004 (0.990-1.0.19)	.59	NA	1.016 (1.007-1.026)	.004
Follow-up time	NA	NA	1.012 (1.004-1.021)	.005	NA	1.014 (1.003-1.025)	.01
Sex							
Female	224	87.7 (57.3-109.7)	1.053 (0.792-1.400)	.74	8.4 (6.3-10.9)	0.868 (0.725-1.040)	.17
Male	129	84.3 (57.7-121.1)	NA	NA	11.8 (5.8-16.7)	NA	NA
BMI[b]	NA	NA	0.996 (0.979-1.013)	.66	NA	0.973 (0.958-0.989)	.002
Disease duration at baseline	NA	NA	1.001 (0.985-1.017)	.94	NA	1.005 (0.994-1.1015)	.42
DMT							
Untreated	48	97.4 (63.8-112.9)	NA	NA	11.7 (8.7-16.4)	NA	NA
Platform	40	68.6 (57.7-90.4)	1.214 (1.072-1.377)	.003	9.7 (6.3-17.4)	0.907 (0.782-1.082)	.24
Oral	118	74.7 (39.4-97.4)	1.045 (0.948-1.151)	.37	7.7 (5.3-9.5)	0.917 (0.812-1.032)	.17
mAB	147	103.6 (68.4-136.5)	1.090 (1.008-1.179)	.03	9.4 (6.8-12.8)	0.918 (0.827-1.010)	.10
EDSS score	NA	NA	1.012 (0.984-1.041)	.41	NA	0.999 (0.968-1.036)	.98
sNfL per doubling, pg/mL	NA	NA	1.141 (1.079-1.207)	<.001	NA	NA	NA
sGFAP per doubling, pg/mL	NA	NA	NA	NA	NA	1.217 (1.120-1.315)	<.001
Progression							
Stable MS	169	63.2 (43.4-90.7)	NA	NA	7.1 (5.4-9.4)	NA	NA
Worsening progressive MS	184	103.0 (81.3-132.5)	1.509 (1.139-1.998)	.01	10.9 (8.2-13.9)	1.099 (0.905-1.339)	.38

Abbreviations: BMI, body mass index; DMT, disease modifying treatment; EDSS, Expanded Disability Status Scale; mAB, monoclonal antibody therapies; NA, not applicable; sGFAP, serum glial fibrillary acidic protein; sNfL, serum neurofilament light chain.

[a] Estimates are back transformed and represent multiplicative effects.

[b] Calculated as weight in kilograms divided by height in meters squared.

Figure 1. Serum Glial Fibrillary Acidic Protein (GFAP) and Serum Neurofilament Light Chain (sNfL) in Worsening Progressive Multiple Sclerosis (MS) and Stable MS



Concentrations of sGFAP (A) and sNfL (B) in worsening progressive MS vs stable MS (stMS) over follow-up time. sGFAP and sNfL concentrations were increased by 64.2% and 42.2%, respectively, in worsening progressive MS vs stMS. Thin lines connect longitudinal data points of individual patients; thick lines show the group regression lines from Table 2, model 1. Only the regression lines are predicted. Estimates including 95% CI and *P* value of differences between worsening progressive MS vs stMS are added to the plots (B). Marginal effects plots of multivariable mixed models showing the association of worsening progressive MS vs stMS with sGFAP (C) and sNfL levels (D) over follow-up time. After adjustment for potential confounders, sGFAP and sNfL concentrations were increased by 57.5% and 24.8%, respectively, in worsening progressive MS vs stMS (Table 2, model 2). Marginal effects plots for worsening status with additional adjustment for sNfL (E) and sGFAP (F). Additional correction for sNfL levels had a minor association with the difference of sGFAP levels between stMS vs worsening progressive MS status (50.9% increase); however, additional correction for sGFAP eliminated the association of progression status with sNfL levels (Table 2, model 3).

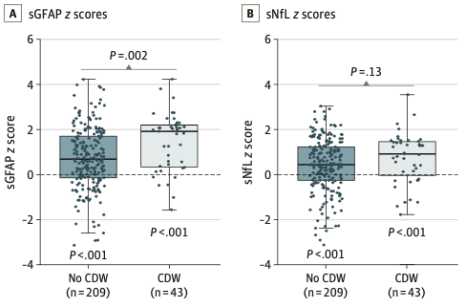
(eFigure 8A in Supplement 1). The associations were all significant for sGFAP, whereas for sNfL (eFigure 8B in Supplement 1), findings were less strong.

sGFAP and Prognostication of Worsening in a Combined Analysis of sNfL and sGFAP Levels

The risk of CDW in patients with high sGFAP levels (ie, *z* score >1.8, cutoff optimized in ROC analysis) compared with low sGFAP levels was 3-fold increased (HR, 3.25; 95% CI, 1.78-5.93; *P* < .001) in a time-to-event analysis. Patients with high sNfL levels (ie, *z* score >1.3) showed a 2-fold increased risk of future CDW (HR, 2.26; 95% CI, 1.24-4.14; *P* = .008) vs patients with low sNfL levels.

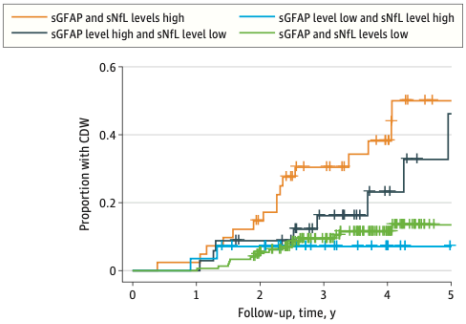
The combination of high sGFAP/high sNfL levels was associated with a 4-fold increased risk of worsening compared with low sGFAP/low sNfL levels (HR, 4.09; 95% CI, 2.04-8.18; *P* < .001 and PIRA only: HR, 4.71; 95% CI, 2.05-9.77; *P* < .001), that of high sGFAP/low sNfL levels showed slightly reduced association (**Figure 3**) (PIRA only: HR, 2.28; 95% CI, 0.92-5.64; *P* = .08). In contrast, the combination of low sGFAP/high sNfL levels did not show an increased risk for future CDW (PIRA only: HR, 1.17; 95% CI, 0.34-4.10; *P* = .80). The Kaplan-Meier analysis indicated that 4 years after initiation of treatment, 38% (95% CI, 20%-53%) of patients in the high sGFAP/high sNfL group will have CDW, compared with 23% (95% CI, 3%-39%) in the high sGFAP/low sNfL group, whereas this will be the case only

Figure 2. Serum Glial Fibrillary Acidic Protein (sGFAP) and Serum Neurofilament Light Chain (sNfL) *z* Scores in Patients With and Without Confirmed Disease Worsening During Follow-up While Receiving B-Cell–Depleting Therapy in Comparison to Healthy Controls



Box plot representation of sGFAP *z* scores (A) and sNfL *z* scores (B). Dashed lines indicate mean values in healthy controls (ie, *z* score = 0) and *P* values below indicate whether observed values differ from *z* scores 0 (Wilcoxon signed rank test). In patients with MS (without and with future confirmed disease worsening [CDW] development), sGFAP levels were increased compared with healthy controls (*z* scores healthy controls = 0; *P* < .001 for both), whereas the increase of sNfL was less pronounced (*P* < .001 for both). Development of CDW was associated with higher sGFAP *z* scores, which was not the case for sNfL.

Figure 3. Kaplan-Meier Curves Using Combined Biomarker Data to Predict Time to Confirmed Disease Worsening (CDW)



No. at risk						
sGFAP and sNfL levels high	41	40	33	22	13	4
sGFAP level high and sNfL level low	34	34	29	20	9	4
sGFAP level low and sNfL level high	28	27	23	16	9	5
sGFAP and sNfL levels low	149	149	132	94	51	19

Optimized cutoffs of serum glial fibrillary acidic protein (sGFAP) and serum neurofilament light chain (sNfL) *z* scores from receiver operating characteristic curve analysis, based on the Youden index, were used to dichotomize patient groups. High sGFAP/high sNfL levels were associated with a 4-fold (hazard ratio [HR], 4.09; 95% CI, 2.04-8.18; *P* < .001) increased risk of CDW compared with low sGFAP/low sNfL levels. The combination of high sGFAP/low sNfL levels showed a slightly reduced risk (HR, 2.32; 95% CI, 0.99-5.42; *P* = .05). The combination of low sGFAP/high sNfL levels, however, did not show an increased risk on CDW (HR, 1.03; 95% CI, 0.30-3.53; *P* = .97).

in 11% (95% CI, 6%-16%) if they fall into the low sGFAP/low sNfL group.

Discussion

The long-term course of disability in MS is driven by 2 partly independent pathomechanisms: focal lesional activity and brain-diffuse neurodegeneration.[23] sNfL has been established in recent years as a biomarker of ongoing neuronal damage in the course of the former process, whereas its association with progression as the clinical manifestation of the latter is relatively weaker.[12] The need for a biomarker that specifically reflects current and prognosticates future disability due to pure progression/PIRA has become urgent on the background that disability worsening often continues despite almost complete suppression of acute disease activity under high-efficacy therapies.[2,3] Increased CSF levels of GFAP have been proposed first by Axelsson et al[14] as a specific biomarker for progression. However, this finding was based on repetitive CSF analysis, which has precluded its entry into routine practice to close this diagnostic gap. Second, the relative contribution of lesional activity and RAW vs PIRA to the overall progression could not be determined in the mixed RRMS and progressive MS population studied. In this cohort study, we attempted to resolve the question about the mechanistic source of GFAP increase in MS by 2 orthogonal methodological approaches where relapse activity was absent in worsening progressive MS and stMS (cohort 1) or lesional activity and relapses were suppressed by BCDT in a mixed MS population (cohort 2). Current results suggest that increased levels of sGFAP were associated with pure progression/PIRA, although this biomarker is largely inert to acute disease activity. Higher baseline sNfL levels were prognostic for white matter volume loss, and baseline sGFAP specifically prognosticated gray matter loss, a previously proposed proxy for disease progression.[27,28] These findings from serum analysis are fully congruent with those of Axelsson et al in CSF.[14]

The increase of GFAP levels in the course of MS progression appears to result from astrocyte proliferation/activation and possibly injury.[29] This seems to be a brain-diffuse process, affecting mainly the normal-appearing white matter resulting in decreased diffusion tensor imaging derived measures.[16,30] In return, the minor increase of sNfL seen, eg, in patients with worsening progressive MS may result from continuous neuronal loss outside of acute lesion formation as part of the pathogenesis of progression due to subclinical neuroinflammation in chronic active lesions and in normal-appearing white matter.[31]

The association of sGFAP levels with future CDW and imaging features of progression is further supported by studies using serum samples.[16,17,32,33] However, these were incompletely controlled for confounding factors such as sex, age, and BMI, which resulted in significant overlap in GFAP levels across different MS groups and also controls, thus limiting clinical usefulness. Moreover, the comparison of raw biomarker concentrations vs *z* scores as an outcome highlights the advantage of

the latter in terms of pathogenetic relevance and ease of interpretability; covariates explained 29% of variation in raw sNfL concentration but only 1.8% of the variation in sNfL *z* score. For sGFAP levels, covariates similarly explained 25% of the variation in sGFAP concentrations but additionally also 13% in the variation of sGFAP *z* scores. Using corrected *z* scores, instead of absolute concentrations, to assess change of these biomarkers compared with normal values, thus substantially increase the sensitivity for detecting pathologic values, a prerequisite for its use in individual patients.

Important from a clinical perspective is that prognostication of future disability can be made based on a single GFAP measurement and from a biofluid (serum) that is easily accessible in clinical practice. A further aspect in our data set for the clinical use of sGFAP levels is the establishment of normative values of sGFAP that allow to define aberrations from physiological values corrected for confounding factors. Although age and BMI were known confounders, also based on the experience from the establishment of normative values for sNfL,[4] the 15% increase of sGFAP values in women vs men was an unexpected finding. Third, the combined evaluation of sNfL and sGFAP levels provides the highest predictive power for disability worsening, specifically in years 2 to 4, as a reflection of a comprehensive coverage of biological processes leading to disability worsening.

Limitations

This study has some limitations. One limitation is that we studied almost exclusively the effect of anti-CD20 antibodies as high-efficacy therapy but less so other types of disease-modifying treatments of this efficacy level (eg, natalizumab). Such evaluations will be necessary to expand on the limited data available whether disease-modifying treatment can lead to decrease of sGFAP levels[34] as a potential sign of attenuation of astrogliosis or pathological astrocyte activation. Second, the current normative database is derived from a relatively small cohort of HCs, where the impact of subclinical comorbidities could not be explored. A much larger cohort of persons, including those with other neurologic disease, may be needed to establish robust normal values for sGFAP levels.

Conclusions

In summary, the findings of this cohort study suggest that sGFAP levels may serve as a biomarker that reflects specifically chronic disease processes conveyed by astrocytes that manifest as pure progression/PIRA in MS. With this property, sGFAP levels are complementary to sNfL, whose levels are strongly associated with neuronal damage due to lesional disease activity.

ARTICLE INFORMATION

Accepted for Publication: November 23, 2022.

Published Online: February 6, 2023. doi:10.1001/jamaneurol.2022.5250

Author Affiliations: Department of Neurology, University Hospital and University of Basel, Basel, Switzerland (Meier, Willemse, Schaedelin, Oechtering, Lorscheider, Melie-Garcia, Cagol, Barakovic, Galbusera, Subramaniam, D'Souza, Orleth, Yaldizli, Derfuss, Fischer, Kappos, Granziera, Leppert, Maleska Maceski, Benkert, Kuhle); Multiple Sclerosis Centre, Departments of Biomedicine and Clinical Research, University Hospital and University of Basel, Basel, Switzerland (Meier, Willemse, Schaedelin, Oechtering, Lorscheider, Melie-Garcia, Cagol, Barakovic, Galbusera, Subramaniam, D'Souza, Orleth, Yaldizli, Derfuss, Fischer, Kappos, Granziera, Leppert, Maleska Maceski, Benkert, Kuhle); Research Center for Clinical Neuroimmunology and Neuroscience Basel, University Hospital and University of Basel, Switzerland (Meier, Willemse, Schaedelin, Oechtering, Lorscheider, Melie-Garcia, Cagol, Barakovic, Galbusera, Subramaniam, D'Souza, Orleth, Yaldizli, Derfuss, Fischer, Kappos, Granziera, Leppert, Maleska Maceski, Benkert, Kuhle); Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland (Willemse, Schaedelin, Subramaniam, Benkert); Translational Imaging in Neurology Basel, Department of Biomedical Engineering, Faculty of Medicine, University Hospital Basel and University of Basel, Basel, Switzerland (Melie-Garcia, Cagol, Barakovic, Galbusera, Yaldizli, Granziera); Department of Neurology, Harvard Medical School, Boston, Massachusetts (Barro); Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, San Francisco (Abdelhak); Department of Medicine and the Ottawa Hospital Research Institute, The University of Ottawa, Ottawa, Ontario, Canada (Thebault); Department of Neurology, Cantonal Hospital Aarau, Aarau, Switzerland (Achtnichts); Unit of Neuroimmunology, Division of Neurology, Department of Clinical Neurosciences, University Hospital of Geneva and Faculty of Medicine, Geneva, Switzerland (Lalive, Bridel); Department of Neurology, Cantonal Hospital St Gallen, St Gallen, Switzerland (Müller); Service of Neurology, Department of Clinical Neurosciences, Lausanne University Hospital, University of Lausanne, Lausanne, Switzerland (Pot); Department of Neurology, Bern University Hospital, University of Bern, Bern, Switzerland (Salmen); Multiple Sclerosis Center, Department of Neurology, Neurocenter of Southern Switzerland, Lugano, Switzerland (Disanto, Zecca, Gobbi); Faculty of Biomedical Sciences, Università della Svizzera Italiana, Lugano, Switzerland (Disanto, Zecca, Gobbi); Department of Neurology, Medical University of Graz, Graz, Austria (Khalil, Buchmann); Service of Neurology, Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland (Du Pasquier); Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany (Berger, Hermesdorf); Department of Neurology with Institute of Translational Neurology, University Hospital Münster, Münster, Germany (Wiendl); Department of Clinical Neuroscience, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden (Piehl); Center for Neurology, Academic Specialist Center, Stockholm Health Services, Stockholm, Sweden (Piehl); Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy (Battaglini).

