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

eMethods

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

		sGFAP (pg/ml), median, IQR	Est.	95%CI	*p*	sNfL (pg/ml), median, IQR	Est.	95%CI	*p*
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)		**R^2=0.133***	-	**-0.27**	**-0.44--0.11**	**0.001**	**R^2=0.018****	-	-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)		0.6 [-0.3-2.1]	-0.50	-1.11-0.12	0.11		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)

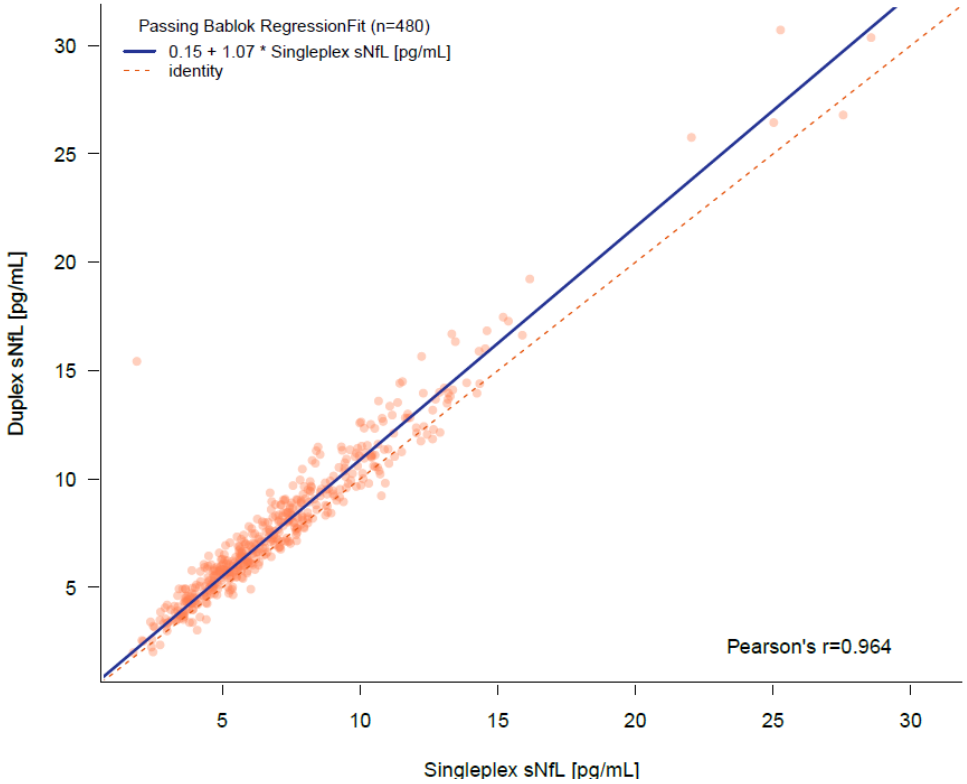
N=252 patients		**Variance explained**	**sGFAP conc. (pg/mL), median, IQR**	**Est.**	**95%CI**	***p***	**Variance explained**	**sNfL conc. (pg/mL), median, IQR**	**Est.**	**95%CI**	***p***
Age (per 10 years)		R^2= 0.251 *		**1.10**	**1.04-1.17**	**0.002**	R^2=0.293 **		**1.22**	**1.15-1.30**	**<0.001**
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).

