Title

Serum neurofilament light chain levels at attack predict post-attack disability worsening and are mitigated by inebilizumab: analysis of four potential biomarkers in neuromyelitis optica spectrum disorder

Authors

Orhan Aktas[1], Hans-Peter Hartung[1,2,3,4], Michael A Smith[5], William A Rees[5], Kazuo Fujihara[6,7], Friedemann Paul[8], Romain Marignier[9], Jeffrey L Bennett[10], Ho Jin Kim[11], Brian G Weinshenker[12], Sean J Pittock[13], Dean M Wingerchuk[14], Gary Cutter[15], Dewei She[5], Michele Gunsior[5ǂ], Daniel Cimbora[5], Eliezer Katz[5ǂ], and Bruce AC Cree[16] on behalf of the N-MOmentum study investigators

Affiliations

[1]Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; [2]Brain and Mind Centre, University of Sydney, Sydney, NSW, Australia; [3]Department of Neurology, Medical University Vienna, Vienna, Austria; [4]Department of Neurology, Palacky University in Olomouc, Olomouc, Czech Republic; [5]Horizon Therapeutics, Gaithersburg, MD, USA; [6]Department of Multiple Sclerosis Therapeutics, Fukushima Medical University; [7]Multiple Sclerosis and Neuromyelitis Optica Center, Southern Tohoku Research Institute for Neuroscience, Koriyama, Japan; [8]Experimental and Clinical Research Center, Max Delbrück Center for Molecular Medicine and Charité - Universitätsmedizin Berlin, Berlin, Germany; [9]Centre de Référence des Maladies Inflammatoires Rares du Cerveau et de la Moelle (MIRCEM), Service de Neurologie, Sclérose en Plaques, Pathologies de la Myéline et Neuro-inflammation, Hôpital Neurologique Pierre Wertheimer, Hospices Civils de Lyon, Lyon, France; [10]University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, CO, USA; [11]Research Institute and Hospital of National Cancer Center, Department of Neurology, Goyang, Republic of Korea; [12]Department of Neurology, University of Virginia, Charlottesville, VA, USA; [13]Department of Neurology and Center for MS and Autoimmune Neurology, Mayo Clinic, Rochester, MN, USA; [14]Department of Neurology, Mayo Clinic, Scottsdale, AZ, USA; [15]The University of Alabama at Birmingham, Birmingham, AL, USA; [16]UCSF Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, CA, USA. [ǂ]Former employee of Horizon Therapeutics

Aktas O, *et al. J Neurol Neurosurg Psychiatry* 2023; 94:757–768. doi: 10.1136/jnnp-2022-330412

Aktas O, *et al. J Neurol Neurosurg Psychiatry* 2023; 94:757–768. doi: 10.1136/jnnp-2022-330412

SUPPLEMENT

Supplementary methods

Study design and participants

The N-MOmentum study (NCT02200770) was an international, multicentre, randomised, double-blind, placebo-controlled phase 2/3 trial with an optional open-label extension phase (OLP), full details of which have been published previously. In brief, adults with NMOSD, an Expanded Disability Status Scale (EDSS) score of 8.0 or less, and a history of either at least one NMOSD attack in the previous year or at least two attacks in the previous 2 years who were seropositive for immunoglobulin G autoantibodies to aquaporin-4 (AQP4+) or seronegative (AQP4−) were eligible. Participants who were AQP4− were assessed at Mayo Clinic Laboratories for presence of myelin oligodendrocyte glycoprotein-IgG (MOG-IgG) by cell-based assay. The primary endpoint was the time to an adjudicated NMOSD attack during the randomised controlled period (RCP), with attack status defined by protocol-defined criteria.

Statistical analyses

The current set of analyses are exploratory; p values are provided for hypothesis generation only and are not adjusted for multiple comparisons. The utility of biomarker concentrations at baseline and at time points during the RCP as a predictor of future attack risk was assessed using multivariate Cox proportional hazards regression, with placebo as reference group and with treatment and baseline EDSS score as explanatory factors. Biomarker concentrations at each scheduled draw were treated as time-varying covariates in the Cox regression model. The Wilcoxon signed-rank test was used to evaluate statistical significance of increases of each biomarker from each time point to attack in paired samples.