Author Contributions: Dr Kuhle had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Ms Meier and Dr Willemse are considered co-first authors and contributed equally to this work. Ms Maleska Maceski and Drs Benkert and Kuhle are considered co-senior authors and contributed equally to this work.
Concept and design: Meier, Schaedelin, Abdelhak, Yaldizli, Derfuss, Leppert, Maleska Maceski, Benkert, Kuhle.
Acquisition, analysis, or interpretation of data: Meier, Willemse, Schaedelin, Oechtering, Lorscheider, Melie-Garcia, Cagol, Barakovic, Galbusera, Subramaniam, Barro, Abdelhak, Thebault, Achtnichts, Lalive, Müller, Pot, Salmen, Disanto, Zecca, D'Souza, Orleth, Khalil, Buchmann, Du Pasquier, Yaldizli, Berger, Hermesdorf, Wiendl, Piehl, Battaglini, Fischer, Kappos, Gobbi, Granziera, Bridel, Leppert, Maleska Maceski, Benkert, Kuhle.
Drafting of the manuscript: Meier, Willemse, Abdelhak, Thebault, Granziera, Leppert, Benkert, Kuhle.
Critical revision of the manuscript for important intellectual content: Meier, Schaedelin, Oechtering, Lorscheider, Melie-Garcia, Cagol, Barakovic, Galbusera, Subramaniam, Barro, Abdelhak, Achtnichts, Lalive, Müller, Pot, Salmen, Disanto, Zecca, D'Souza, Orleth, Khalil, Buchmann, Du Pasquier, Yaldizli, Derfuss, Berger, Hermesdorf, Wiendl, Piehl, Battaglini, Fischer, Kappos, Gobbi, Granziera, Bridel, Leppert, Maleska Maceski, Benkert, Kuhle.
Statistical analysis: Meier, Schaedelin, Barakovic, Abdelhak, Benkert, Kuhle.
Obtained funding: Berger, Leppert, Kuhle.
Administrative, technical, or material support: Oechtering, Cagol, Barro, Achtnichts, Lalive, Müller, Pot, Salmen, Zecca, D'Souza, Orleth, Buchmann, Derfuss, Berger, Hermesdorf, Wiendl, Maleska

Maceski, Kuhle.
Supervision: Derfuss, Kappos, Gobbi, Granziera, Leppert, Maleska Maceski, Kuhle.

Conflict of Interest Disclosures: Dr Oechtering reported receiving travel grants from Biogen Idec, Novartis, and Bayer; grants from Swiss MS Society; and serving on an advisory board by Roche outside the submitted work. Dr Lorscheider reported receiving grants from Novartis and Biogen; personal fees from Novartis, Roche, and Teva; serving on the advisory boards for Roche and Teva and receiving grants from Innosuisse Innovation Agency outside the submitted work. Dr Cagol reported receiving grant support from the Horizon 2020 Eurostar program. Dr Barro reported receiving travel support from Novartis and Teva. Dr Abdelhak reported receiving research grants from the German Multiple Sclerosis Society, Roche, and Denali Therapeutics outside the submitted work. Dr Achtnichts reported serving on the scientific advisory boards for Celgene, Novartis Pharmaceuticals, Merck, Biogen, Sanofi Genzyme, Roche and Bayer; receiving funding for travel and/or speaker honoraria from Celgene, Biogen, Sanofi Genzyme, Novartis, Merck Serono, Roche, Teva and the Swiss Multiple Sclerosis Society; and research support from Biogen, Sanofi Genzyme, and Novartis. Dr. Lalive reported receiving speaker honoraria 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, Novartis. Dr Müller reported receiving honoraria for travel, honoraria for lectures/consulting, and/or grants for studies from Almirall, Biogen, Celgene, Novartis, Teva, Merck Serono, Genzyme, Roche, and Bayer Schweiz. Dr Pot reported receiving consulting fees and/or travel compensation, used exclusively for research support, for activities with Biogen, Merck, Novartis, Roche, and Sanofi Genzyme. Dr Salmen reported receiving speaker honoraria and/or travel compensation for activities with Almirall Hermal GmbH, Biogen, Merck, Novartis, Roche, and Sanofi Genzyme; personal fees from Bristol Myers Squibb, CSL Behring, Novartis, and Roche; and grants from Baasch Medicus Foundation, Medical Faculty of the University of Bern, and the Swiss Multiple Sclerosis Society outside the submitted work. Dr Zecca reported receiving honoraria for speaking/consulting fees or grants from Abbvie, Almirall, Biogen Idec, Celgene, Janssen, Genzyme, Lilly, Merck Serono, Novartis, Roche, Sanofi, and Teva Pharma outside the submitted work. Dr Khalil reported receiving funding for travel and speaker honoraria from Bayer, Novartis, Merck, Biogen Idec, and Teva Pharmaceutical Industries; serving on scientific advisory boards for Biogen Idec, Merck, Roche, Novartis, Bristol Myers Squibb, and Gilead; and receiving grants from Teva Pharmaceutical Industries, Biogen, and Novartis. Dr Yaldizli reported receiving grants from European Committee for Treatment and Research in Multiple Sclerosis/European Magnetic Resonance Imaging in Multiple Sclerosis network, 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, Almirall, Biogen, and Novartis. Dr Derfuss reported receiving speaker fees, research support, travel support, and/or served on advisory boards, data safety monitoring boards, or Steering Committees of Actelion, Alexion, Celgene, Polyneuron, Novartis Pharma, Merck Serono, Sanofi, Biogen, Teva, Bayer-Schering, GeNeuro, Mitsubishi Pharma, MedDay, Roche, and Genzyme. Dr Berger reported receiving a grant from the German Ministry of Education and Research (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 patients with multiple sclerosis. Dr Wiendl reported receiving honoraria for acting as a member of scientific advisory boards from Abbvie, Alexion, Argenx, Bristol Myers Squibb/Celgene, Janssen, Merck, and Novartis; speaker honoraria and travel support from Alexion, Biogen, Bristol Myers Squibb, F. Hoffmann-La Roche Ltd, Genzyme, Merck, Neurodiem, Novartis, Roche Pharma AG, Teva Pharma, and WebMD Global; consultant fees from Abbvie, Actelion, Argenx, Biogen, Bristol Myers Squibb, EMD Serono, Fondazione Cariplo, Gossamer Bio, Idorsia, Immunic, Immunovant, Janssen, Lundbeck, Merck, NexGen, Novartis, PSI CRO, Roche, Sanofi, Swiss Multiple Sclerosis Society, UCB, and Worldwide Clinical Trials; and research funding from the German Ministry for Education and Research, Deutsche Forschungsgesellschaft, Deutsche Myasthenie Gesellschaft e.V., Alexion, Amicus Therapeutics Inc, Argenx, Biogen, CSL Behring, F. Hoffmann - La Roche, Genzyme, Merck KgaA, Novartis Pharma, Roche Pharma, and UCB Biopharma. Dr Piehl reported receiving grants from Merck KGaA and UCB; data safety monitoring board member fees from Chugai, Lundbeck, and Roche; and fees for the preparation of expert witness statement from Novartis outside the submitted work. Dr Fischer reported receiving grants from the Swiss National Science Foundation, Medtronic, Stryker, Penumbra, Phenox, Rapid Medical; consultant fees from Medtronic, CSL Behring, and Stryker; and fees for participation in an advisory board from Alexion/Portola and Boehringer Ingelheim; being a member of a clinical event committee of the COATING study (Phenox) and of the data and safety monitoring committee of the TITAN, LATE MT and IN EXTREMIS trials; and serving as vice presidency of the Swiss Neurological Society. Dr Kappos' employer (University Hospital Basel) has received and dedicated to research support fees for board membership, consultancy or speaking, or grants in the past 3 years from Abbvie, Actelion, Advancell, Allozyne, Auriga Vision AG, Bayer, Bayhill, Biogen Idec, BioMarin, Celgene, CSL Behring, df-mp Molnia & Pohlman, Eisai, Eli Lilly EU, EMD Serono, Genentech, Genmab, GeNeuro SA, Genzyme, Gianni Rubatto Foundation, Glaxo Smith Kline, Glenmark, Innosuisse, Janssen, Japan Tobacco, Merck Serono, MediciNova, Minoryx Therapeutics, Mitsubishi Pharma, MH Consulting, Neurostatus-UHB AG, Novartis, Novartis Research Foundation, Novo Nordisk, Österreichische Gesellschaft für Neurologie, Peptimmune, Roche, Roche Research Foundation, Santhera, Sanofi-Aventis, Senda Biosciences Inc, Swiss MS Society, Swiss National Research Foundation, Teva Pharmaceutical Industries Ltd, TG Therapeutics, UCB, Wellmera AG, and Wyeth; and Dr Kappos reported having a patent for Neurostatus UHB-AG with royalties paid Payments made to institution (University Hospital Basel); being CEO (employment by University Hospital Basel)MAGNIMS Steering Committee) and a board member of the European Charcot Foundation. Dr Gobbi reported receiving honoraria for speaking/consulting or grants from Abbvie, Almirall, Biogen Idec, Celgene, Genzyme, Merck Serono, Novartis, Roche, Teva Pharma. Dr Granziera reported The University Hospital Basel (USB), as the employer of C.G., has received the following fees which were used exclusively for research support: (1) advisory boards and consultancy fees from Actelion, Novartis, Genzyme-Sanofi, GeNeuro, Hoffmann La Roche and Siemens Healthineers; (2) speaker fees from Biogen, Hoffmann La Roche, Teva, Novartis, Janssen and Genzyme-Sanofi; and (3) research grants from Hoffmann La Roche, GeNeuro, Genzyme, Biogen. Dr Bridel reported serving on scientific advisory boards for Biogen, Novartis, and BMS. Dr Leppert reported receiving grants from Progressive MS Alliance outside the submitted work; and being chief medical officer of GeNeuro. Dr Kuhle reported receiving grants from Swiss National Science Foundation, Swiss MS Society, Biogen, Celgene, Merck, Novartis, Roche, Sanofi, Progressive MS Alliance, University of Basel, Octave Bioscience. No other disclosures were reported.

Funding/Support: This investigation was supported in part by grant 320030189140/1 from the Swiss National Science Foundation, award PA-2007-36872 from Progressive MS Alliance, 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.

Role of the Funder/Sponsor: The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Sharing Statement: See Supplement 2.

Additional Contributions: We thank the patients and relatives for their participation and support, the study nurses in participating centers for their motivated collaboration and recruitment efforts, and the administrative personnel of the Swiss Multiple Sclerosis Cohort. We thank Juan Francisco Vilchez Gomez, MSc, and Amar Zadic, MSc, University of Basel, Switzerland, for valuable technical assistance. No one received financial compensation for their contribution.

REFERENCES

1. Giovannoni G, Popescu V, Wuerfel J, et al. Smouldering multiple sclerosis: the real MS. *Ther Adv Neurol Disord*. 2022;15:17562864211066751. doi:10.1177/17562864211066751

2. Cree BAC, Gourraud PA, Oksenberg JR, et al; University of California, San Francisco MS-EPIC Team. Long-term evolution of multiple sclerosis disability in the treatment era. *Ann Neurol*. 2016;80(4):499-510. doi:10.1002/ana.24747

3. Cree BAC, Hollenbach JA, Bove R, et al; University of California, San Francisco MS-EPIC Team. Silent progression in disease activity-free relapsing multiple sclerosis. *Ann Neurol*. 2019;85(5):653-666. doi:10.1002/ana.25463

4. Benkert P, Meier S, Schaedelin S, et al; NfL Reference Database in the Swiss Multiple Sclerosis Cohort Study Group. Serum neurofilament light chain for individual prognostication of disease

activity in people with multiple sclerosis: a retrospective modelling and validation study. *Lancet Neurol*. 2022;21(3):246-257. doi:10.1016/S1474-4422(22)00009-6

5. Novakova L, Zetterberg H, Sundström P, et al. Monitoring disease activity in multiple sclerosis using serum neurofilament light protein. *Neurology*. 2017;89(22):2230-2237. doi:10.1212/WNL.0000000000004683

6. Bittner S, Steffen F, Uphaus T, et al; KKNMS consortium. Clinical implications of serum neurofilament in newly diagnosed MS patients: a longitudinal multicentre cohort study. *EBioMedicine*. 2020;56:102807. doi:10.1016/j.ebiom.2020.102807

7. Gafson AR, Jiang X, Shen C, et al. Serum neurofilament light and multiple sclerosis progression independent of acute inflammation. *JAMA Netw Open*. 2022;5(2):e2147588. doi:10.1001/jamanetworkopen.2021.47588

8. Thebault S, Reaume M, Marrie RA, et al. High or increasing serum NfL is predictive of impending multiple sclerosis relapses. *Mult Scler Relat Disord*. 2022;59:103535. doi:10.1016/j.msard.2022.103535

9. Bridel C, Leurs CE, van Lierop ZYGJ, et al. Serum neurofilament light association with progression in natalizumab-treated patients with relapsing-remitting multiple sclerosis. *Neurology*. 2021;97(19):e1898-e1905. doi:10.1212/WNL.0000000000012752

10. Cantó E, Barro C, Zhao C, et al. Association between serum neurofilament light chain levels and long-term disease course among patients with multiple sclerosis followed up for 12 years. *JAMA Neurol*. 2019;76(11):1359-1366. doi:10.1001/jamaneurol.2019.2137

11. Manouchehrinia A, Stridh P, Khademi M, et al. Plasma neurofilament light levels are associated with risk of disability in multiple sclerosis. *Neurology*. 2020;94(23):e2457-e2467. doi:10.1212/WNL.0000000000009571