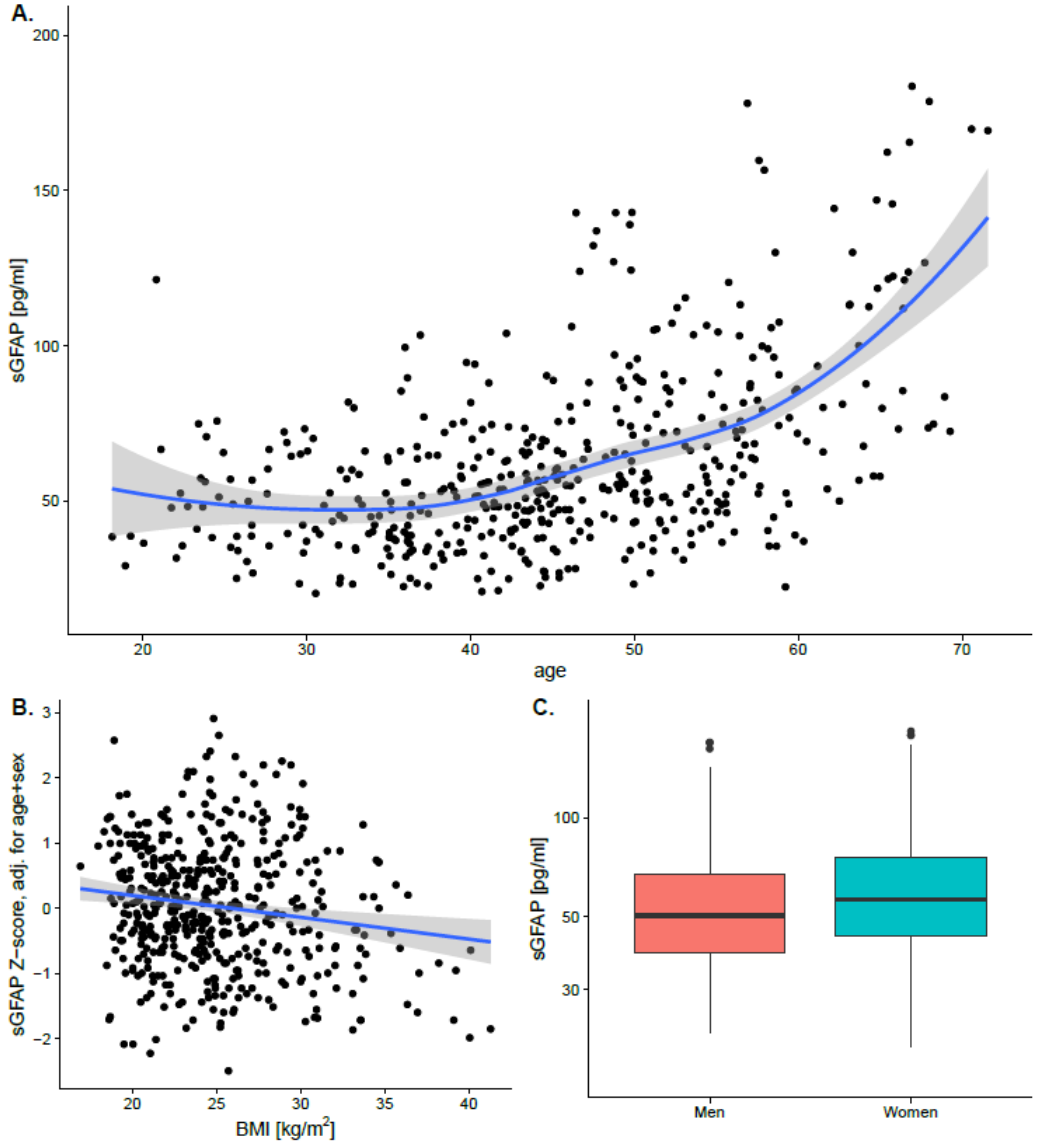
Abbreviations: BMI: body mass index; CDW: confirmed disease worsening; DMT: disease; EDSS: Expanded Disability Status Scale score; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain; modifying treatment; IQR: interquartile range; OCR: ocrelizumab; PPMS: primary progressive 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)