A mixed-effects logistic regression model was used to evaluate the significance of elevation in biomarker concentration in attack samples versus those drawn at scheduled visits. Baseline biomarker concentration, age, and baseline EDSS score were included as covariates with a random intercept term for each week of scheduled draws. Given that EDSS assessments were conducted less frequently than biomarker assessments, ‘last observation carried forward’ was used to impute EDSS scores at each scheduled biomarker assessment. The area under the curve (AUC) was

Aktas O, *et al. J Neurol Neurosurg Psychiatry* 2023; 94:757–768. doi: 10.1136/jnnp-2022-330412

predicted from 10 iterations of fivefold cross-validation. AUCs were evaluated between the full model and serum glial fibrillary acidic protein (sGFAP) alone as predictors, and a sensitivity analysis was conducted to assess the performance of the model across participants in different treatment arms and in those who did or did not experience attacks during the RCP.

Correlation between changes in EDSS scores, EDSS component scores, and biomarker concentrations from baseline to attack were evaluated using Spearman's Rho. Multiple linear regression was used to assess independent correlation of each biomarker with EDSS score change at attack and proceeding attack after controlling for baseline EDSS score and age. The Mann–Whitney U test, estimated from 500 iterations of bootstrap resampling (pROC package), was used to evaluate significance of differences further in biomarker concentrations between those who experienced protocol-defined EDSS score worsening and those who experienced no worsening of EDSS score.

The significance of changes in biomarker concentrations from baseline to attack were evaluated in both RCP treatment groups using the Wilcoxon signed-rank test. Fold changes from baseline between treatment groups and in participants who did or did not experience attacks were assessed for significance using the Mann–Whitney U test. Serum neurofilament light chain (sNfL) concentrations between treatment groups were also assessed using a mixed linear model, including baseline sNfL, EDSS score and age as covariates and a per-subject random intercept term. Significance of sNfL changes in response to treatment was assessed on the interaction term of treatment and individual time points using the lmerTest package. Significance of the interaction term was also assessed through a likelihood ratio test. All statistical analysis was performed in R 4.1.3.