12. Leppert D, Kropshofer H, Häring DA, et al. Blood neurofilament light in progressive multiple sclerosis: post hoc analysis of 2 randomized controlled trials. *Neurology*. 2022;98(21):e2120-e2131. doi:10.1212/WNL.0000000000200258

13. Norgren N, Sundström P, Svenningsson A, Rosengren L, Stigbrand T, Gunnarsson M. Neurofilament and glial fibrillary acidic protein in multiple sclerosis. *Neurology*. 2004;63(9):1586-1590. doi:10.1212/01.WNL.0000142988.49341.D1

14. Axelsson M, Malmeström C, Nilsson S, Haghighi S, Rosengren L, Lycke J. Glial fibrillary acidic protein: a potential biomarker for progression in multiple sclerosis. *J Neurol*. 2011;258(5):882-888. doi:10.1007/s00415-010-5863-2

15. Petzold A, Eikelenboom MJ, Gveric D, et al. Markers for different glial cell responses in multiple sclerosis: clinical and pathological correlations. *Brain*. 2002;125(pt 7):1462-1473. doi:10.1093/brain/awf165

16. Högel H, Rissanen E, Barro C, et al. Serum glial fibrillary acidic protein correlates with multiple sclerosis disease severity. *Mult Scler*. 2020;26(2):210-219. doi:10.1177/1352458518819380

17. Abdelhak A, Huss A, Kassubek J, Tumani H, Otto M. Serum GFAP as a biomarker for disease severity in multiple sclerosis. *Sci Rep*. 2018;8(1):14798. doi:10.1038/s41598-018-33158-8

18. Abdelhak A, Foschi M, Abu-Rumeileh S, et al. Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat Rev Neurol*. 2022;18(3):158-172. doi:10.1038/s41582-021-00616-3

19. Martínez MAM, Olsson B, Bau L, et al. Glial and neuronal markers in cerebrospinal fluid predict progression in multiple sclerosis. *Mult Scler*. 2015;21(5):550-561. doi:10.1177/1352458514549397

20. Mañé-Martínez MA, Olsson B, Bau L, et al. Glial and neuronal markers in cerebrospinal fluid in different types of multiple sclerosis. *J Neuroimmunol*. 2016;299:112-117. doi:10.1016/j.jneuroim.2016.08.004

21. Gunnarsson M, Malmeström C, Axelsson M, et al. Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab. *Ann Neurol*. 2011;69(1):83-89. doi:10.1002/ana.22247

22. Burman J, Zetterberg H, Fransson M, Loskog AS, Raininko R, Fagius J. Assessing tissue damage in multiple sclerosis: a biomarker approach. *Acta Neurol Scand*. 2014;130(2):81-89. doi:10.1111/ane.12239

23. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*. 2014;83(3):278-286. doi:10.1212/WNL.0000000000000560

24. Disanto G, Barro C, Benkert P, et al; Swiss Multiple Sclerosis Cohort Study Group. Serum neurofilament light: a biomarker of neuronal damage in multiple sclerosis. *Ann Neurol*. 2017;81(6):857-870. doi:10.1002/ana.24954

25. 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(8):2382-2391. doi:10.1093/brain/awy154

26. Teunissen CE, Petzold A, Bennett JL, et al. A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology*. 2009;73(22):1914-1922. doi:10.1212/WNL.0b013e3181c47cc2

27. Calabrese M, Agosta F, Rinaldi F, et al. Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. *Arch Neurol*. 2009;66(9):1144-1150. doi:10.1001/archneurol.2009.174

28. Cagol A, Schaedelin S, Barakovic M, et al. Association of brain atrophy with disease progression independent of relapse activity in patients with relapsing multiple sclerosis. *JAMA Neurol*. 2022;79(7):682-692. doi:10.1001/jamaneurol.2022.1025

29. Prineas JW, Lee S. Multiple sclerosis: destruction and regeneration of astrocytes in acute lesions. *J Neuropathol Exp Neurol*. 2019;78(2):140-156. doi:10.1093/jnen/nly121

30. Saraste M, Bezukladova S, Matilainen M, et al. Increased serum glial fibrillary acidic protein associates with microstructural white matter damage in multiple sclerosis: GFAP and DTI. *Mult Scler Relat Disord*. 2021;50(January):102810. doi:10.1016/j.msard.2021.102810

31. Maggi P, Kuhle J, Schädelin S, et al. Chronic white matter inflammation and serum neurofilament levels in multiple sclerosis. *Neurology*. 2021;97(6):e543-e553. doi:10.1212/WNL.0000000000012326

32. Abdelhak A, Hottenrott T, Morenas-Rodríguez E, et al. Glial activation markers in CSF and serum from patients with primary progressive multiple sclerosis: potential of serum GFAP as disease severity marker? *Front Neurol*. 2019;10(March):280. doi:10.3389/fneur.2019.00280

33. Ayrignac X, Le Bars E, Duflos C, et al. Serum GFAP in multiple sclerosis: correlation with disease type and MRI markers of disease severity. *Sci Rep*. 2020;10(1):10923. doi:10.1038/s41598-020-67934-2

34. Kuhle J, Maceski AM, Meinert R, et al. Plasma neurofilament light chain and glial fibrillary acidic protein levels are prognostic of disability worsening: a biosignature that helps in differentiating active from nonactive SPMS. Poster presented at: the American Academy of Neurology Meeting; April 21, 2021; online.

Supplementary Online Content

Meier S, Willemse EAJ, Schaedalin S, et al. Serum glial fibrillary acidic protein compared with neurofilament light chain as a biomarker for disease progression in multiple sclerosis. *JAMA Neurol*. Published online February 6, 2023. doi:10.1001/jamaneurol.2022.5250

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

Study design and MS patients

The Swiss Multiple Sclerosis Cohort (SMSC; NCT02433028) is a prospective multicentre cohort study performed across eight Swiss academic medical centres.[1–3] Demographic, neuroimaging, and clinical data as well as blood samples are collected every 6 or 12 months. Standardized clinical assessments with EDSS score calculations are performed by certified raters.[1,4,5] Confirmed disease worsening (CDW) was defined as an increase in EDSS of ≥1.5 points from an EDSS score of 0, ≥1.0 points from an EDSS score of 1.0–5.5 or ≥0.5 points from an EDSS score ≥6.0 confirmed at a subsequent visit ≥6 months later. Due to the observational setting, a roving baseline (BL) definition was used.[6] Relapses were defined as new, worsening or recurrent neurologic symptoms that lasted for at least 24 hours without fever, infection, or adverse reaction to a prescribed medication and that were preceded by a stable or improving neurologic status of at least 30 days. Disease modifying treatments (DMTs) were categorized into high-efficacy monoclonal antibody therapies (mAB; natalizumab, rituximab and ocrelizumab), oral therapies (orals; dimethyl fumarate, fingolimod, siponimod, ozanimod and teriflunomide), platform compounds (platform; interferon beta and glatiramer acetate), and untreated.

Cohort 1

stMS was defined as having no relapses or CDW during the entire FU. RMS patients in active disease phase experienced a relapse within the prior 30 days or/and had one or more CEL in an MRI scan <30 days before serum sampling. For remission timepoints, samples within one year prior of, or six months after a relapse, or CEL in MRI within 30 days from sampling were excluded. This selection underwent careful and independent inspection by two neurologists (JO and JK) to confirm worsening as captured by EDSS (e.g. patients with objectively worsening ataxia in the upper limbs but stable EDSS scores were excluded). Patients with relevant comorbidities (diabetes mellitus, hypertension, surgical orthopedic interventions influencing walking distance) were excluded. Based on these criteria, patients with most pronounced disease progression or signs of active disease were selected from the SMSC patients followed at the University Hospital Basel (n=745).

sGFAP and sNfL measurements

All measurements were performed with reagents from one lot for cohort 1 and one lot for cohort 2. A total of 7 runs for cohort 1 and 10 runs for cohort 2 on two HD-X analyser were required to measure all samples. All longitudinal samples from the same healthy control/patient were measured in the same run. The runs consisted of evenly distributed numbers of patients and controls samples across all runs.

Cohort 1: Inter-assay coefficients of variation (CV) for six native serum samples (sGFAP concentrations ranging from 43 to 121 pg/mL) showed a mean CV of 10.5% (range: 9-12%). A duplicate CV <20% was accepted (few samples were repeated) and the mean intra-assay CV of duplicate measurements in all samples was 4.3%.

Cohort 2: Inter-assay CVs for sGFAP in five human serum controls (one spiked with human cerebrospinal fluid; concentrations ranging from 26.2 to 349.8pg/mL) showed a mean CV of 8.1% (range: 6.0-11.7%). A duplicate CV of <20% was accepted (few samples were repeated) and the mean intra-assay CV of duplicate measurements for sGFAP in all samples was 6.2%. The inter-assay CV for sNfL (concentrations ranging from 7.7 to 120.3pg/mL) was 8.1% (range: 6.1-9.9%). The mean intra-assay CV of duplicate measurements for sNfL in all samples was 5.2%. Nine samples showed sGFAP levels below 16.6pg/ml (lower limit of quantification[7]) and were excluded from the analysis. Parallel comparison of sNfL results measured with the Nf-Light kit and the Neurology 2-plex B assay showed excellent congruency (Pearson's r = 0.964; **eFigure 1**). sNfL Z-scores from the Neurology 2-plex B assay were therefore calculated using the sNfL reference data generated with the Nf-Light kit.[3]

MRI assessment methods

Brain MRI scans were performed annually in the SMSC. A standardized imaging protocol was applied across centers including a 3D Magnetization Prepared – Rapid Gradient Echo (MPRAGE), a 3D Fluid Attenuated Inversion Recovery (FLAIR) sequence, and a post contrast T1w sequence acquired at a spatial resolution of 1 mm^3. T2 lesion volume (T2LV) was calculated automatically on FLAIR images using the multidimensional gated recurrent units algorithm,[8] and results were manually reviewed by experts. Longitudinal changes of white matter lesions were automatically assessed with LeMan-PV,[9] and the outputs, in terms of new and enlarged lesions (NEL), were manually reviewed and corrected. The number of CEL was assessed manually. T1w images

were lesion-filled using the FSL-lesion filling tool[10] and segmented by applying the SPM12 unified segmentation tool[11] to compute gray matter (GMV), white matter (WMV) and CSF (CSFV) volumes. The total intracranial volume (TIV) was calculated as TIV = GMV + WMV + CSFV and total brain volume (TBV) as the sum of GMV and WMV.

Statistical analysis

a) Cohort 1

Demographic and clinical characteristics were described as counts and percentages as well as median and interquartile range (IQR), as appropriate, and were compared using Fisher's exact test for categorical variables and Wilcoxon test for continuous variables (all non-normally distributed). Raw biomarker concentrations in healthy controls (HC) were analyzed using mixed models with log-transformed sGFAP or sNfL as dependent variable and age and sex as independent variables with a random intercept for person to account for the repeated structure of the data. The correlation between sGFAP and sNfL in BL samples of HC was quantified with the Spearman correlation coefficient. Comparison of sGFAP and sNfL levels in stMS/wPMS and RMS groups versus (vs) HC was performed using a linear mixed model with log-transformed sGFAP or sNfL as dependent variable and age, BMI, sex and phenotype group (stMS, wPMS, RMS in remission and RMS with active disease) as independent variables as well as person as random intercept to correct for repeated measures. Estimates were back-transformed and represent percentage change in the geometric mean of the biomarker level per unit change in the independent variable.

To assess the association between disease progression and sGFAP or sNfL levels (dependent variable, log-transformed), univariable and multivariable models with stMS vs wPMS status as well as age, sex, BMI, FU time, disease duration, DMT and EDSS scores as independent variables and a random intercept for person were used. Similar models were built to investigate the effect of active disease vs remission on sGFAP or sNfL levels in the RMS cohort. In an attempt to evaluate the independent association between progression status or active disease status and sGFAP or sNfL that is not explained by the other biomarker, the respective $\log_2$ transformed marker (estimates indicating effects per doubling) was additionally added to the above models, i.e. $\log_2$(sNfL) was added as covariate to the model with log(sGFAP) as dependent variable and vice versa. Sensitivity analyses for both biomarkers including T2LV, and number of NEL and CEL were additionally performed. For visualization purposes, estimates (marginal effects) from the above-described models were plotted which show the association of a given variable with the endpoint while accounting for repeated measures and correcting for the other covariates.

The within person variation of sGFAP or sNfL was assessed by the intraclass correlation coefficient (ICC) with 95% confidence interval obtained by bootstrapping. The ICC is calculated by fitting a separate mixed model for each biomarker containing solely an intercept term as well as a random intercept for patient. The ICC is estimated by dividing the variation which was due to the subject-to-subject difference by the total variance observed. The ICC can take values between 0 and 1 and can be interpreted as the proportion of the variation of the data which can be attributed to subject-to-subject variability. An ICC of 1 indicates that all differences in observed data are explainable by variability between subjects and lower values indicate higher within patient variation.

Atrophy rates per year in the combined stMS and wPMS cohort were assessed with a linear mixed model with log2-transformed TBV as dependent variable and TIV, age at BL, sex, disease duration at BL, and the interaction between stMS/wPMS and FU time (quantifying the group difference in atrophy rates) as independent variables, with a random intercept for person. Similarly, models using interaction terms between BL sGFAP and FU time as well as BL sNfL and FU time to assess the association between biomarker levels and log-transformed GMV or WMV as dependent variable were built. To compare the prognostic power of BL sGFAP and sNfL levels for PIRA, uni- and multivariable Cox regression analyses were performed in the combined stMS and wPMS cohort with $\log_2$-transformed sGFAP or sNfL at BL as predictors. Both unadjusted hazard ratios and estimates adjusted for sex, age, BMI and disease duration at BL are presented. P-values below 0.05 were considered statistically significant. Analyses were performed in R version 4.2.0.

b) Cohort 2

Demographic and clinical characteristics were described as counts and percentages as well as median and interquartile range (IQR), as appropriate, and were compared using Fisher's exact test for categorical variables and Wilcoxon test for continuous variables (all non-normally distributed). In HC, the association between log-transformed biomarker concentrations as dependent variable and age, sex and body mass index (BMI) as independent variables were analysed using mixed models with a random intercept for person. In analogy with age- and BMI-adjusted sNfL reference values[3], we calculated sGFAP Z-scores as follows: the above multivariable analysis confirmed age, sex and BMI as significant predictors of sGFAP. We used a generalized additive model for location, scale and shape (GAMLSS) based on a Box-Cox t distribution with sGFAP as

dependent variable and the three covariates. Based on investigating univariable associations graphically (**eFigure 2**) and by taking into account model fit of alternative models based on the Akaike information criterion, we defined a final parsimonious model which included age modelled with splines using three degrees of freedom, BMI (linear) and sex.