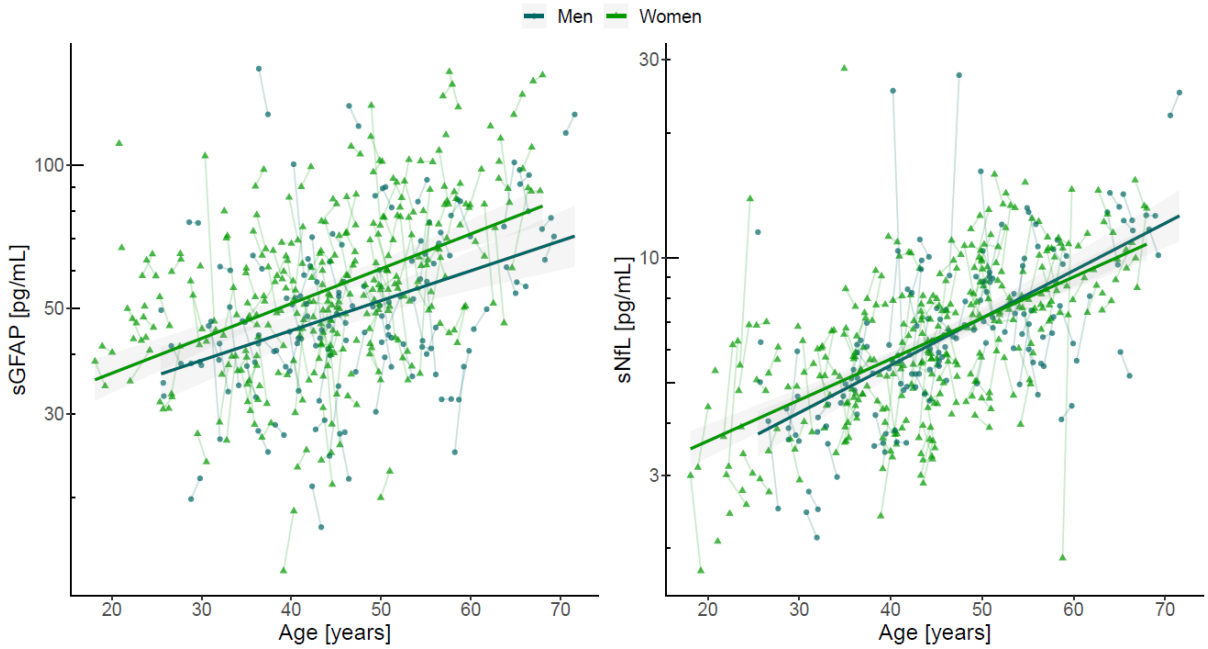
Legend: Each dot indicates an individual data point. The solid line indicates the Passing-Bablok regression line. The dotted line indicates the x=y identity line. Parallel comparison of sNfL results measured with the Nf-Light kit and the Neurology 2-plex B assay showed excellent congruency (Pearson's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



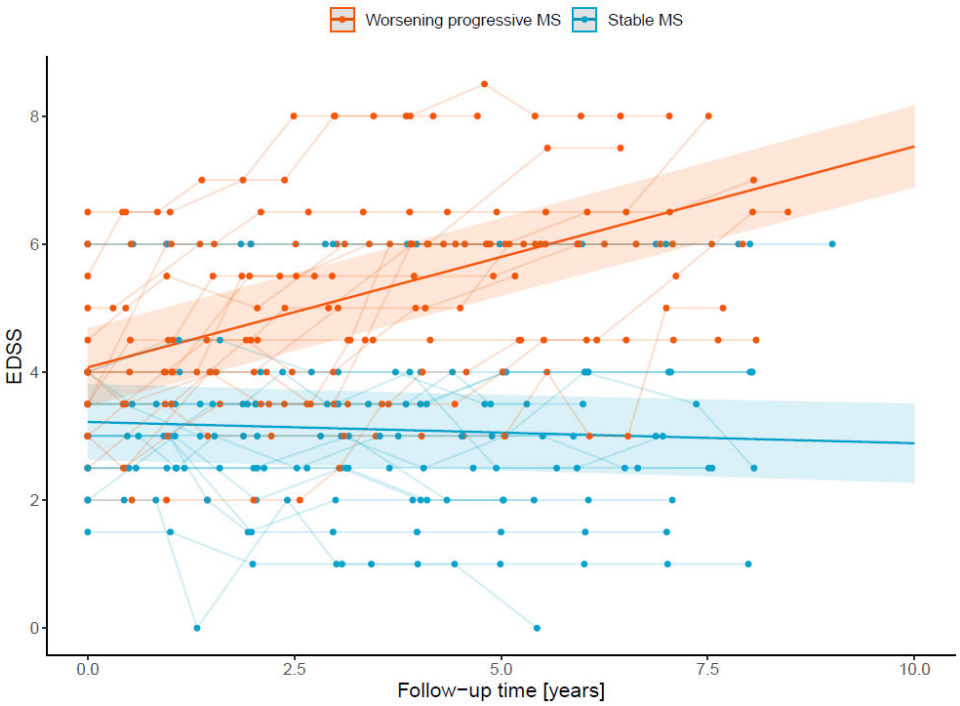
Legend: Graphical representation of the associations between sGFAP and age (A.), BMI (B.) as well as sex (C.): sGFAP increases with age in a non-linear manner (line represents a non-linear smoothing function with confidence band (A.) and a linear regression line with confidence band (B.)), decreases 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