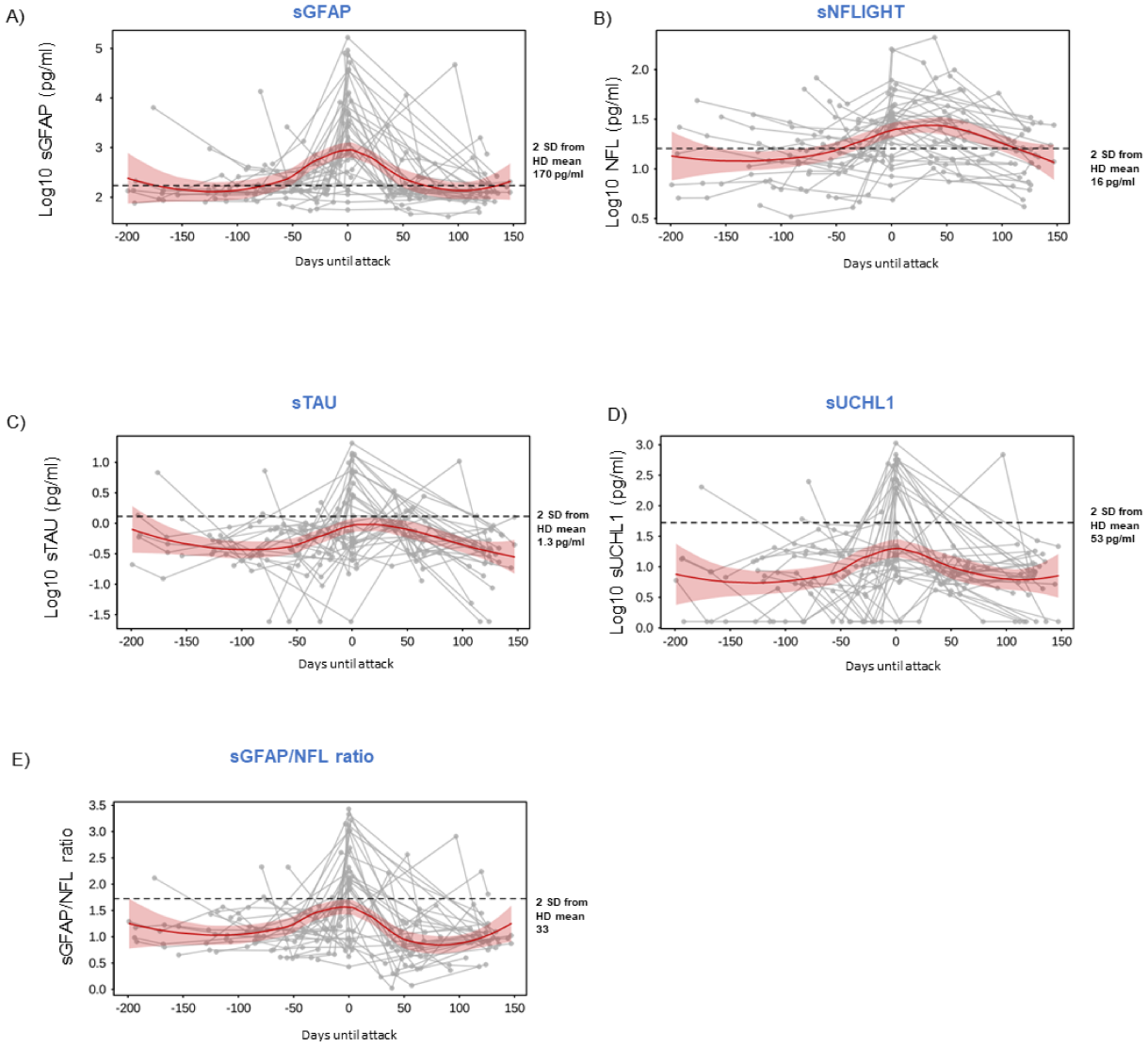
Data availability

Data from the study will be made available to others in accordance with the other elements of this statement. Horizon is committed to responsibly sharing data from the clinical trials we sponsor, provided that the trials are not part of an ongoing or planned regulatory submission (including requests for clinical trial data for unlicensed products and indications). Access to anonymised