Biomarker levels in patients with and without later CDW were visualised using boxplots and compared using Wilcoxon signed rank test. Levels were considered increased compared to HC (a Z-score of 0 (50^{th} percentile) indicates the physiologic mean level of HC[3]) when being significantly above Z=0 in the univariate Wilcoxon signed rank tests.

A cross-sectional analysis was performed using linear models with individual biomarker Z-score as dependent variable and following predictors: age, sex, BMI, EDSS, disease subtype, disease modifying treatment (DMT), time since DMT therapy start and whether the patient developed CDW during follow-up (FU) ("CDW status"). Estimated additive effects on biomarker Z-scores are reported based on the full models including all covariates. Analyses using log-transformed biomarker levels instead of Z-scores are provided as supplementary data. Whereas the latter models capture variables explaining the variation in observed raw biomarker levels, the former models identify factors explaining increased biomarker levels in B-cell depleted MS patients compared to healthy controls while differences due to confounding effects of physiological aspects (age, BMI, and sex for sGFAP) have already been eliminated when building the Z-scores. However, these 3 variables are still included as covariates in the multivariable models with Z-scores as endpoints since they now quantify potential disease-related effects.

The association between biomarker levels and time to CDW/PIRA was investigated using Kaplan-Meier curves and Cox regression models, using Z-scores as continuous predictor as well as dichotomised in high versus (vs) low levels based on increasing cut-offs. As a sensitivity analysis, multivariable models adjusted for the above-mentioned covariates were performed.

Receiver operating characteristics (ROC) analyses were used to identify optimal cut-points for sGFAP and sNfL Z-score values to dichotomize the respective biomarker levels in high and low groups in studying the association with future CDW/PIRA. The performance of a composite of both biomarkers in prognosticating CDW/PIRA was investigated by categorizing patients into four groups according to high and low levels for each biomarker, using the constellation of "$sGFAP_{low}/sNfL_{low}$" as reference. P-values below 0.05 were considered statistically significant. Analyses were performed in R version 4.1.0.

eTable 1. Multivariable Mixed Models Testing Associations Between sGFAP and sNfL and Age, Sex, BMI, and MS Extreme Phenotypes vs Healthy Controls

		sGFAP (pg/ml), median, IQR	Est.	95%CI	*p*	sNfL (pg/ml), median, IQR	Est.	95%CI	*p*
Group	HC (485)	51.8 [41.2-69.7]	-	-	-	6.3 [4.7-8.5]	-	-	-
	stMS (169)	63.2 [43.4-90.7]	1.141	0.970-1.343	0.12	7.2 [5.4-9.4]	1.164	1.013-1.337	0.03
	wPMS (184)	103.0 [81.3-132.5]	1.770	1.498-2.091	<0.001	10.9 [8.2-13.9]	1.502	1.304-1.730	<0.001
	RRMS Remission (66)	52.9 [40.2-70.9]	1.143	1.030-1.270	0.01	6.7 [5.5-8.9]	1.264	1.142-1.399	<0.001
	RRMS Active (66)	59.1 [45.4-79.3]	1.225	1.102-1.360	<0.001	10.2 [7.7-16.2]	1.986	1.793-2.199	<0.001
Age			1.016	1.013-1.019	<0.001		1.023	1.020-1.026	<0.001
BMI			0.985	0.978-0.993	<0.001		0.973	0.966-0.981	<0.001
Sex	F (654)	61.7 [46.4-89.7]	1.127	1.039-1.223	0.004	7.5 [5.4-9.9]	0.987	0.917-1.063	0.73
	M (316)	58.9 [42.0-86.5]	-	-	-	7.3 [5.4-12.3]	-	-	-

Estimates (Est.) are multiplicative effects. Numbers in parentheses in the first column state the number of samples.
Abbreviations; BMI: body mass index; CI: confidence interval; F: female; HC: healthy control; IQR: interquartile range; M: male; RRMS: relapsing remitting MS; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; stMS: stable MS; wPMS: worsening progressive MS.

eTable 2. Sensitivity Analysis of Multivariable Mixed Linear Models Investigating the Association Between Worsening Status and sGFAP Levels (Left) and sNfL Levels (Right) With Additional Correction for MRI Variables

		sGFAP (pg/ml), median, IQR	Est.	95%CI	*p*	sNfL (pg/ml), median, IQR	Est.	95%CI	*p*
Sensitivity analysis: MRI (n=184)									
Age at BL			1.005	0.998-1.021	0.62		1.017	1.008-1.026	0.002
FU time			1.012	1.000-1.025	0.06		1.026	1.011-1.041	0.001
Sex	F (115)	87.1 [52.4-108.1]	1.067	0.775-1.470	0.71	8.4 [6.1-10.9]	0.946	0.800-1.118	0.57
	M (69)	81.5 [59.760-120.2]	-	-	-	12.2 [5.8-17.2]	-	-	-
BMI				0.982-1.025	0.72		0.980	0.963-0.996	0.03
Disease duration at BL				0.984-1.021	0.82		1.006	0.996-1.016	0.32
DMT	Untreated (21)	105.8 [82.7-123.0]	-	-	-	14.0 [10.9-17.7]	-	-	-
	Platform (22)	70.1 [56.757-91.592]	1.602	1.154-2.199	0.006	10.8 [6.4-18.5]	1.342	1.024-1.810	0.06
	Orals (71)	71.5 [36.637-97.998]	1.055	0.872-1.268	0.58	7.4 [5.4-9.5]	1.030	0.841-1.231	0.77
	mAB (70)	91.7 [62.663-134.3]	1.065	0.920-1.233	0.41	9.5 [7.0-12.7]	0.994	0.836-1.155	0.95
EDSS score				0.999-1.092	0.06		1.033	0.991-1.079	0.16
T2w lesion volume (log+1)*				0.962-1.148	0.28		1.099	1.018-1.178	0.02
NEL*				0.981-1.001	0.08		1.005	0.996-1.015	0.35
CEL*				1.006-1.510	0.05		1.291	1.017-1.637	0.04
Progression	**stMS (99)**	62.2 [40.4-93.4]	-	-	-	6.8 [5.5-9.5]	-	-	-
	wPMS (85)	103.0 [84.1-138.6]	1.692	1.218-2.347	0.006	11.3 [8.6-14.3]	1.256	1.040-1.523	0.04

*Information on T2LV, NEL and CEL were available for 184/352 visits (Stable MS: n: 99 and worsening progressive MS (wPMS): n: 85). wPMS and stable MS had CEL or at least 2 NEL at some point during FU: 9 wPMS (in 4 patients twice; in 2 patients 3 times; overall: 20% of visits) and 5 stable MS patients (in one patient twice; overall: 6% of visits). Numbers in parentheses in the first column state the number of samples.

Abbreviations: BMI: body mass index; CEL: contrast enhancing lesion; CI: confidence interval; DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale; F: female; IQR: interquartile range; M: male; mAB: monoclonal antibody therapies; NEL: new enlarging T2 lesions; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; stMS: stable MS; T2LV: T2w lesion volume; wPMS: worsening progressive MS.

eTable 3. Multivariable Mixed Linear Models Investigating the Effect of Focal Inflammation (Remission vs Active State) on sGFAP Levels (Left) and sNfL Levels (Right)

		sGFAP (pg/ml), median, IQR	Est.	95%CI	*p*	sNfL (pg/ml), median, IQR	Est.	95%CI	*p*
Model 1: Univariate									
Focal inflammation	**Remission (66)**	52.9 [40.2-70.9]	-	-	-	6.7 [5.5-8.9]	-	-	-
	Active (66)	59.1 [45.4-79.3]	1.073	1.002-1.150	0.05	10.2 [7.7-16.2]	1.584	1.338-1.874	<0.001
Model 2: multivariable									
Age			1.008	0.998-1.018	0.12		1.009	0.997-1.021	0.15
Sex	F (100)	59.6 [42.9-80.0]	1.032	0.861-1.238	0.75	8.7 [6.1-13.9]	1.157	0.926-1.447	0.23
	M (32)	53.4 [46.1-61.1]	-	-	-	7.6 [5.6-9.0]	-	-	-
BMI			0.969	0.951-0.987	0.002		0.967	0.945-0.991	0.01
Disease duration			1.006	0.994-1.017	0.34		0.996	0.982-1.010	0.58
DMT	Untreated (31)	71.9 [42.0-130.1]	-	-	-	10.5 [6.4-16.9]	-	-	-
	Platform (14)	56.7 [44.9-64.3]	0.904	0.766-1.066	0.25	7.7 [6.1-10.3]	0.885	0.645-1.215	0.46
	Orals (71)	52.2 [42.7-64.7]	0.872	0.768-0.982	0.03	7.9 [5.9-10.3]	0.881	0.707-1.111	0.28
	mAB (16)	63.6 [50.9-86.1]	0.996	0.843-1.169	0.96	7.8 [5.7-16.9]	1.070	0.785-1.487	0.68
EDSS score			1.122	1.058-1.186	<0.001		1.238	1.128-1.353	<0.001
Focal inflammation	**Remission (66)**	52.9 [40.2-70.9]	-	-	-	6.7 [5.5-8.9]	-	-	-
	Active (66)	59.1 [45.4-79.3]	1.048	0.977-1.122	0.20	10.2 [7.7-16.2]	1.532	1.308-1.814	<0.001
Model 3: plus sNfL/sGFAP									
Age			1.007	0.998-1.016	0.16		1.004	0.993-1.015	0.54
Sex	F (100)	59.6 [42.9-80.0]	1.012	0.855-1.199	0.90	8.7 [6.1-13.9]	1.146	0.939-1.403	0.21
	M (32)	53.4 [46.1-61.1]	-	-	-	7.6 [5.6-9.0]	-	-	-
BMI			0.977	0.960-0.995	0.01		0.989	0.967-1.013	0.39
Disease duration			1.007	0.997-1.018	0.20		0.991	0.979-1.004	0.21
DMT	Untreated (31)	71.9 [42.0-130.1]	-	-	-	10.5 [6.4-16.9]	-	-	-
	Platform (14)	56.7 [44.9-64.3]	0.930	0.797-1.083	0.37	7.7 [6.1-10.3]	0.966	0.723-1.298	0.82
	Orals (71)	52.2 [42.7-64.7]	0.913	0.809-1.022	0.12	7.9 [5.9-10.3]	1.026	0.833-1.293	0.82
	mAB (16)	63.6 [50.9-86.1]	1.009	0.863-1.172	0.91	7.8 [5.7-16.9]	1.130	0.850-1.544	0.43
EDSS score			1.063	1.003-1.125	0.04		1.181	1.085-1.282	<0.001
sNfL (pg/ml) per doubling			1.145	1.081-1.215	<0.001		n.a	n.a.	n.a.
sGFAP (pg/ml) per doubling			n.a	n.a.	n.a.		1.528	1.287-1.806	<0.001
Focal inflammation	**Remission (66)**	52.9 [40.2-70.9]	-	-	-	6.7 [5.5-8.9]	-	-	-
	Active (66)	59.1 [45.4-79.3]	0.973	0.903-1.044	0.47	10.2 [7.7-16.2]	1.506	1.300-1.770	<0.001

Estimates (Est.) are multiplicative effects. Numbers in parentheses in the second column state the number of samples.
Abbreviations: BMI: body mass index; CI: confidence interval; DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale; F: female; IQR: interquartile range; M: male; mAB: monoclonal antibody therapies; n.a.: not applicable; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

eTable 4. Multivariable Mixed Models to Assess the Association Between BL sGFAP and BL sNfL and Longitudinal GMV or WMV

		Est.	95% CI	*p*
GMV				
TIV		0.9999	0.9998-1.0000	0.03
Age at BL		0.9980	0.9943-1.0017	0.33
Sex	F	0.8799	0.8171-0.9503	0.004
	M	-	-	
Disease duration at BL		0.9981	0.9943-1.0019	0.36
BL sGFAP (log2)		1.0479	0.9985-1.0993	0.09
BL sNfL (log2)		0.9400	0.8910-0.9926	0.05
FU time (years)		1.0111	1.0043-1.0178	0.002
Interaction BL sGFAP * FU time**		0.9976	0.9965-0.9988	<0.001
Interaction BL sNfL * FU time		0.9999	0.9989-1.0009	0.78
WMV				
TIV		1.0002	1.0001-1.0004	<0.001
Age at BL		1.0002	0.9962-1.0042	0.93
Sex	F	0.8868	0.8178-0.9639	0.01
	M	-	-	
Disease duration at BL		0.9955	0.9914-0.9996	0.05
BL sGFAP (log2)		1.0269	0.9745-1.0817	0.36
BL sNfL (log2)		0.9516	0.8977-1.0093	0.13
FU time (years)		1.0038	0.9957-1.0117	0.35
Interaction BL sGFAP * FU time		1.0005	0.9991-1.0018	0.48
Interaction BL sNfL * FU time**		0.9974	0.9962-0.9985	<0.001

**Reading example: Doubling of BL sGFAP levels is associated with a 0.24% increase in gray matter atrophy per year whereas doubling of BL sNfL levels is associated with a 0.26% increase in white matter atrophy. n=198 timepoints with volumetric endpoints available. Est. are multiplicative effects.
Abbreviations: BL: baseline; CI: confidence interval; Est: estimates; FU: follow-up; GMV: gray matter volume; MS: multiple sclerosis; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; TIV: total intracranial volume; WMV: white matter volume.

eTable 5. Patient Characteristics at Time of Sample Collection (Baseline)

	Total	Without CDW	With CDW	*p*
N	**252**	209	43	
Sex = female	**156 (61.9)**	131 (62.7)	25 (58.1)	0.70
BMI	**24.1 [21.8-27.4]**	24.1 [21.8-27.2]	24.1 [21.2-28.2]	0.85
Age	**44.3 [33.3-54.7]**	42.9 [33.1-53.7]	49.9 [38.0-59.5]	0.03
Disease duration, years	**9.9 [5.0-18.5]**	10.4 [5.0-19.6]	9.3 [4.8-17.4]	0.65
Disease subtype (at entry into the SMSC)				<0.001
RRMS	**181 (71.8)**	160 (76.6)	21 (48.8)	
SPMS	**34 (13.5)**	25 (12.0)	9 (20.9)	
PPMS	**37 (14.7)**	24 (11.5)	13 (30.2)	
EDSS	**3.0 [2.0-4.5]**	3.0 [2.0-4.5]	4.0 [2.8-6.0]	0.002
DMT				0.001
OCR	**169 (67.1)**	147 (70.3)	22 (51.2)	
RTX	**83 (32.9)**	62 (29.7)	21 (48.8)	
FU time, years	**3.1 [2.1-4.0]**	3.1 [2.1-3.9]	3.1 [2.0-4.0]	0.95
Time from treatment start to sampling, months	**12.2 [10.7-16.8]**	12.4 [10.7-17.5]	11.4 [10.7-14.8]	0.15
DMT during FU				<0.001
Only OCR	**164 (65.1)**	143 (68.4)	21 (48.8)	
Only RTX	**51 (20.2)**	43 (20.6)	8 (18.6)	
RTX --> OCR	**37 (14.7)**	23 (11.0)	14 (32.6)	
CDW during FU				<0.001
PIRA	**39 (15.5)**	0 (0.0)	39 (90.7)	
RAW	**4 (1.6)**	0 (0.0)	4 (9.3)	
Relapses during FU				0.79
0	**235 (93.3)**	194 (92.8)	41 (95.3)	
1	**16 (6.3)**	14 (6.7)	2 (4.7)	
3	**1 (0.4)**	1 (0.5)	0 (0)	
T2w lesion volume (ml)*	**7.0 [3.1-17.3]**	6.6 [3.1-13.5]	7.8 [3.3-42.7]	0.14
T2w lesion number*	**33.0 [23.0-50.5]**	32.5 [22.0-50.2]	35.0 [24.0-49.0]	0.76

