BRAIN
ORIGINAL ARTICLE

Neurofilament light chain and glial fibrillary acidic protein levels in metachromatic leukodystrophy

Shanice Beerepoot,[1,2,3] Hans Heijst,[4] Birthe Roos,[5] Mirjam M. C. Wamelink,[5] Jaap Jan Boelens,[2,6] Caroline A. Lindemans,[3,7] Peter M. van Hasselt,[8] Edwin H. Jacobs,[9] Marjo S. van der Knaap,[1,10] Charlotte E. Teunissen[4,†] and Nicole I. Wolf[1,†]

[†]These authors contributed equally to this work.

Metachromatic leukodystrophy is a lethal metabolic leukodystrophy, with emerging treatments for early disease stages. Biomarkers to measure disease activity are required for clinical assessment and treatment follow-up.

This retrospective study compared neurofilament light chain and glial fibrillary acidic protein (GFAP) levels in CSF ($n = 11$) and blood ($n = 92$) samples of 40 patients with metachromatic leukodystrophy (aged 0–42 years) with 38 neurologically healthy children (aged 0–17 years) and 38 healthy adults (aged 18–45 years), and analysed the associations between these levels with clinical phenotype and disease evolution in untreated and transplanted patients. Metachromatic leukodystrophy subtype was determined based on the (expected) age of symptom onset. Disease activity was assessed by measuring gross motor function deterioration and brain MRI. Longitudinal analyses with measurements up to 23 years after diagnosis were performed using linear mixed models.

CSF and blood neurofilament light chain and GFAP levels in paediatric controls were negatively associated with age (all $P < 0.001$). Blood neurofilament light chain level at diagnosis (median, interquartile range; picograms per millilitre) was significantly increased in both presymptomatic (14.7, 10.6–56.7) and symptomatic patients (136, 40.8–445) compared to controls (5.6, 4.5–7.1), and highest among patients with late-infantile (456, 201–854) or early-juvenile metachromatic leukodystrophy (291.0, 104–445) and those ineligible for treatment based on best practice (291, 57.4–472). GFAP level (median, interquartile range; picogram per millilitre) was only increased in symptomatic patients (591, 224–1150) compared to controls (119, 78.2–338) and not significantly associated with treatment eligibility ($P = 0.093$). Higher blood neurofilament light chain and GFAP levels at diagnosis were associated with rapid disease progression in late-infantile ($P = 0.006$ and $P = 0.051$, respectively) and early-juvenile patients ($P = 0.048$ and $P = 0.039$, respectively). Finally, blood neurofilament light chain and GFAP levels decreased during follow-up in untreated and transplanted patients but remained elevated compared with controls. Only neurofilament light chain levels were associated with MRI deterioration ($P < 0.001$).

This study indicates that both proteins may be considered as non-invasive biomarkers for clinical phenotype and disease stage at clinical assessment, and that neurofilament light chain might enable neurologists to make better informed treatment decisions. In addition, neurofilament light chain holds promise assessing treatment response. Importantly, both biomarkers require paediatric reference values, given that their levels first decrease before increasing with advancing age.

Received May 11, 2021. Revised June 22, 2021. Accepted July 14, 2021. Advance access publication August 16, 2021

1 Amsterdam Leukodystrophy Center, Department of Child Neurology, Emma Children's Hospital, Amsterdam University Medical Centers, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, 1081 HV Amsterdam, The Netherlands
2 Center for Translational Immunology, University Medical Center Utrecht, 3584 CX Utrecht, The Netherlands
3 Nierkens and Lindemans Group, Princess Máxima Center for Pediatric Oncology, 3584 CS Utrecht, The Netherlands
4 Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam University Medical Centers, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, 1081 HV Amsterdam, The Netherlands
5 Metabolic Unit, Department of Clinical Chemistry, Amsterdam University Medical Centers, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam Gastroenterology and Metabolism, 1081 HV Amsterdam, The Netherlands
6 Department of Pediatrics, Stem Cell Transplant and Cellular Therapies, Memorial Sloan Kettering Cancer Center, 10065 New York, USA
7 Department of Pediatrics, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht University, 3584 EA Utrecht, The Netherlands
8 Department of Pediatric Metabolic Diseases, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht University, 3584 EA Utrecht, The Netherlands
9 Department of Clinical Genetics, Erasmus University Medical Center, 3015 GD Rotterdam, The Netherlands
10 Department of Functional Genomics, Center for Neurogenomics and Cognitive Research, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands

Correspondence to: Nicole I. Wolf
Amsterdam UMC, location VUmc
De Boelelaan 1118
1081 HV Amsterdam, The Netherlands
E-mail: n.wolf@amsterdamumc.nl

Keywords: arylsulfatase A; metachromatic leukodystrophy; neurofilament light; glial fibrillary acidic protein; biomarker

Abbreviations: b/cGFAP = glial fibrillary acidic protein level in blood/CSF; b/cNfL = neurofilament light chain level in blood/CSF; FSIQ = full-scale intelligence quotient; HSC-GT = haematopoietic stem cell gene therapy; HSCT = haematopoietic stem cell transplantation; MLD = metachromatic leukodystrophy; NfL = neurofilament light chain

Introduction

Metachromatic leukodystrophy (MLD) (OMIM #250100) is an autosomal recessive lysosomal storage disease caused by deficient activity of arylsulfatase A. Accumulation of sulfatides, the macromolecular substrate of arylsulfatase A, in glia cells and neurons results in progressive central and peripheral demyelination. Consequently, patients show severe neurological deterioration characterized by loss of all motor and communication skills and eventually premature death.[1]

Based on the age of symptom onset, a late-infantile, early-juvenile, late-juvenile and adult clinical phenotype can be distinguished (onset at <2.5 years, 2.5–6 years, 6–16 years and >6 years, respectively).[1] The natural disease course can be roughly divided into four clinical stages. The first is a presymptomatic stage with normal motor and cognitive development, albeit evidence for demyelination may already be present both at brain imaging and nerve conduction studies.[2] The second is an early plateau stage, characterized by developmental slowing and stagnation, followed by the onset of first symptoms. The third is the stage of rapid disease progression when accelerated central demyelination results in rapid loss of gross motor function.[3–5] The final stage is marked by very slow deterioration or stabilization at a low functional level prior to death.[3] The duration of these stages can be shorter (late-infantile and early-juvenile MLD) or longer and more variable (late-juvenile and adult MLD).[3,6,7] Importantly, patients with late-juvenile and adult onset may present with mainly behavioural and cognitive symptoms, without the rapid motor deterioration seen in the patients with younger onset. Prediction of disease progression for these patients is particularly challenging.[7]

During the past three decades, allogeneic haematopoietic stem cell transplantation (HSCT) has proved to be a valuable therapeutic approach for pre- or early-symptomatic patients with a juvenile or adult MLD phenotype. It aims to correct arylsulfatase A deficiency in the brain, halt sulfatide accumulation and enable remyelination.[8–11] In addition, haematopoietic stem cell gene therapy (HSC-GT) has recently been approved as a treatment for presymptomatic patients with late-infantile MLD and pre- or early-symptomatic patients with early-juvenile MLD, while HSC-GT for later-onset MLD forms and intrathecal enzyme replacement therapy are still being evaluated in clinical trials (NCT04283227 and NCT01887938, respectively).[12] Evaluation of treatment eligibility and monitoring of therapeutic effects in both clinical and research setting require reliable biomarkers (with age-matched reference data) that are ideally cheap, reproducible and observer independent. In addition, such biomarkers in MLD would preferably be able to reflect different disease stages and mark rapid disease progression during which treatment with HSCT and HSC-GT is ineffective or even detrimental.[11,12] In this regard, neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) are considered potential candidates.[13,14]