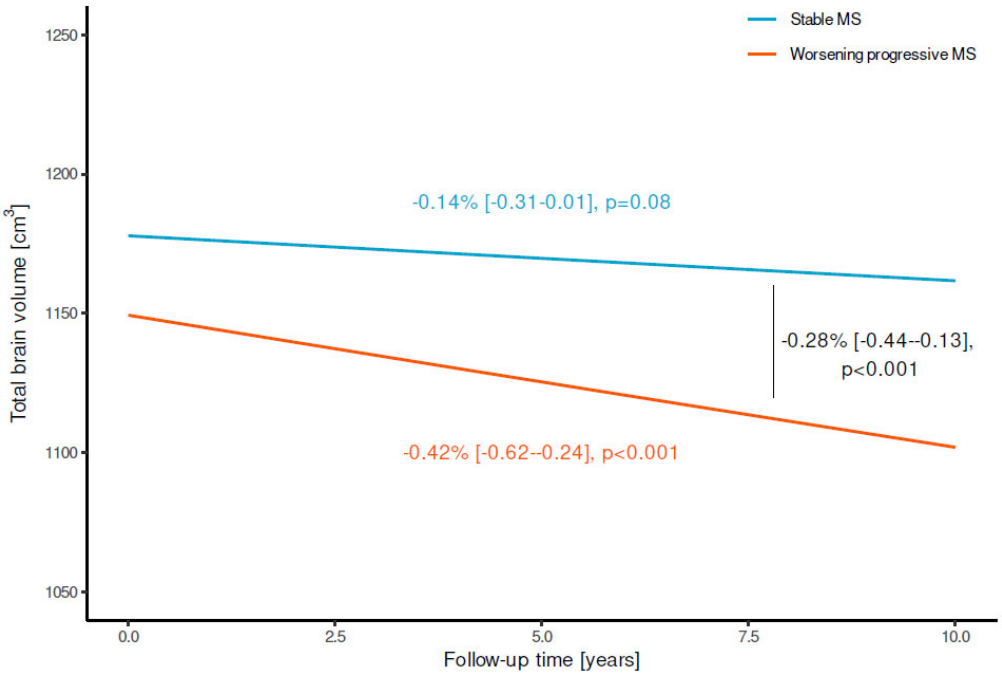
Legend: sGFAP (left) and sNfL (right) concentrations in samples from healthy controls (259 at baseline and 226 at follow-up) in relation to age, stratified for sex (men represented by blue circles; women by green triangles). Samples from one individual are connected through lines; thick lines show the 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