Data are represented as number (percentage) or as median [IQR]. *Available for 53.1% of the cohort.
Abbreviations: CDW: confirmed disease worsening; DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale score; FU: follow-up; IQR: interquartile range; n.a.: not applicable; OCR: ocrelizumab; PIRA: progression independent of relapse activity; PPMS: primary progressive MS; RAW: relapse associated worsening; RRMS: relapsing remitting MS; RTX: rituximab; SPMS: secondary progressive MS.

eTable 6. Multivariable Linear Models Investigating the Effect of Demographic and MS-Related Characteristics on sGFAP *Z* Scores (Left) and sNfL *Z* Scores (Right)

N= 252 patients		Variance explained	sGFAP Z-score, median, IQR	Est.	95%CI	*p*	Variance explained	sNfL Z-score, median, IQR	Est.	95%CI	*p*
Age (per 10 years)			-	**-0.27**	**-0.44--0.11**	**0.001**		-	-0.10	-0.25-0.05	0.18
Sex	Men (96)		0.7 [-0.4-1.8]		-			0.4 [-0.4-1.3]		-	
	Women (156)		0.9 [0.2-2.0]	**0.36**	**0.02-0.70**	**0.04**		0.6 [-0.0-1.3]	0.18	-0.13-0.49	0.25
BMI (per 5 units)			-	-0.05	-0.21-0.11	0.54		-	-0.09	-0.24-0.05	0.21
EDSS			-	**0.23**	**0.11-0.35**	**<0.001**		-	0.09	-0.03-0.20	0.13
Disease course	RRMS (181)		0.9 [0.1-1.9]		-			0.5 [-0.2-1.2]		-	
	SPMS (34)	R^2=0.133*	0.6 [-0.3-2.1]	-0.50	-1.11-0.12	0.11	R^2=0.018**	0.8 [-0.3-1.5]	-0.14	-0.70-0.42	0.62
	PPMS (37)		0.6 [-0.2-2.1]	-0.28	-0.83-0.28	0.33		0.4 [-0.3-1.4]	-0.04	-0.54-0.46	0.88
DMT	RTX (83)		1.0 [0.2-2.1]		-			0.7 [-0.1-1.5]		-	
	OCR (169)		0.7 [-0.2-1.7]	**-0.36**	**-0.72--0.00**	**0.05**		0.4 [-0.3-1.2]	-0.15	-0.48-0.18	0.37
Months since DMT start			-	**-0.05**	**-0.09--0.01**	**0.03**		-	-0.03	-0.07-0.00	0.08
CDW status	No CDW (209)		0.7 [-0.1-1.7]		-			0.4 [-0.3-1.2]		-	
	CDW (43)		1.9 [0.4-2.2]	**0.59**	**0.14-1.04**	**0.01**		0.9 [-0.0-1.5]	0.24	-0.16-0.65	0.24

Legend: Significant associations are indicated in bold. Independent covariables: estimates per unit change are shown. *13.3%, **1.8%. N=number of patients. Numbers in parentheses in the first column state the number of patients.
Estimates represent additive effects (e.g. 0.59 Z-score units higher sGFAP Z-score in patients with vs without CDW during FU).
Abbreviations: BMI: body mass index; CDW: confirmed disease worsening; DMT: disease modifying treatment; EDSS: Expanded Disability Status Scale score; sGFAP: serum glial fibrillary acidic protein; IQR: interquartile range; sNfL: serum neurofilament light chain; OCR: ocrelizumab; PPMS: primary progressive MS; RRMS: relapsing remitting MS; RTX: rituximab; SPMS: secondary progressive MS.

eTable 7. Multivariable Linear Models Investigating the Effect of Demographic and MS-Related Characteristics on sGFAP (Left) and sNfL Concentrations (Right)

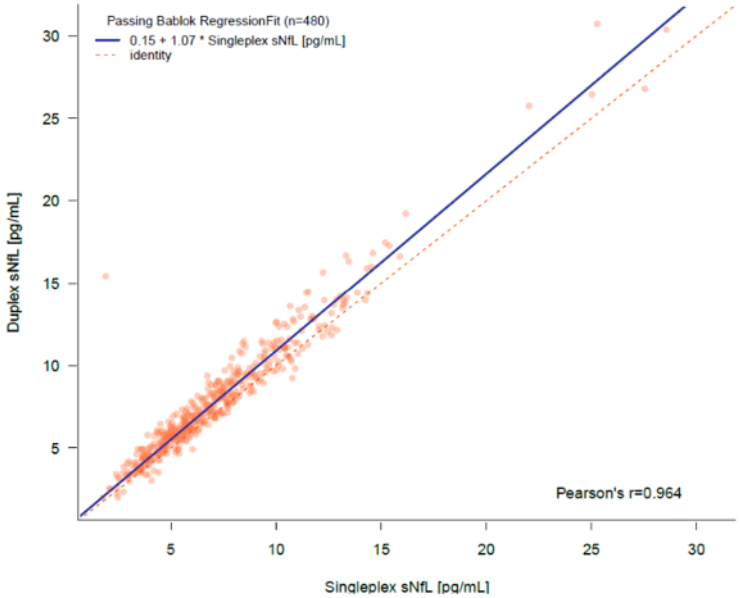
N=252 patients		**Variance explained**	**sGFAP conc. (pg/mL), median, IQR**	**Est.**	**95%CI**	***p***	**Variance explained**	**sNfL conc. (pg/mL), median, IQR**	**Est.**	**95%CI**	***p***
Age (per 10 years)		**R^2= 0.251 ***		**1.10**	**1.04-1.17**	**0.002**	**R^2=0.293 ****		**1.22**	**1.15-1.30**	**<0.001**
Sex	Men (96)		**71.4 [43.3-99.0]**	-				8.1 [6.5-11.7]	-		
	Women (156)		**84.3 [58.8-120.8]**	**1.28**	**1.13-1.45**	**<0.001**		8.0 [5.8-11.3]	1.04	0.92-1.18	0.50
BMI (per 5 units)				**0.91**	**0.86-0.97**	**0.003**			**0.92**	**0.87-0.97**	**0.003**
EDSS				**1.08**	**1.03-1.13**	**0.001**			1.02	0.98-1.07	0.31
Disease course	RRMS (181)		72.3 [51.9-105.8]	-				7.4 [5.8-10.1]	-		
	SPMS (34)		95.6 [63.7-146.5]	0.89	0.71-1.11	0.30		10.9 [7.9-15.3]	1.04	0.84-1.29	0.72
	PPMS (37)		92.2 [55.1-121.8]	0.94	0.77-1.15	0.56		11.0 [9.1-17.2]	1.08	0.88-1.31	0.46
DMT	RTX (83)		91.2 [63.4-122.5]	-				9.3 [6.8-12.2]	-		
	OCR (169)		72.4 [47.4-106.4]	0.88	0.77-1.01	0.06		7.9 [5.8-11.1]	1.01	0.89-1.14	0.92
Months since DMT start				**0.98**	**0.97-1.00**	**0.02**			1.00	0.98-1.01	0.76
CDW status	No CDW (209)		**73.1 [52.4-102.0]**	-				7.9 [6.1-11.3]	-		
	CDW (43)		**114.5 [70.4-144.1]**	**1.25**	**1.06-1.48**	**0.008**		10.0 [7.2-14.0]	1.08	0.92-1.27	0.32

Legend: Significant associations are indicated in bold. Independent covariables: estimates per unit change are shown. *25.1 %; ** 29.3%. N=number of patients. Numbers in parentheses in the first column state the number of patients.

Biomarker levels were log-transformed and estimates back-transformed representing multiplicative effects (e.g. 22% higher sNfL levels per 10 years of age).

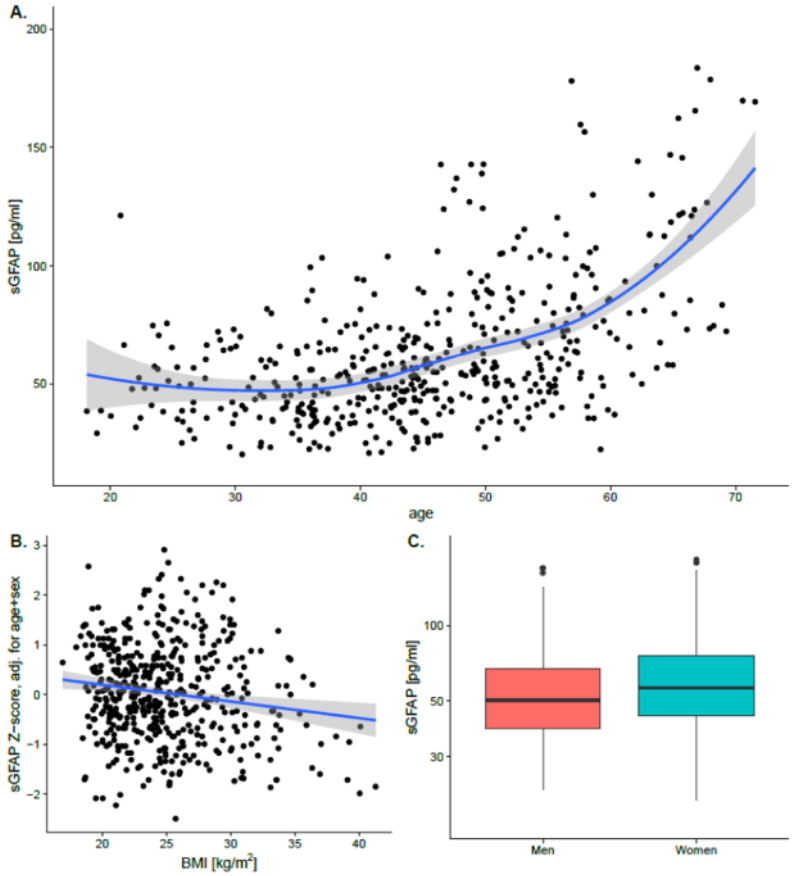
Abbreviations: BMI: body mass index; CDW: confirmed disease worsening; DMT: disease; EDSS: Expanded Disability Status Scale score; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; modifying treatment; IQR: interquartile range; OCR: ocrelizumab; PPMS: primary progressive MS; RTX: rituximab; RRMS: relapsing remitting MS; SPMS: secondary progressive MS.

eFigure 1. Comparison of sNfL Results From the Nf-Light Kit (Singleplex) and Neurology 2-Plex B Assay (Duplex) (n: 480)



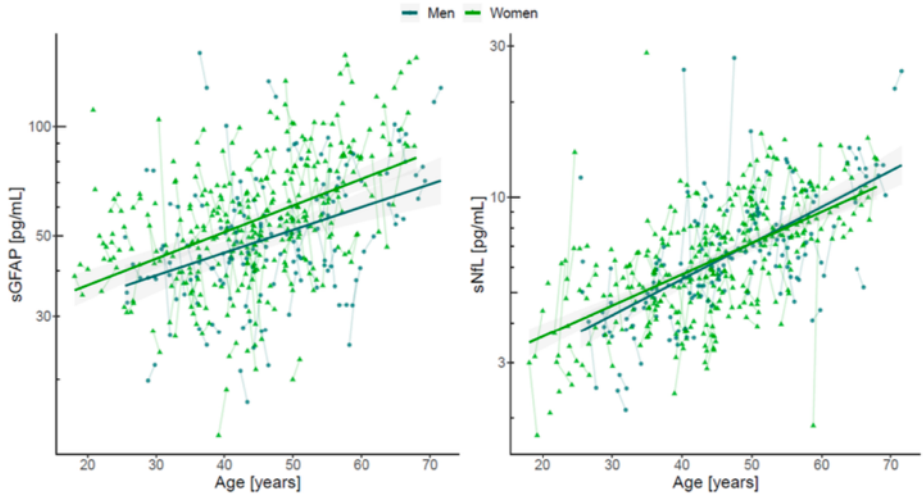
Legend: Each dot indicates an individual data point. The solid line indicates the Passing-Bablok regression line. The dotted line indicates the x=y identity line. Parallel comparison of sNfL results measured with the Nf-Light kit and the Neurology 2-plex B assay showed excellent congruency (Pearson's r = 0.964).
Abbreviations: sNfL: serum neurofilament light chain.

eFigure 2. Associations Between Age (A), BMI (B), and Sex (C) and sGFAP Concentrations in Healthy Controls



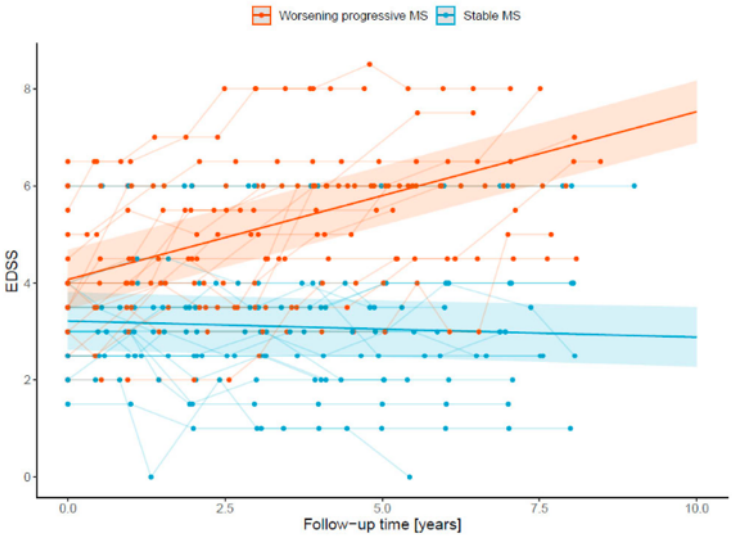
Legend: Graphical representation of the associations between sGFAP and age (A.), BMI (B.) as well as sex (C.): sGFAP increases with age in a non-linear manner (line represents a non-linear smoothing function with confidence band (A.) and a linear regression line with confidence band (B.)), decreases with BMI (sGFAP values adjusted for age are shown in B.) and are higher in women compared to men (see also **eFigure 3**).
Abbreviations: adj.: adjusted; BMI: body mass index; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

eFigure 3. Serum GFAP (Left) and sNfL (Right) and Age in Healthy Controls Stratified by Sex