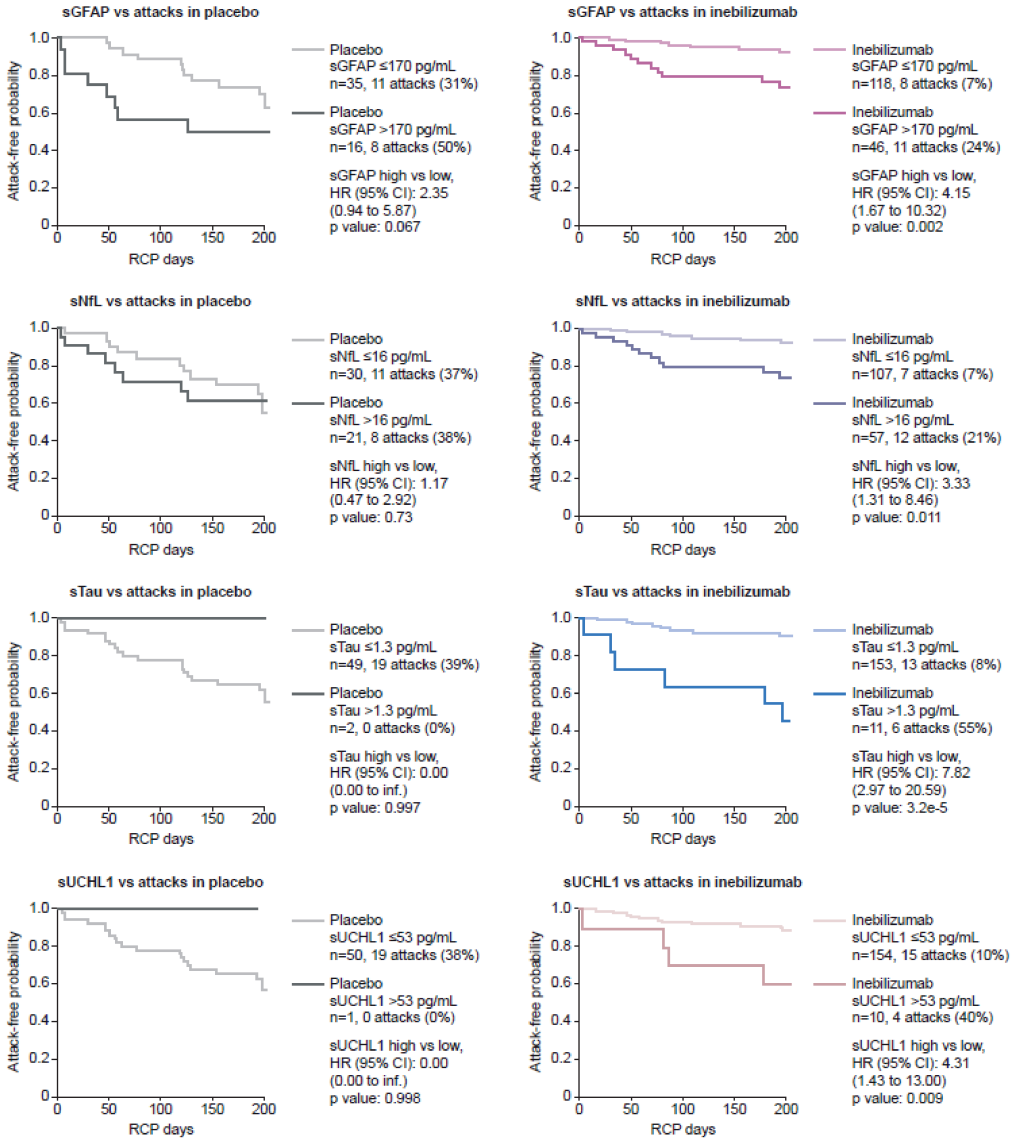
individual and trial-level data (analysis datasets) may be granted. Clinical trial data may be requested by submitting a research proposal and statistical analysis plan. Data will be provided following review and approval of the plan, and execution of a data sharing agreement. For more information, or to submit a request, please email: medicalinformation@horizontherapeutics.com.