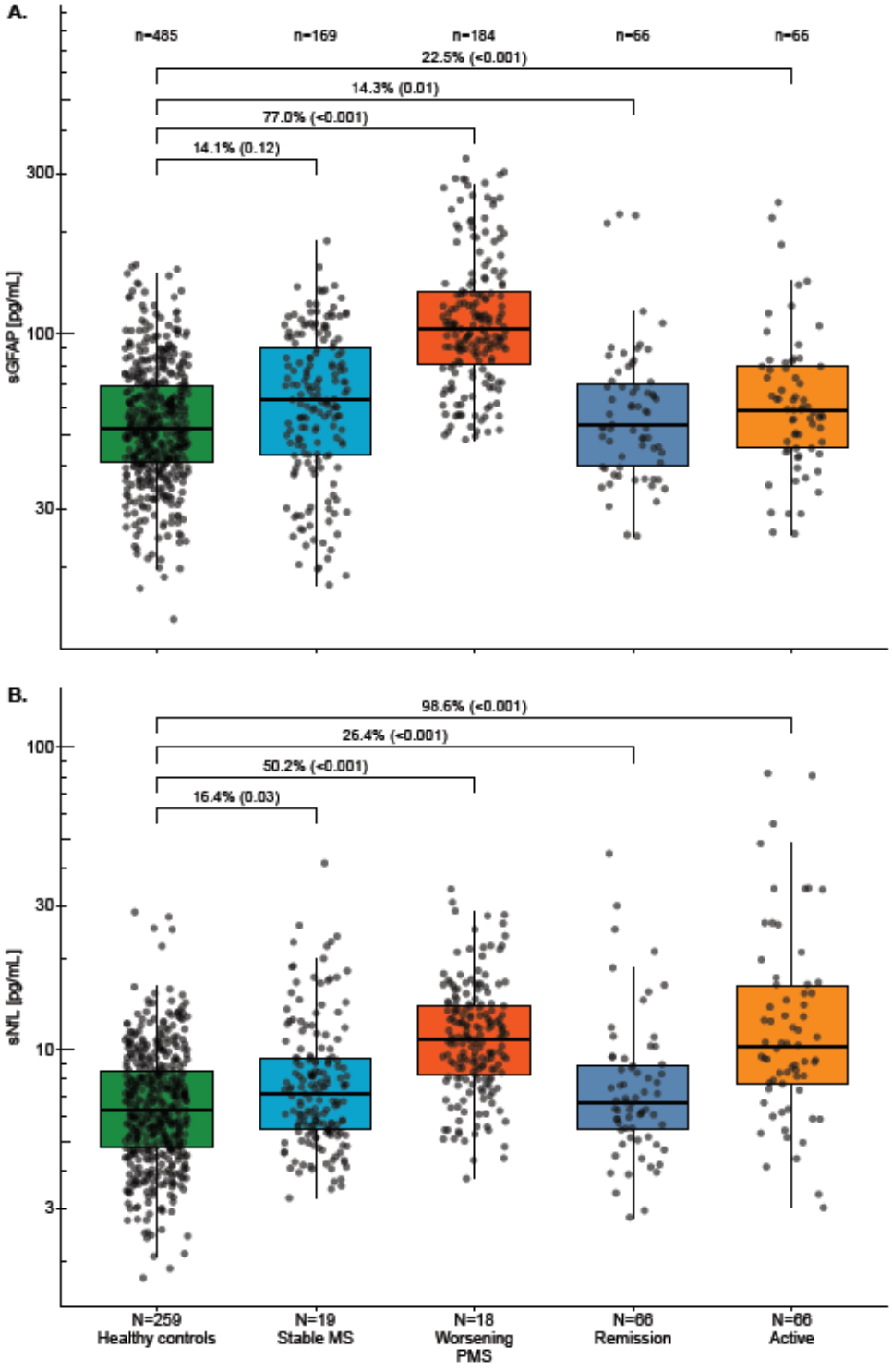
Legend: Patients with worsening progressive MS (red) showed an increase in EDSS score while stable patients (blue) maintain stable EDSS scores. Thin lines connect individual data points; thick lines including 95% CI show marginal effects from a mixed model with EDSS explained by an interaction term between 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

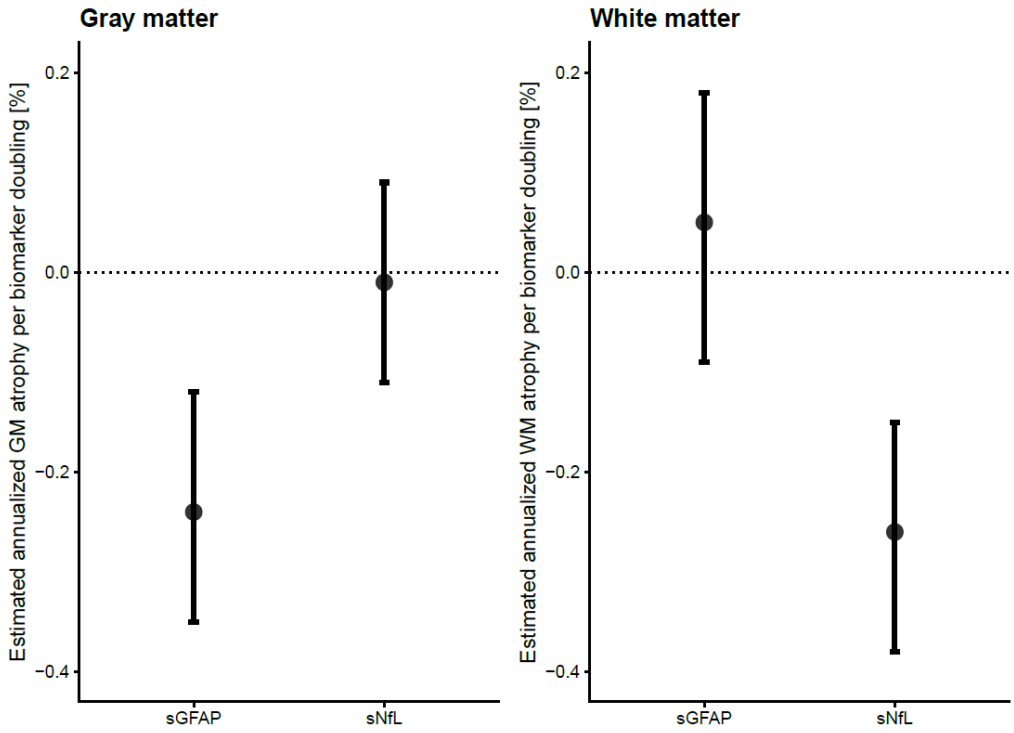
eFigure 6. Serum GFAP (A) and sNfL (B) in Different MS Groups vs Healthy Controls