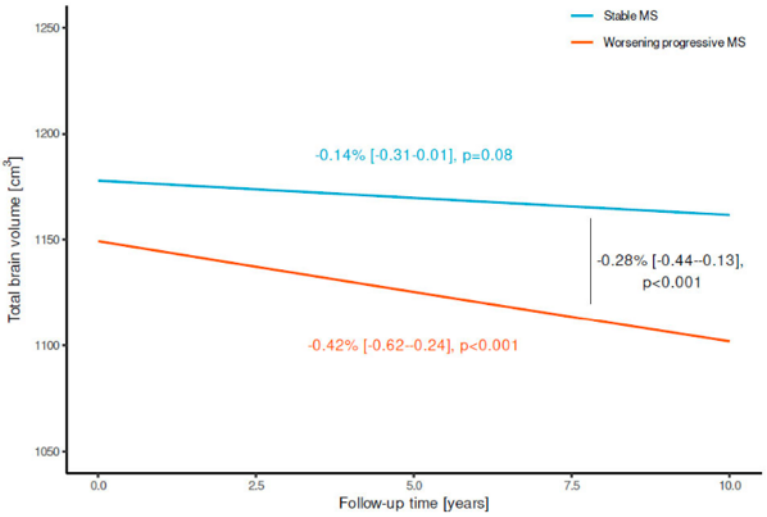
Legend: sGFAP (left) and sNfL (right) concentrations in samples from healthy controls (259 at baseline and 226 at follow-up) in relation to age, stratified for sex (men represented by blue circles; women by green triangles). Samples from one individual are connected through lines; thick lines show the group regression lines.
Serum GFAP levels increased with age (1.5% per year, estimate (est.) [95% CI] 1.015 [1.012-1.019], p<0.001; A.), and showed 14.9% higher levels in women compared to men (est. 1.149 [1.047-1.260], p=0.004). Serum NfL increased by 2.5% per year (est. 1.025 [1.022-1.028], p<0.001; B.), and showed no differences between sexes (est. 0.98 [0.90-1.06], p=0.62).
Abbreviations: sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

eFigure 4. EDSS Score Over Time in Stable MS and Worsening Progressive MS



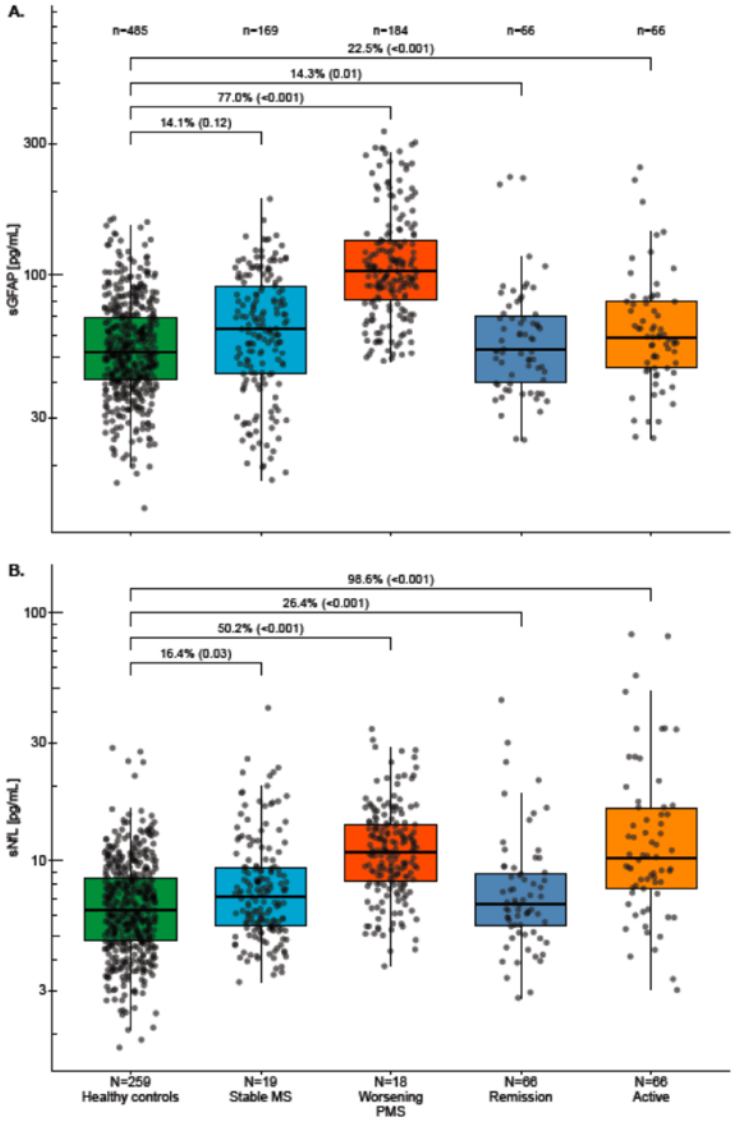
Legend: Patients with worsening progressive MS (red) showed an increase in EDSS score while stable patients (blue) maintain stable EDSS scores. Thin lines connect individual data points; thick lines including 95% CI show marginal effects from a mixed model with EDSS explained by an interaction term between follow up time and wPMS versus stMS plus a random intercept per patient. Abbreviations: EDSS: Expanded Disability Status Scale.

eFigure 5. Total Brain Volume Loss in Stable MS and Worsening Progressive MS



Legend: Worsening progressive MS patients showed an annual total brain volume (TBV) loss of -0.42% [95% CI: -0.62--0.24], p<0.001, which was significantly increased compared to TBV loss in stable MS patients (stMS) (-0.14% [-0.31-0.01], p=0.08; p-value of interaction wPMS/stMS * FU time: p<0.001).
Abbreviations: CI: confidence interval; FU: follow-up; MS: multiple sclerosis; stMS: stable MS; TBV: total brain volume; wPMS: worsening progressive MS

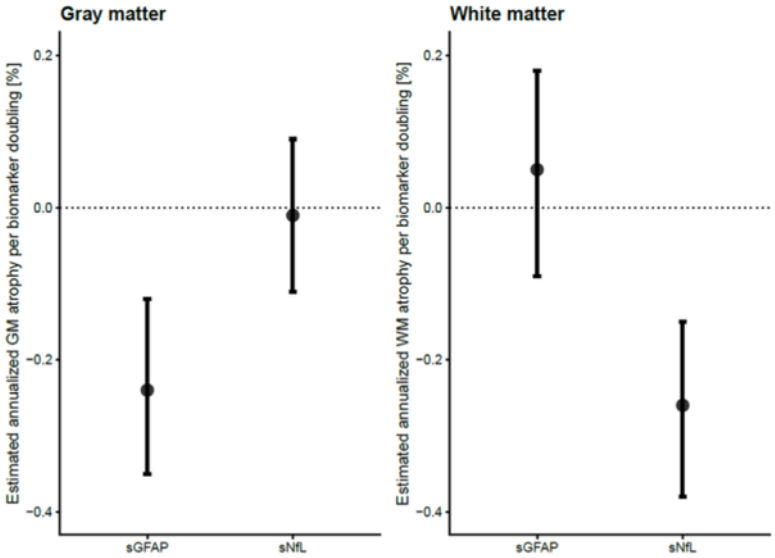
eFigure 6. Serum GFAP (A) and sNfL (B) in Different MS Groups vs Healthy Controls



Legend: Comparisons of sGFAP (A) and sNfL concentrations (B) in stable MS (stMS), worsening progressive MS (wPMS), patients in remission and active status vs healthy controls (HC). N: number of healthy controls/patients; n: number of samples. Boxplots show median and interquartile range and whiskers show the total range without outliers (defined as <1.5 times the interquartile range). Percentages increase versus HC and adjusted p values (in brackets) according **eTable 1** are shown. Serum GFAP levels were highest in wPMS, followed by RMS in active state, RMS during remission, and stMS patients. Conversely, sNfL levels were highest in active RMS, followed by wPMS, stMS, and RMS in remission.

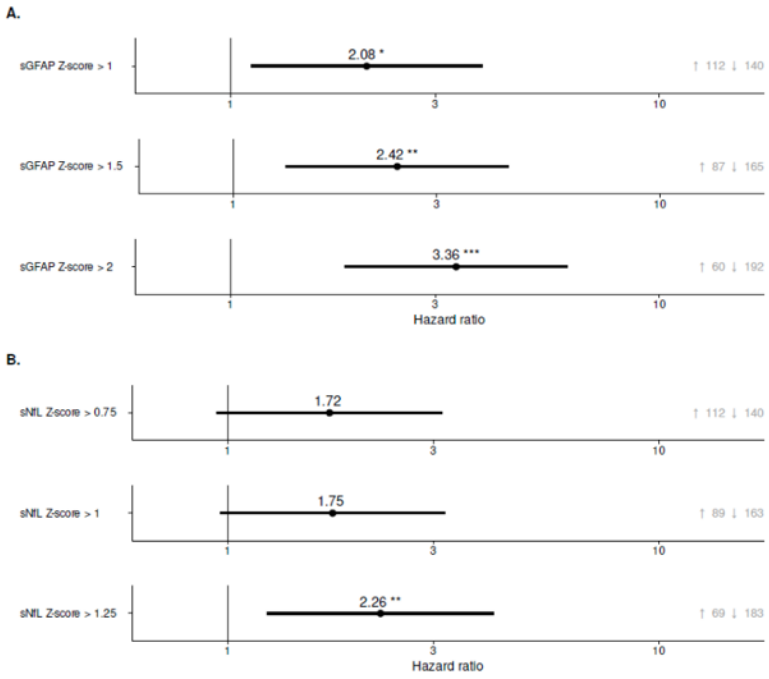
Abbreviations: PMS: progressive MS; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

eFigure 7. Associations of sGFAP and sNfL With Gray (A) and White Matter (B) Atrophy



Legend: Dots show estimated annualized atrophy per biomarker doubling and vertical bars show their 95% confidence intervals according to the multivariable mixed model in **eTable 4**.
Each doubling of BL sGFAP led to an additional loss of GMV (-0.24%/y [-0.35--0.12], p<0.001) but not WMV (-0.05% [-0.09-0.18], p=0.48), while doubling of BL sNfL resulted in additional loss of WMV (0.26% [-0.38--0.15], p<0.001) but not GMV (0.01% [-0.11-0.09], p=0.78).
Abbreviations: GM: gray matter; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; WM: white matter.

eFigure 8. Hazard Ratios for CDW Using Increasing *Z* Score Cut Points for sGFAP (A) and sNfL (B)



Legend: Dots show CDW hazard ratios and horizontal bars show their 95% confidence intervals from Cox regression models. Numbers in gray indicate the number of patients above (arrow up) or below (arrow down) the cut-point. Z-score cut-points were chosen with respect to keeping an acceptable distribution between patients above and below the cut-point. sGFAP Z-score cut-points of 1, 1.5 and 2 led to increasing hazards for CDW (A.). The associations for sNfL (B.) were less strong (and were not significant for cut-off above 1.25 (data not shown)). ***: p<0.001; **: p<0.01; *: p<0.05.
Abbreviations: CDW: confirmed disease worsening; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

eReferences

1. 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(3):1–13.
2. Disanto G, Barro C, Benkert P, et al. Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. Ann. Neurol. 2017;81(6):857–870.
3. Benkert P, Meier S, Schaedelin S, et al. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. Lancet Neurol. 2022;21(3):246–257.
4. Kappos L, D'Souza M, Lechner-Scott J, Lienert C. On the origin of Neurostatus. Mult. Scler. Relat. Disord. 2015;4(3):182–185.
5. Neurostatus-UHB Ltd c/o University Hospital Basel Switzerland. www.neurostatus.net. 2016;
6. Kappos L, Butzkueven H, Wiendl H, et al. Greater sensitivity to multiple sclerosis disability worsening and progression events using a roving versus a fixed reference value in a prospective cohort study. Mult. Scler. J. 2018;24(7):963–973.
7. Quanterix Corp. Simoa® Neurology 2-Plex B Kit HD-X Data Sheet, Item No 103520. [date unknown].
8. Andermatt S, Pezold S, Cattin P. Multi-Dimensional Gated Recurrent Units for the Segmentation of Biomedical 3D-Data. In: Deep Learning and Data Labeling for Medical Applications. 2016 p. 142–151.
9. Fartaria MJ, Kobera T, Granziera C, Cuadra MB. Longitudinal analysis of white matter and cortical lesions in multiple sclerosis. NeuroImage Clin. 2019;23(July):101938.
10. Battaglini M, Jenkinson M, De Stefano N. Evaluating and reducing the impact of white matter lesions on brain volume measurements. Hum. Brain Mapp. 2012;33(9):2062–2071.
11. Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26(3):839–851.

Chapter 4: Summary, discussion and future steps

In this project we aimed to advance the use of blood-based biomarkers to measure disease activity, worsening, progression[4,77] and as a secondary aim also treatment response in MS patients; we focused on two fluid biomarkers, sNfL and sGFAP, and their value in three of the currently most pressing issues in MS management: detecting subclinical disease activity, treatment monitoring and detecting disease progression in MS. Since sNfL is a step ahead of sGFAP in its development status towards clinical application, the focus of the sNfL part lay not only on measuring sNfL in MS patients, but on a critical missing piece which is the establishment of reference values of sNfL in healthy persons.

In the first study, the sNfL measurements of a large cohort of control persons showed a physiological dependence of sNfL concentrations on age and BMI with a non-linear association between age and sNfL and a further increase after the age of 50 years. Importantly we also noticed that below an estimated glomerular filtration rate (eGFR) of 60 mL/min/1.73 m^2, sNfL levels rapidly increased in control persons. Although in earlier studies age as a variable of sNfL levels has been recognized,[24,52] these findings have not been explored in a way to normalize values, i.e. to derive a measure that allows to compare measures across different ages. Here we show that the use of fixed cut-off levels of absolute sNfL concentrations to define pathological values is suboptimal in that concurrent disease activity in young persons may remain unrecognized, whereas older persons have a higher likelihood to have higher sNfL levels driven by a higher age, i.e. false positive results. To overcome this fundamental limitation, we created a reference database (RDB), based on normative sNfL values from 5'390 persons with 4'532 individual and 10'133 overall samples from Europe and the USA that provides a basis of comparison for pathological sNfL values. Based on the RDB, we created a web-based tool to calculate sNfL percentiles and Z scores that are adjusted for age and BMI. With this tool, researchers and clinicians can determine the age- and BMI-adjusted sNfL values of their patients and detect pathological levels more accurately than with absolute sNfL concentrations (832 active users sNfL reference app: https://shiny.dkfbasel.ch/baselnflreference/, as of May 12, 2023). The availability of sNfL Z scores or percentiles (which are interchangeable, however physicians may find percentiles more intuitive) strongly increases the applicability of sNfL measurements in clinical practice. Furthermore, the reference database may give rise to a stronger collaboration among MS researchers as a basis for streamlining the use of sNfL as a biomarker for MS. Indeed, following the publication of our reference database,[78] the American Food and Drug Administration (FDA) granted Quanterix Breakthrough Device Designation for

the SIMOA NfL test used in our publication, allowing for accelerated assessment and review processes for the potential approval of this immunoassay.[79]

During the past year, several other reference databases for sNfL values were published.[80–83] Nevertheless, these studies included smaller numbers of participants and used different measurement platforms and, in part, sample types (serum,[82] plasma[81] or both sample types[80,83]). Furthermore, our RDB is the only one correcting for age and BMI, as opposed to only age, leading to more accurate sNfL values. Despite these differences, a consensus should be reached on a European or global level to develop more comparable sNfL Z score ranges, specifically across different analytical platforms. A first step towards such a solution was taken in a commutability study for NfL measurements, which is an inter-center collaboration that our group has initiated and led together with the colleagues in Gothenburg.[84] Serum and plasma samples were evaluated on 4 different analytical platforms for NfL to determine the correlation between these different measurement methods towards a certified reference material (CRM) for making concentrations between platforms comparable.