Aktas O, *et al. J Neurol Neurosurg Psychiatry* 2023; 94:757–768. doi: 10.1136/jnnp-2022-330412

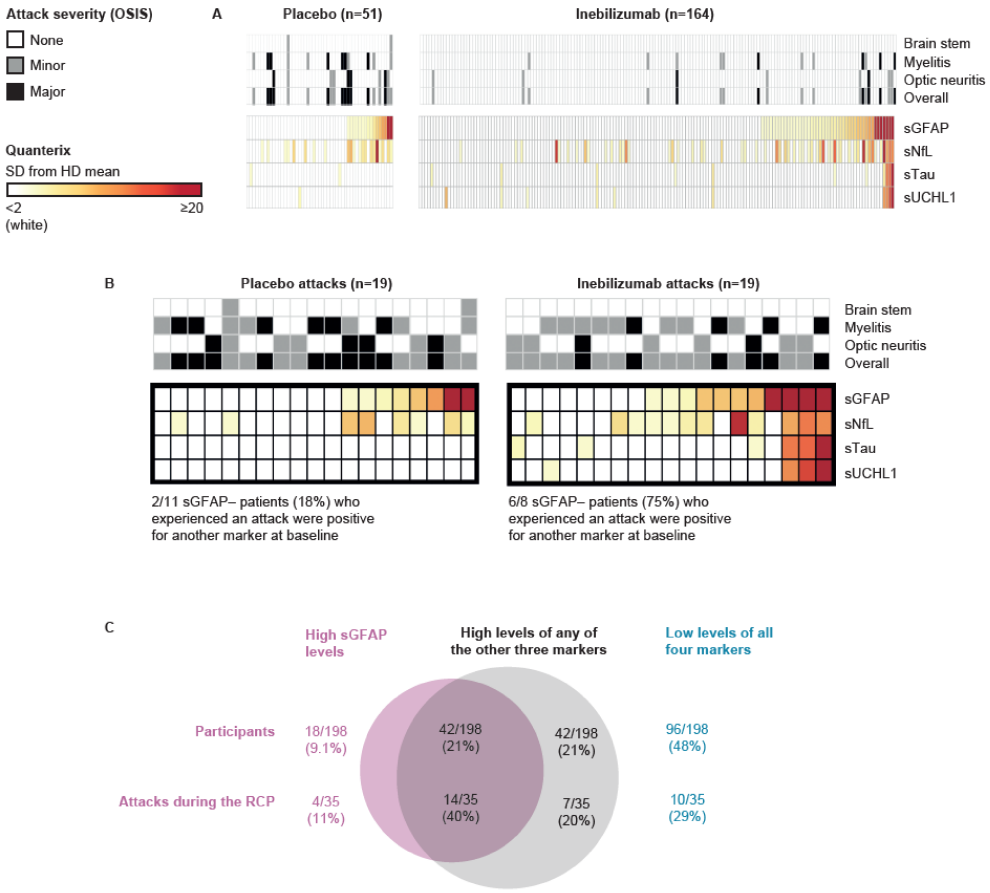
Figures



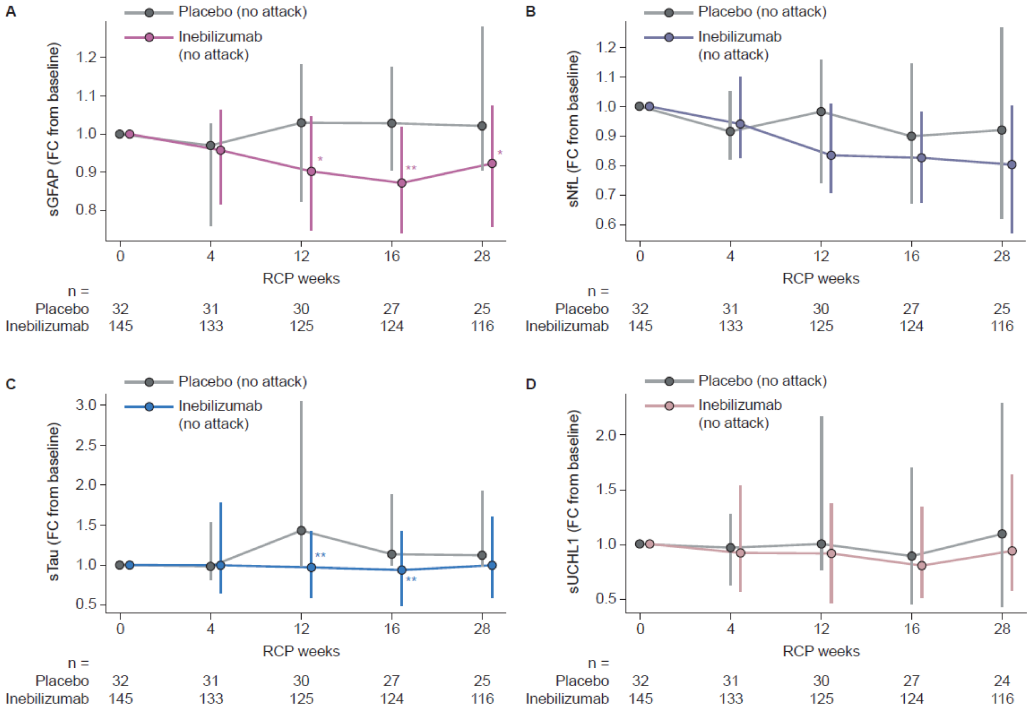
efigure 1 Lineplots displaying sGFAP (A) sNfL (B) sTAU (C) sUCHL1 (D) sGFAP:NfL ratio (E) leading up to NMOSD attack and in the days after the attack. Each gray line represents a single patient's longitudinal profile. Red line represents smooth curve +/- 95% CI estimated through LOESS regression (span = 0.75, degree=2)



efigure 2 Attack-free survival by treatment group according to cut-off values for CNS damage biomarkers. Graphs show Kaplan−Meier plots of time to first adjudicated attack during the RCP in AQP4+ participants. Cut-offs were two SDs from the healthy donor mean after trimming the most extreme value for each biomarker from the SD and mean. AQ4P+, seropositive for immunoglobulin G autoantibodies to aquaporin-4; CI, confidence interval; CNS, central nervous system; HR, hazard ratio; RCP, randomised controlled period; SD, standard deviation; sGFAP, serum glial fibrillary acidic protein; sNfL, serum neurofilament light chain; sTau, serum tau; sUCHL1, serum ubiquitin C-terminal hydrolase L1.