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.

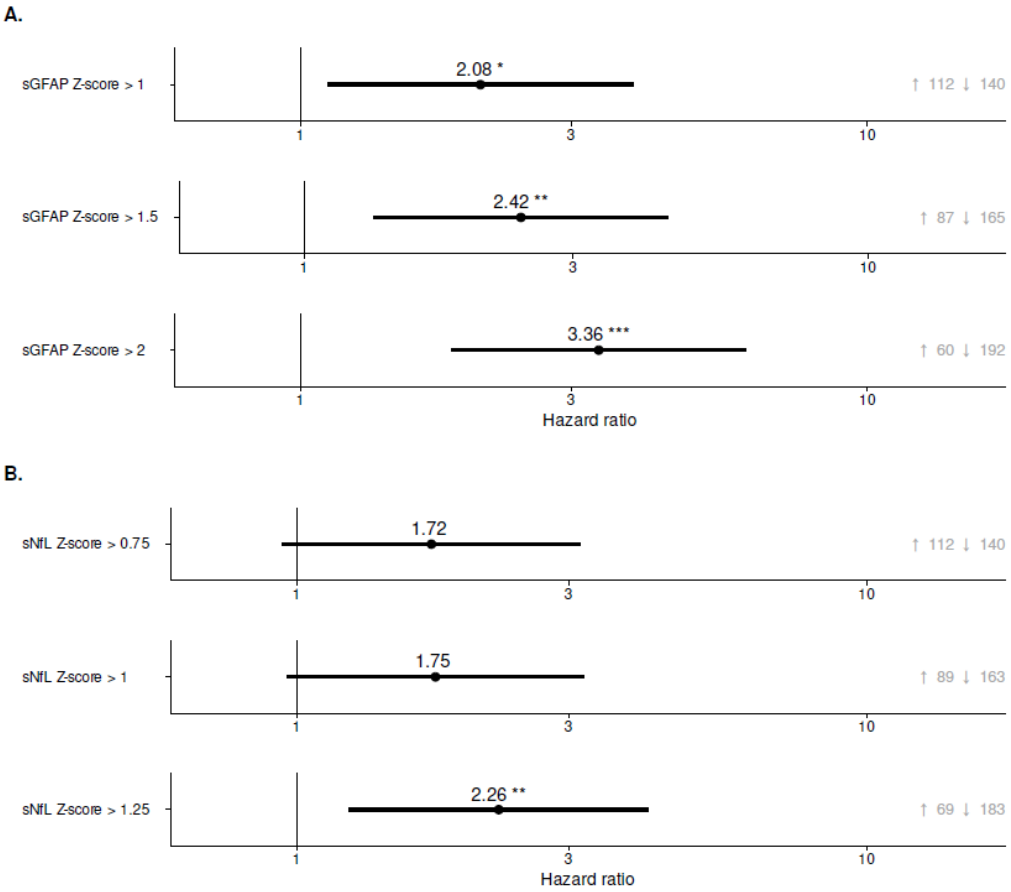
Abbreviations: PMS: progressive MS; sGFAP: serum glial fibrillary acidic protein; sNfL: serum neurofilament light chain.

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.