A further part of this project was the association of biomarkers and the response to treatment in MS patients. The development of high-efficacy treatments in MS has also increased the need for stringent monitoring of their efficacy, especially in seemingly stable patients (NEDA-3), where importantly sNfL showed additional prognostic value, i.e. experiencing EDA-3 in/during the next year of clinical and MRI follow-up. We show that sNfL could in fact be used to monitor treatment response, as sNfL levels strongly decreased with monoclonal antibody and oral therapies, confirming findings from previous studies of MS patients undergoing a specific treatment.[54,55] Recent studies have used sNfL as an endpoint for their drug trials,[54,55] albeit only few of them compared the effects of different types of DMT's on blood-based biomarkers.[25,42,56] In these studies, similarly to our study, the rate of sNfL decrease was related to the treatment used, although they investigated different compounds to those examined herein. In our case, the grouping of treatment categories allowed for the detection of changes between treatment types on a larger scale. The strong differences in efficacy between therapies based on the sNfL values may have an influence on the choice of treatment depending on the patient, as well as potential treatment adaptation depending on the development of the disease. Due to this variation between treatment efficacies, regular monitoring of disease activity is all the more important, highlighting the value of an easily measurable biomarker such as sNfL.

A common endpoint of clinical trials in MS and aim of treatment in clinical practice is the state of NEDA-3 that describes a stable patient, based on clinical and radiologic criteria.[85–87] When

investigating the connection of sNfL levels with NEDA-3 in patients under DMT, our study showed a prognostic value of sNfL in NEDA-3 patients for future disease activity, which proposes superior sensitivity versus clinical and conventional MRI measures (chapter 3.1). Despite its value in predicting relapses, EDSS worsening or brain volume loss, attempts to investigate sNfL's prognostic potential for disease progression have so far been less successful. However, PIRA events may be preceded by increased sNfL levels 1 to 3 years earlier before overt disease progression is taking place (Abdelhak et al, under review). The difficulty to accurately capture our clinical phenotype outcomes may be an additional factor limiting such correlations or even prognostications.

In our second study, we focused on the capacity of sGFAP and sNfL in detecting PIRA in MS. Here, we examined whether sGFAP and sNfL have the ability to capture and prognosticate 'pure progression'/PIRA in MS patients. We found significantly higher levels of sGFAP, independent of other metrics, in relapse-free PMS patients with 'pure progression' compared to those who remained clinically stable. Conversely, while sGFAP was largely inert to changes of acute inflammatory activity, sNfL was strongly associated with this disease state. Increased sNfL levels were also related to continuous disease worsening, confirming earlier findings,[63,74] however this correlation was relatively weak and less prognostic for future PIRA than sGFAP. The association of sGFAP with PIRA as a reflection of chronic subclinical disease activity with astrocytic involvement hence may provide a valuable addition to sNfL. These results are currently under validation in large cohorts of the SMSC and Swedish MS Registry. At the same time and as a consequence of the clear age and interestingly sex association of sGFAP my group is currently also developing a large reference database for sGFAP.

Interestingly, in the worsening progression cohort of our study, baseline sGFAP was associated with gray matter (GM) volume loss, while baseline sNfL was associated with white matter (WM) volume loss. Previous studies examining the association of sNfL and sGFAP with brain atrophy in different brain compartments have not reached a clear consensus,[47,75,88–90] i.e. this needs to be further investigated to be better understood. Our preliminary results in our well characterized cohort of MS patients with very specific disease phenotypes provide a basis to better understand the dynamics of sNfL and sGFAP under different conditions of the disease, and provide the opportunity to compare with the course of pathology in MRI.

Additional to the MS patients we also investigated sGFAP in a cohort of healthy persons to determine the physiological levels and confounding factors of sGFAP. Similar to sNfL, the sGFAP levels in the control cohort were correlated with age and BMI. In addition, sGFAP

concentrations were higher in women than men, again highlighting the importance of including healthy persons as a comparison to the patient groups in order to prevent any skewing of our results in the patient cohort. A similar establishment of a RDB of GFAP values from healthy controls is the next step to bring GFAP closer to its potential introduction into clinical practice.

4.2 Summary and Outlook

In summary, this project has advanced sNfL considerably as a biomarker for MS disease activity on an individual patient level through the establishment of an RDB of control persons and the corresponding tool for calculation of sNfL percentiles and Z scores. Further we could show differential effects of DMT's through sNfL measurements and the benefit of combining sNfL with an additional biomarker, sGFAP, to prognosticate PIRA.

The establishment of the RDB has advanced the recognition of sNfL as a biomarker for MS in its potential clinical use on a larger scale. A further study in our group with our national collaborators from the SMSC aims to prospectively test the applicability and added value of sNfL measurements in a clinical setting by incorporating the sNfL measurements into the examinations of MS patients together with clinical and MRI tools. By regularly monitoring these classical parameters together with sNfL, the added value of sNfL may be more clearly established on an individual patient level. The RDB will serve as a basis for percentile and Z score calculations that will be used to interpret individual measuring results in clinical practice.

The measurements of sGFAP in healthy controls gave an important insight into confounding factors of sGFAP concentrations and showed the need to also here apply percentiles or Z scores to reach the most accurate and meaningful conclusions. Therefore, a database for the calculation of reliable Z scores or a larger reference database, alike to the one for sNfL, is currently being established. This is a necessary next step in the development of sGFAP as a biomarker complementary to sNfL. Ideally, these findings may add to the establishment of a range of biomarkers that could be combined and used to more precisely examine disease worsening in MS in the future. Furthermore, the clinical usefulness of sGFAP in capturing and prognosticating PIRA and the association of sGFAP with gray matter atrophy will be further investigated by measuring sGFAP in all available samples from the SMSC, allowing for an overarching analysis of sGFAP in a large cohort of MS patients.

References

1. Noseworthy, J., Lucchinetti, C., Rodriguez, M. & Weinshenker, B. Multiple Sclerosis. *N Engl J Med* **343**, 938–52 (2000).

2. Weinshenker, B. G. Epidemiology of multiple sclerosis. *Neurol. Clin.* **14**, 291–308 (1996).

3. Kappos, L. *et al.* Contribution of Relapse-Independent Progression vs Relapse-Associated Worsening to Overall Confirmed Disability Accumulation in Typical Relapsing Multiple Sclerosis in a Pooled Analysis of 2 Randomized Clinical Trials. *JAMA Neurol.* **77**, 1132–1140 (2020).

4. Lublin, F. D. *et al.* Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology* **83**, 278–86 (2014).

5. Rovaris, M. *et al.* Secondary progressive multiple sclerosis: Current knowledge and future challenges. *Lancet Neurology* vol. 5 343–354 (2006).

6. McDonald, W. I. *et al.* Recommended diagnostic criteria for multiple sclerosis: Guidelines from the International Panel on the Diagnosis of Multiple Sclerosis. *Ann. Neurol.* **50**, 121–127 (2001).

7. Lublin, F. D. & Reingold, S. C. Defining the clinical course of multiple sclerosis: Results of an international survey. *Neurology* **46**, 907–911 (1996).

8. Cree, B. A. C. *et al.* Long-term evolution of multiple sclerosis disability in the treatment era. *Ann. Neurol.* **80**, 499–510 (2016).

9. Cree, B. A. C. *et al.* Silent progression in disease activity–free relapsing multiple sclerosis. *Ann. Neurol.* **85**, 653–666 (2019).

10. Tremlett, H., Yinshan Zhao & Devonshire, V. Natural history of secondary-progressive multiple sclerosis. *Mult. Scler. J.* **14**, 314–324 (2008).

11. Giovannoni, G. *et al.* Smouldering multiple sclerosis: the 'real MS'. *Ther. Adv. Neurol. Disord.* **15**, 175628642110667 (2022).

12. Tur, C. *et al.* Association of Early Progression Independent of Relapse Activity with Long-term Disability after a First Demyelinating Event in Multiple Sclerosis. *JAMA*

Neurol. **80**, 151–160 (2023).

13. McGinley, M. P., Goldschmidt, C. H. & Rae-Grant, A. D. Diagnosis and Treatment of Multiple Sclerosis: A Review. *JAMA - J. Am. Med. Assoc.* **325**, 765–779 (2021).

14. Thompson, A. J. *et al.* Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol.* **17**, 162–173 (2018).

15. Link, H. & Huang, Y. M. Oligoclonal bands in multiple sclerosis cerebrospinal fluid: An update on methodology and clinical usefulness. *J. Neuroimmunol.* **180**, 17–28 (2006).

16. Khalil, M. *et al.* Neurofilaments as biomarkers in neurological disorders. *Nat. Rev. Neurol.* **14**, 577–589 (2018).

17. Yabe, J. T. *et al.* Neurofilaments consist of distinct populations that can be distinguished by C-terminal phosphorylation, bundling, and axonal transport rate in growing axonal neurites. *J. Neurosci.* **21**, 2195–2205 (2001).

18. Barro, C. *et al.* Fluid biomarker and electrophysiological outcome measures for progressive MS trials. *Mult. Scler.* **23**, 1600–1613 (2017).

19. Norgren, N. *et al.* Neurofilament and glial fibrillary acidic protein in multiple sclerosis. *Neurology* **63**, 1586–1590 (2004).

20. Kuhle, J. *et al.* A comparative study of CSF neurofilament light and heavy chain protein in MS. *Mult. Scler. J.* **19**, 1597–1603 (2013).

21. Lycke, J. N., Karlsson, J. E., Andersen, O. & Rosengren, L. E. Neurofilament protein in cerebrospinal fluid: A potential marker of activity in multiple sclerosis. *J. Neurol. Neurosurg. Psychiatry* **64**, 402–404 (1998).

22. Rissin, D. *et al.* Single-molecule enzyme-linked immunosorbent assay detects serum proteins at subfemtomolar concentrations. *Nat Biotechnol* **28**, 595–9 (2010).

23. Bacioglu, M. *et al.* Neurofilament Light Chain in Blood and CSF as Marker of Disease Progression in Mouse Models and in Neurodegenerative Diseases. *Neuron* **91**, 56–66 (2016).

24. Barro, C. *et al.* Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis. *Brain* **141**, 2382–2391 (2018).

25. Disanto, G. *et al.* Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann. Neurol.* **81**, 857–870 (2017).

26. Piehl, F. *et al.* Plasma neurofilament light chain levels in patients with MS switching from injectable therapies to fingolimod. *Mult. Scler. J.* **24**, 1046–1054 (2018).

27. Novakova, L. *et al.* Monitoring disease activity in multiple sclerosis using serum neurofilament light protein. *Neurology* **89**, 2230–2237 (2017).

28. Lu, C. H. *et al.* Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis. *Neurology* **84**, 2247–2257 (2015).

29. Gaiottino, J. *et al.* Increased Neurofilament Light Chain Blood Levels in Neurodegenerative Neurological Diseases. *PLoS One* **8**, 1–9 (2013).

30. Weston, P. S. J. *et al.* Serum neurofilament light in familial Alzheimer disease: A marker of early neurodegeneration. *Neurology* **89**, 2167–2175 (2017).

31. Shahim, P. *et al.* Serum neurofilament light protein predicts clinical outcome in traumatic brain injury. *Sci. Rep.* **6**, 1–9 (2016).

32. Oliver, J. M. *et al.* Serum Neurofilament Light in American Football Athletes over the Course of a Season. *J. Neurotrauma* **33**, 1784–1789 (2016).

33. Shahim, P., Zetterberg, H., Tegner, Y. & Blennow, K. Serum neurofilament light as a biomarker for mild traumatic brain injury in contact sports. *Neurology* **88**, 1788–1794 (2017).

34. Kuhle, J. *et al.* Serum neurofilament light chain is a biomarker of human spinal cord injury severity and outcome. *J. Neurol. Neurosurg. Psychiatry* **86**, 273–279 (2015).

35. Weydt, P. *et al.* Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis. *Ann. Neurol.* **79**, 152–158 (2016).

36. Kuhle, J. *et al.* Fingolimod and CSF neurofilament light chain levels in relapsing-remitting multiple sclerosis. *Neurology* **84**, 1639–1643 (2015).

37. Leppert, D. & Kuhle, J. Blood neurofilament light chain at the doorstep of clinical application. *Neurol. Neuroimmunol. NeuroInflammation* **6**, 4–6 (2019).

38. Bjornevik, K. *et al.* Serum Neurofilament Light Chain Levels in Patients with

Presymptomatic Multiple Sclerosis. *JAMA Neurol.* **77**, 58–64 (2020).

39. Calabresi, P. A. *et al.* Temporal profile of serum neurofilament light in multiple sclerosis: Implications for patient monitoring. *Mult. Scler. J.* 11–13 (2020).

40. Chitnis, T. *et al.* Neurofilament light chain serum levels correlate with 10-year MRI outcomes in multiple sclerosis. *Ann. Clin. Transl. Neurol.* **5**, 1478–1491 (2018).

41. Kuhle, J. *et al.* Serum neurofilament light chain in early relapsing remitting MS is increased and correlates with CSF levels and with MRI measures of disease severity. *Mult. Scler.* **22**, 1550–1559 (2016).

42. Kuhle, J. *et al.* Blood neurofilament light chain as a biomarker of MS disease activity and treatment response. *Neurology* **92**, E1007–E1015 (2019).

43. Siller, N. *et al.* Serum neurofilament light chain is a biomarker of acute and chronic neuronal damage in early multiple sclerosis. *Mult. Scler. J.* **25**, 678–686 (2019).

44. Jakimovski, D. *et al.* Serum neurofilament light chain levels associations with gray matter pathology: a 5-year longitudinal study. *Ann. Clin. Transl. Neurol.* **6**, 1757–1770 (2019).