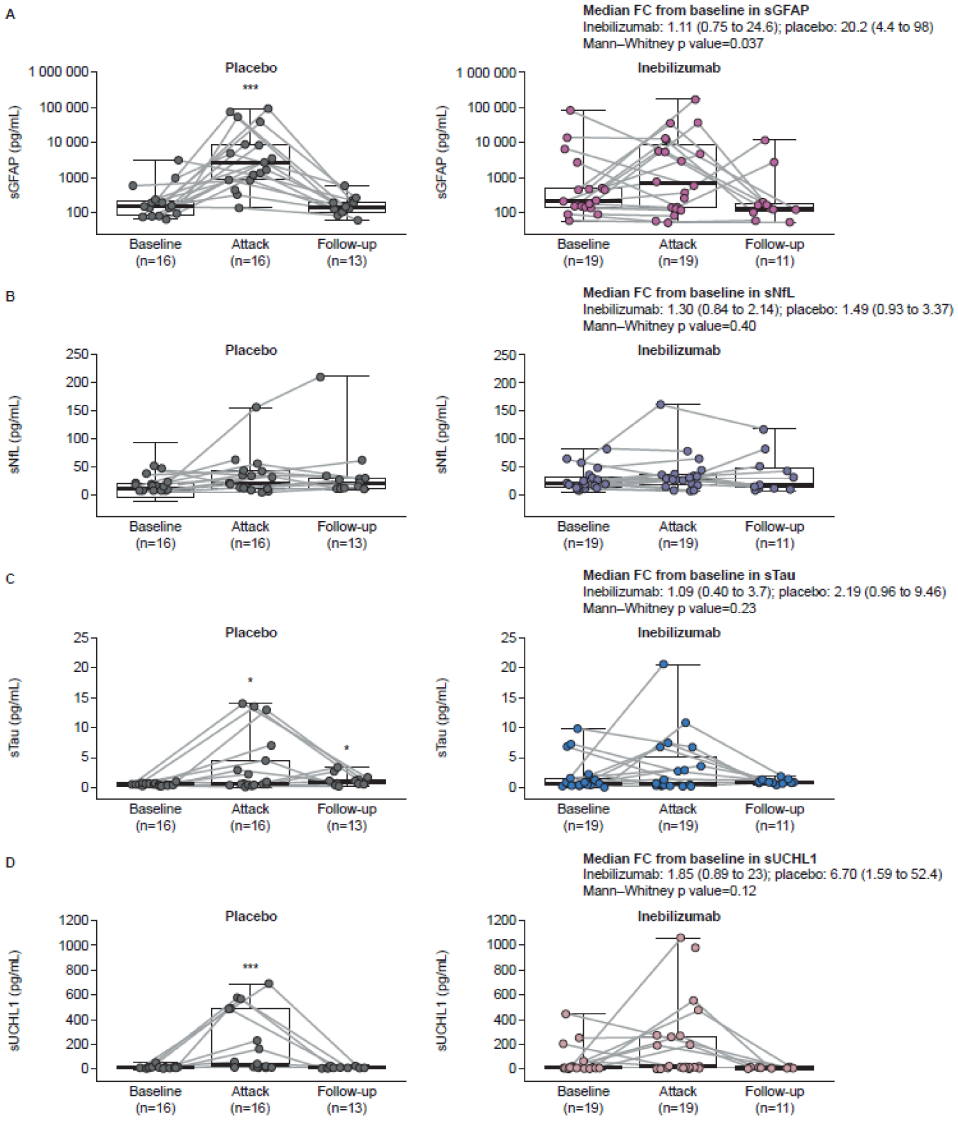


efigure 3. Heatmaps of day1/RCP CNS biomarker concentrations displaying overlap of biomarkers. A) Heatmap displaying day 1/RCP serum CNS damage biomarker concentrations relative to HDs in all inebilizumab- and placebo-treated participants. B) Heatmap displaying day 1/RCP biomarker concentrations only in those who later experienced attacks during the RCP. Colour bar coloured by attack severity as determined by OSIS. Heatmaps ordered by sGFAP concentration in study participants. C) Venn diagram displaying overlap between participants 'high' for sGFAP relative to any of the other 3 markers. Bottom row of text displays the proportion of attacks that occurred within each baseline subgroup. CNS, central nervous system; HD, healthy donor; OSIS, Opticospinal Impairment Scale; RCP, randomised controlled period; SD, standard deviation; sGFAP, serum glial fibrillary acidic protein; sNfL, serum neurofilament light chain; sTau, serum tau; sUCHL1, serum ubiquitin C-terminal hydrolase L1.



efigure 4. Change of different biomarkers over the RCP in participants without attacks

Quanterix CNS damage biomarkers. A) sGFAP, B) sNfL, C) sTau, and D) sUCHL1 measured at regular intervals during the RCP. Significance of changes between treatment groups assessed using Mann-Whitney U test comparing FC from baseline calculated from each treatment group. p<0.10; *p<0.05; **p<0.01; ***p<0.001. CNS, central nervous system; FC, fold change; RCP, randomised controlled period; sGFAP, serum glial fibrillary acidic protein; sTau, serum tau; sUCHL1, serum ubiquitin C-terminal hydrolase L1.



efigure 5 Change in concentrations of sGFAP, sNfL, sTau, and sUCHL1 across attacks. Biomarker concentrations measured by Quanterix on day 1/RCP, samples proximal to attack (+/− 7 days) and following the attack in inebilizumab-treated and placebo-treated participants. A) sGFAP, B) sNfL, C) sTau, and (D) sUCHL1. Difference between both treatment groups estimated using Fisher's exact test. Significance of changes from baseline in each treatment group assessed using Wilcoxon-signed rank test. p<0.10; *p<0.05; **p<0.01; ***p<0.001. Significance of changes between treatment groups assessed using Mann-Whitney U test comparing FC from baseline calculated from each treatment group. FC, fold change; RCP, randomised controlled period; sGFAP, serum glial fibrillary acidic

protein; sNfL, serum neurofilament light chain; sTau, serum tau; sUCHL1, serum ubiquitin C-terminal hydrolase L1.

eTable 1 Mixed-effects logistic regression model for identifying attacks from samples drawn at scheduled visits.