45. Srpova, B. *et al.* Serum neurofilament light chain reflects inflammation-driven neurodegeneration and predicts delayed brain volume loss in early stage of multiple sclerosis. *Mult. Scler. J.* **27**, 52–60 (2021).

46. Kuhle, J. *et al.* Serum neurofilament is associated with progression of brain atrophy and disability in early MS. *Neurology* **88**, 826–831 (2017).

47. Plavina, T. *et al.* Association of Serum Neurofilament Light Levels with Long-term Brain Atrophy in Patients with a First Multiple Sclerosis Episode. *JAMA Netw. Open* **3**, 1–11 (2020).

48. Varhaug, K. N. *et al.* Neurofilament light chain predicts disease activity in relapsing-remitting MS. *Neurol. Neuroimmunol. NeuroInflammation* **5**, 1–7 (2018).

49. Ziemssen, T., Akgün, K. & Brück, W. Molecular biomarkers in multiple sclerosis. *J. Neuroinflammation* **16**, 1–11 (2019).

50. Altmann, P. *et al.* Serum neurofilament light chain withstands delayed freezing and repeated thawing. *Sci. Rep.* **10**, 1–8 (2020).

51. Teunissen, C. E., Tumani, H., Engelborghs, S. & Mollenhauer, B. Biobanking of CSF: International standardization to optimize biomarker development. *Clin. Biochem.* **47**, 288–292 (2014).

52. Khalil, M. *et al.* Serum neurofilament light levels in normal aging and their association with morphologic brain changes. *Nat. Commun.* **11**, 1–9 (2020).

53. Manouchehrinia, A. *et al.* Confounding effect of blood volume and body mass index on blood neurofilament light chain levels. *Ann. Clin. Transl. Neurol.* **7**, 139–143 (2020).

54. Hauser, S. L. *et al.* Five years of ocrelizumab in relapsing multiple sclerosis: OPERA studies open-label extension. *Neurology* **95**, e1854–e1867 (2020).

55. Kuhle, J. *et al.* Sustained reduction of serum neurofilament light chain over 7 years by alemtuzumab in early relapsing–remitting MS. *Mult. Scler. J.* 573–582 (2021).

56. Delcoigne, B. *et al.* Blood neurofilament light levels segregate treatment effects in multiple sclerosis. *Neurology* **94**, e1201–e1212 (2020).

57. Gunnarsson, M. *et al.* Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab. *Ann. Neurol.* **69**, 83–89 (2011).

58. Axelsson, M. *et al.* Immunosuppressive therapy reduces axonal damage in progressive multiple sclerosis. *Mult. Scler. J.* **20**, 43–50 (2014).

59. Petzold, A. Glial fibrillary acidic protein is a body fluid biomarker for glial pathology in human disease. *Brain Res.* **1600**, 17–31 (2015).

60. Lundgaard, I., Osório, M. J., Kress, B. T., Sanggaard, S. & Nedergaard, M. White matter astrocytes in health and disease. *Neuroscience* **276**, 161–173 (2014).

61. Petzold, A. *et al.* Markers for different glial cell responses in multiple sclerosis: Clinical and pathological correlations. *Brain* **125**, 1462–1473 (2002).

62. Rosengren, L. E., Lycke, J. & Andersen, O. Glial fibrillary acidic protein in CSF of multiple sclerosis patients: relation to neurological deficit. *J. Neurol. Sci.* **133**, 61–65 (1995).

63. Abdelhak, A., Huss, A., Kassubek, J., Tumani, H. & Otto, M. Serum GFAP as a biomarker for disease severity in multiple sclerosis. *Sci. Rep.* **8**, 1–7 (2018).

64. Aktas, O. *et al.* Serum Glial Fibrillary Acidic Protein: A Neuromyelitis Optica Spectrum Disorder Biomarker. *Ann. Neurol.* **89**, 895–910 (2021).

65. Watanabe, M. *et al.* Serum GFAP and neurofilament light as biomarkers of disease activity and disability in NMOSD. *Neurology* **93**, E1299–E1311 (2019).

66. Heller, C. *et al.* Plasma glial fibrillary acidic protein is raised in progranulin-associated frontotemporal dementia. *J. Neurol. Neurosurg. Psychiatry* **91**, 263–270 (2020).

67. Abu-Rumeileh, S. *et al.* CSF biomarkers of neuroinflammation in distinct forms and subtypes of neurodegenerative dementia. *Alzheimer's Res. Ther.* **12**, 1–15 (2019).

68. Abdelhak, A. *et al.* Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nat. Rev. Neurol.* **18**, 158–172 (2022).

69. US Food & Drug Administration. FDA authorizes marketing of first blood test to aid in the evaluation of concussion in adults. https://www.fda.gov/news-events/press-announcements/fda-authorizes-marketing-first-blood-test-aid-evaluation-concussion-adults (2018).

70. Axelsson, M. *et al.* Glial fibrillary acidic protein: A potential biomarker for progression in multiple sclerosis. *J. Neurol.* **258**, 882–888 (2011).

71. Burman, J. *et al.* Assessing tissue damage in multiple sclerosis: A biomarker approach. *Acta Neurol. Scand.* **130**, 81–89 (2014).

72. Mañé-Martínez, M. A. *et al.* Glial and neuronal markers in cerebrospinal fluid in different types of multiple sclerosis. *J. Neuroimmunol.* **299**, 112–117 (2016).

73. Thebault, S. *et al.* High or increasing serum NfL is predictive of impending multiple sclerosis relapses. *Mult. Scler. Relat. Disord.* **59**, 103535 (2022).

74. Högel, H. *et al.* Serum glial fibrillary acidic protein correlates with multiple sclerosis disease severity. *Mult. Scler. J.* **26**, 210–219 (2020).

75. Ayrignac, X. *et al.* Serum GFAP in multiple sclerosis: correlation with disease type and MRI markers of disease severity. *Sci. Rep.* **10**, 1–5 (2020).

76. Thelin, E. P. *et al.* Serial sampling of serum protein biomarkers for monitoring human traumatic brain injury dynamics: A systematic review. *Frontiers in Neurology* vol. 8 1–23 (2017).

77. Lublin, F. D., Coetzee, T., Cohen, J. A., Marrie, R. A. & Thompson, A. J. The 2013 clinical course descriptors for multiple sclerosis: A clarification. *Neurology* **94**, 1088–1092 (2020).

78. Benkert, P. *et al.* Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. *Lancet Neurol.* **21**, 246–257 (2022).

79. Quanterix Corp. Quanterix Granted Breakthrough Device Designation From U.S. FDA For NfL Test For Multiple Sclerosis. *22.04.2022* https://www.quanterix.com/press-releases/quanterix-granted-breakthrough-device-designation-from-us-fda-for-nfl-test-for-multiple-sclerosis/.

80. Simren, J. *et al.* Establishment of reference values for plasma neurofilament light based on healthy individuals aged 5-90 years. *Brain Commun.* **4**, 1–8 (2022).

81. Bornhorst, J. A. *et al.* Plasma neurofilament light chain (NfL) reference interval determination in an Age-stratified cognitively unimpaired cohort. *Clin. Chim. Acta* **535**, 153–156 (2022).

82. Fitzgerald, K. C. *et al.* Contributors to Serum NfL Levels in People without Neurologic Disease. *Ann. Neurol.* (2022) doi:10.1002/ana.26446.

83. Vermunt, L. *et al.* Age- and disease-specific reference values for neurofilament light presented in an online interactive support interface. *Ann. Clin. Transl. Neurol.* **9**, 1832–1837 (2022).

84. Andreasson, U. *et al.* Assessing the commutability of candidate reference materials for the harmonization of neurofilament light measurements in blood. *Clin. Chem. Lab. Med.* (2023) doi:10.1515/cclm-2022-1181.

85. Kappos, L. *et al.* Inclusion of brain volume loss in a revised measure of 'no evidence of disease activity' (NEDA-4) in relapsing-remitting multiple sclerosis. *Mult. Scler.* **22**, 1297–1305 (2016).

86. Rotstein, D. L., Healy, B. C., Malik, M. T., Chitnis, T. & Weiner, H. L. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. *JAMA Neurol.* **72**, 152–158 (2015).

87. Giovannoni, G. *et al.* Sustained disease-activity-free status in patients with relapsing-

remitting multiple sclerosis treated with cladribine tablets in the CLARITY study: a post-hoc and subgroup analysis. *Lancet Neurol.* **10**, 329–337 (2011).

88. Nyberg, L. *et al.* Elevated plasma neurofilament light in aging reflects brain white-matter alterations but does not predict cognitive decline or Alzheimer's disease. *Alzheimer's Dement. Diagnosis, Assess. Dis. Monit.* **12**, 1–9 (2020).

89. Papuć, E. & Rejdak, K. Increased CSF NFL in Non-demented Parkinson's Disease Subjects Reflects Early White Matter Damage. *Front. Aging Neurosci.* **12**, 1–6 (2020).

90. Saraste, M. *et al.* Increased serum glial fibrillary acidic protein associates with microstructural white matter damage in multiple sclerosis: GFAP and DTI. *Mult. Scler. Relat. Disord.* **50**, 102810 (2021).

Appendix A: Publications

** denotes equal contribution*

1. Leppert D, Watanabe M, Schaedelin S, Piehl F, Furlan R, Gastaldi M, Lambert J, Evertsson B, Fink K, Matsushita T, Masaki K, Isobe N, Kira J, Benkert P, Maceski A, Willemse E, Oechtering J, Orleth A, **Meier S**, Kuhle J. Granulocyte activation markers in cerebrospinal fluid differentiate acute neuromyelitis spectrum disorder from multiple sclerosis. *Submitted.*

2. Amrein M*, **Meier S***, Schäfer I, Schaedelin S, Willemse E, Benkert P, Walter J, Puelacher C, Zimmermann T, Median D, Egli C, Leppert D, Twerenbold R, Zellweger M, Kuhle J, Mueller C. Serum neurofilament light chain in functionally relevant coronary artery disease and adverse cardiovascular outcomes. Biomarkers. 2023 Feb 12:1-11. doi: 10.1080/1354750X.2023.2172211. Epub ahead of print. PMID: 36714921.

3. **Meier S***, Willemse EAJ*, Schaedelin S, Oechtering J, Lorscheider J, Melie-Garcia L, Cagol A, Barakovic M, Galbusera R, Subramaniam S, Barro C, Abdelhak A, Thebault S, Achtnichts L, Lalive P, Müller S, Pot C, Salmen A, Disanto G, Zecca C, D'Souza M, Orleth A, Khalil M, Buchmann A, Du Pasquier R, Yaldizli Ö, Derfuss T, Berger K, Hermesdorf M, Wiendl H, Piehl F, Battaglini M, Fischer U, Kappos L, Gobbi C, Granziera C, Bridel C, Leppert D, Maleska Maceski A, Benkert P, Kuhle J. Serum Glial Fibrillary Acidic Protein Compared With Neurofilament Light Chain as a Biomarker for Disease Progression in Multiple Sclerosis. *JAMA Neurol.* 2023;80(3):287–297. doi:10.1001/jamaneurol.2022.5250

4. Benkert P*, **Meier S***, Schaedelin S, Yaldizli Ö, Maceski A, Oechtering J, Achtnichts L, Aeschbacher S, Barakovic M, Buser A, Chan A, Conen D, Derfuss T, Disanto G, D'Souza M, Du Pasquier R, Findling O, Galbusera R, Hrusovsky K, Khalil M, Lalive PH, Lorscheider J, Mathias A, Mueller C, Müller S, Naegelin Y, Oksenberg JR, Orleth A, Pot C, Radue EW, Rahmanzadeh R, Salmen A, Sinnecker T, Subramaniam S, Vehoff J, Wellmann S, Wuerfel J, Zecca C, Willemse E, Blennow K, Zetterberg H, Gobbi C, Kappos L, Wiendl H, Berger K, Sormani MP, Granziera C, Leppert D, Kuhle J, for the Swiss Multiple Sclerosis Cohort Study. Serum neurofilament light chain for individual prognostication of disease activity in people with multiple sclerosis: a retrospective modelling and validation study. *Lancet Neurol* 2022; 21: 246–57.

5. Wai CH, Jin J, Cyrklaff M, Genoud C, Funaya C, Sattler J, Maceski A, **Meier S**, Heiland S, Lanzer M, Frischknecht F, Kuhle J, Bendszus M, Hoffmann A. Neurofilament light chain plasma levels are associated with area of brain damage in experimental cerebral malaria. Sci

Rep. 2022 Jun 24;12(1):10726. doi: 10.1038/s41598-022-14291-x. PMID: 35750882; PMCID: PMC9232608.

6. Uyar M, Lezius S, Buhmann C, Pötter-Nerger M, Schulz R, **Meier S**, Gerloff C, Kuhle J, Choe CU. Diabetes, Glycated Hemoglobin (HbA1c), and Neuroaxonal Damage in Parkinson's Disease (MARK-PD Study). Mov Disord. 2022 Jun;37(6):1299-1304. doi: 10.1002/mds.29009. Epub 2022 Apr 6. PMID: 35384057.

7. Oechtering J, Lincke T, Schaedelin S, Décard BF, Maceski A, Orleth A, **Meier S**, Willemse E, Buchmann A, Khalil M, Derfuss T, Benkert P, Heijnen I, Regeniter A, Müller S, Achtnichts L, Lalive P, Salmen A, Pot C, Gobbi C, Kappos L, Granziera C, Leppert D, Schlaeger R, Lieb JM, Kuhle J; Swiss MS Cohort Study. Intrathecal IgM Synthesis Is Associated with Spinal Cord Manifestation and Neuronal Injury in Early MS. Ann Neurol. 2022 Jun;91(6):814-820. doi: 10.1002/ana.26348. Epub 2022 Apr 9. PMID: 35293622; PMCID: PMC9320956.

8. Oechtering J, Schaedelin S, Benkert P, Müller S, Achtnichts L, Vehoff J, Disanto G, Findling O, Fischer-Barnicol B, Orleth A, Chan A, Pot C, Barakovic M, Rahmanzadeh R, Galbusera R, Heijnen I, Lalive P, Wuerfel J, Subramaniam S, Aeschbacher S, Conen D, Naegelin Y, Maceski A, **Meier S**, Berger K, Wiendl H, Lincke T, Lieb J, Yaldizli Ö, Sinnecker T, Derfuss T, Regeniter A, Zecca C, Gobbi C, Kappos L, Granziera C, Leppert D, Kuhle J. Intrathecal Immunoglobulin M Synthesis is an Independent Biomarker for Higher Disease Activity and Severity in Multiple Sclerosis. *Ann Neurol* 2021; 90: 477–89.

9. Uher T, Kubala Havrdova E, Benkert P, Bergsland N, Krasensky J, Srpova B, Dwyer M, Tyblova M, **Meier S**, Vaneckova M, Horakova D, Zivadinov R, Leppert D, Kalincik T, Kuhle J. Measurement of neurofilaments improves stratification of future disease activity in early multiple sclerosis. *Mult Scler J* 2021; 27: 2001–13.