Full model:

m1 = glmer(attack ~ GFAP + `Day 1 GFAP`+ NFLIGHT + `Day 1 NFLIGHT`+ TAU + `Day 1 TAU`+ UCHL1+ `Day 1 UCHL1`+ age+edss+trx+(1|RCPweeks))

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.63	1.32	-2.75	5.95E-03
GFAP	1.61	0.39	4.16	3.17E-05
Day 1/RCP GFAP	0.05	0.33	0.16	8.75E-01
NFLIGHT	1.10	0.45	2.42	1.54E-02
Day 1/RCP NFLIGHT	-1.50	0.48	-3.15	1.65E-03
TAU	-0.86	0.39	-2.19	2.85E-02
Day 1/RCP TAU	1.09	0.43	2.53	1.13E-02
UCHL1	0.10	0.49	0.21	8.37E-01
Day 1/RCP UCHL1	0.04	0.45	0.09	9.28E-01
Age	-0.01	0.28	-0.03	9.76E-01
EDSS	0.05	0.36	0.14	8.89E-01
Inebilizumab treatment	-0.47	0.69	-0.68	4.95E-01

Reduced model:

m2 = glmer(attack ~ GFAP + `Day 1 GFAP+age+edss+trx+(1|RCPweeks))

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.97	1.06	-2.79	5.19E-03
GFAP	1.23	0.19	6.62	3.52E-11
Day 1/RCP GFAP	0.35	0.19	1.83	6.70E-02
Age	-0.14	0.23	-0.63	5.31E-01
EDSS	-0.11	0.29	-0.37	7.10E-01
Inebilizumab treatment	-0.91	0.59	-1.54	1.25E-01

Likelihood ratio test on nested models

	logLik	Chisq	Df	Pr(>Chisq)
Reduced Model	-67.52			
Full model	-59.56	15.90	6	0.014

Mixed-effect logistic regression models for identifying samples drawn during attack relative to samples drawn during scheduled draws. CNS damage markers were sampled at day 1/RCP, and at weeks 4, 12, 16, and 28 of the RCP, along with attack assessments. EDSS was conducted at day 1, and weeks 12 and 28 of the RCP, along with attack assessments, and was imputed as LOCF for time points with biomarker measurements, but without an EDSS assessment. EDSS score was also imputed as LOCF during attack assessments because EDSS score was included in the attack definition, thus collinear with attack status at that time point.

CNS, central nervous system; EDSS, Expanded Disability Status Scale; GFAP, serum glial fibrillary acidic protein; LOCF, last observation carried forward; NFLIGHT, neurofilament light chain; Tau, serum tau; RCP, randomised controlled period; UCHL1, serum ubiquitin C-terminal hydrolase L1.

Aktas O, *et al. J Neurol Neurosurg Psychiatry* 2023; 94:757–768. doi: 10.1136/jnnp-2022-330412

eTable 2 Cox regression model using Quanterix 4-plex as time-varying covariates.

Full model:

coxph(Surv(tstart, tstop, attack) ~Age+ Treatment+ `Previous Attack`+EDSS+log(GFAP)+log(NFL)+log(TAU)+log(UCHL1), cluster = subject)

Variable	HR (-/+ 95% CI)	p-value
Age Years	0.99 (0.96,1.02)	4.74e-01
Inebizumab Treatment	***0.28 (0.13,0.59)***	***9.21e-04***
Previous attack w/in 90 days	0.49 (0.14,1.76)	2.75e-01
Baseline EDSS	0.9 (0.71,1.15)	4.01e-01
log GFAP concentration	***2.56 (1.7,3.86)***	***6.81e-06***
log NFL concentration	1.28 (0.67,2.46)	4.53e-01
log TAU concentration	1.02 (0.63,1.65)	9.28e-01
log UCHL1 concentration	0.78 (0.52,1.19)	2.53e-01

Full model:

coxph(Surv(tstart, tstop, attack) ~Age+ Treatment+ `Previous Attack`+EDSS+log(GFAP), cluster = subject)

Variable	HR (-/+ 95% CI)	p-value
Age Years	0.99 (0.96,1.02)	5.50e-01
Inebizumab Treatment	***0.29 (0.13,0.62)***	***1.49e-03***
Previous attack w/in 90 days	0.48 (0.14,1.7)	2.56e-01
Baseline EDSS	0.94 (0.75,1.18)	5.90e-01
log GFAP concentration	***2.3 (1.84,2.89)***	***4.17e-13***

CI, confidence interval; EDSS, Expanded Disability Status Scale; GFAP, serum glial fibrillary acidic protein; HR, hazard ratio; NfL, serum neurofilament light chain; Tau, serum tau; UCHL1, serum ubiquitin C-terminal hydrolase L1.

eTable 3 Multiple regression of Quanterix 4-plex versus change from baseline in EDSS score to attack.

Results from 4-way multiple regression on change in Quanterix measurements vs change in EDSS:

- Multiple R-squared: 0.38
- Adjusted R-squared: 0.27
- p-value: 0.02
- T-tests on individual coefficients (log_2 FC from baseline):

Analyte Name	Estimate (+/- 95% CI)*	p-value
sGFAP	0.32 (-0.097, 0.74)	0.41
sNFL	0.64 (0.013, 1.27)	0.046
sTAU	0.28 (-0.16,0.72)	0.20
sUCHL1	-0.49 (-1.07, 0.083)	0.09

CI, confidence interval; EDSS, Expanded Disability Status Scale; FC, fold change; sGFAP, serum glial fibrillary acidic protein; sNfL, serum neurofilament light chain; sTau, serum tau; sUCHL1, serum ubiquitin C-terminal hydrolase L1.

eTable 4 Mixed linear model of sNfL versus EDSS score at attack and at attack follow-up.

Parameter	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.35	0.61	36.30	3.86	4.53E-04
Age	0.05	0.01	34.00	3.35	2.01E-03
NFL	0.28	0.21	60.48	1.36	1.80E-01
Attack	1.23	0.19	70.00	6.47	1.14E-08
Follow-up	0.64	0.19	70.00	3.34	1.34E-03
NFL:Attack	0.34	0.19	70.00	1.77	8.11E-02
NFL:Follow-up	0.40	0.19	70.00	2.10	3.95E-02

EDSS, Expanded Disability Status Scale; NFL, neurofilament light chain.

eTable 5 Mixed-effects linear regression model for identifying attacks from samples drawn at scheduled visits.

```
lmy = lmer(NFL ~ `baseline NFL`+Timepoint + Treatment + Treatment:Timepoint + Age + EDSS +(1|subject))
```

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	**0.15**	**0.08**	**716.70**	**2.03**	**4.30E-02**
Baseline sNFL concentration	**0.73**	**0.03**	**206.33**	**26.62**	**1.19E-68**
Inebilizumab Treatment	0.04	0.09	714.61	0.50	6.19E-01
Age	**0.12**	**0.03**	**194.24**	**4.79**	**3.36E-06**
EDSS	**0.08**	**0.03**	**276.88**	**3.25**	**1.31E-03**
Week 4	-0.16	0.09	702.25	-1.70	8.96E-02
Week 12	-0.19	0.10	724.38	-1.88	6.09E-02
Week 16	-0.17	0.10	733.66	-1.63	1.04E-01
Week 28	-0.15	0.11	748.08	-1.37	1.71E-01
Attack	**0.78**	**0.14**	**817.58**	**5.72**	**1.53E-08**
Week4:Inebilizumab	0.07	0.11	702.83	0.64	5.25E-01
Week12:Inebilizumab	-0.18	0.11	723.76	-1.56	1.20E-01
Week16:trxInebilizumab	-0.21	0.12	731.69	-1.79	7.42E-02
Week28:trxInebilizumab	**-0.34**	**0.13**	**744.22**	**-2.69**	**7.21E-03**
Attack:trxInebilizumab	**-0.36**	**0.18**	**827.14**	**-2.00**	**4.55E-02**

Likelihood ratio test on nested models (reduced model excludes treatment)

	logLik	Chisq	Df	Pr(>Chisq)
Reduced Model	-687.52			
Full model	-678.21	18.63	6	0.0048

EDSS, Expanded Disability Status Scale; NFL, neurofilament light chain.