NfL and GFAP are cytoskeletal proteins that are released into CSF and eventually blood on neuroaxonal and astroglial injury, respectively. Both biomarkers can be easily and reliably quantified in CSF, plasma and serum by a Single Molecule Array, and have been proven to serve as biomarkers of disease activity in several central and, in case of NfL, peripheral demyelinating diseases.[13–22] In addition, NfL and GFAP levels in CSF correlate strongly with levels in blood in most neurodegenerative diseases, highlighting the potential for NfL and GFAP in blood as minimally invasive measures for MLD disease activity.[13,14,16,21,22] Nevertheless, data on NfL and GFAP levels in paired CSF and blood samples of paediatric patients and controls are currently lacking.[23–25]

The main objective of this study was to determine the potential of NfL and GFAP levels in blood (bNfL and bGFAP) for MLD disease activity. We hypothesized that bNfL and bGFAP levels would be increased in presymptomatic and symptomatic patients with MLD compared to similar aged controls, and that higher levels would be associated with a more severe disease phenotype and clinical indicators of disease progression, reflecting higher disease activity. In addition, we presumed that bNfL and bGFAP levels would decline over time in transplanted patients, and that bNfL and bGFAP levels would strongly correlate with NfL and GFAP levels in CSF (cNfL and cGFAP). Confirmation of these hypotheses would establish bNfL and bGFAP levels as relatively non-invasive biomarkers to measure disease activity for clinical assessment and treatment follow-up for MLD.

Materials and methods

Subjects

In this retrospective study, we included 27 subjects from the MLD cohort (total $n = 90$) of the Amsterdam Leukodystrophy Center, a Dutch nationwide expertise centre, with a confirmed diagnosis of MLD[26]; available data on the age of symptom onset and loss of gross motor function over time and at least one available CSF, plasma or serum sample. In addition, we included 13 patients from the same cohort with an already performed NfL measurement during clinical assessment in our centre, resulting in a total inclusion of 40 study participants.

Two anonymous reference cohorts were included. The first cohort consisted of children with suspected neurological disorders for whom both results of additional investigations, including brain imaging and follow-up did not confirm neurological disease. Thirty-eight of these 'neurologically healthy' children between the age of 0.3 and 17.5 years with available (paired) blood and CSF samples were included. The second cohort yields from a previously reported in-house reference cohort, consisting of 38 healthy volunteers between the age of 18 and 45 years with available NfL and GFAP measurements in serum.[15] Paediatric controls older than 16 years were included in the adult group for the statistical comparison between patients and controls.

Standard protocol approvals and patient consents

The local Institutional Review Board approved the study, and appropriate written consent was obtained according to the Declaration of Helsinki.

Clinical assessment

Patient characteristics and follow-up data were collected from patient records. According to the age of symptom onset, patients were grouped into a 'late-infantile' (<2.5 years), 'early-juvenile' (2.5–6 years), 'late-juvenile' (6–16 years) and 'adult' (>16 years) clinical phenotype. Presymptomatic patients diagnosed by family screening were grouped into the same clinical phenotype as their MLD affected sibling. Presence and type of first symptoms were assessed at diagnosis and categorized as 'presymptomatic' (no symptoms), 'motor phenotype' (only motor symptoms), 'cognitive phenotype' (only cognitive symptoms) or 'mixed phenotype' (both motor and cognitive symptoms).[7] Cognitive performance was evaluated by full-scale intelligence quotient (FSIQ) at standardized neuropsychological testing at diagnosis. Peripheral neuropathy at diagnosis was dichotomized into 'no to mild peripheral neuropathy' and 'moderate to severe peripheral neuropathy' based on neurophysiologist's conclusions in nerve conduction study reports, since examinations were performed in multiple hospitals using different protocols and reference values. One late-juvenile patient was assigned 'no to mild peripheral neuropathy' based on the absence of clinical signs of peripheral neuropathy only, and one early-juvenile patient was assigned 'moderate to severe peripheral neuropathy' based on clinical signs and nerve pathology. Treatment characteristics included whether or not the patient proceeded to treatment with HSCT or HSC-GT based on clinical eligibility and, in the case of treatment with HSCT, the type of HSCT: 'bone marrow transplantation' or 'umbilical cord blood transplantation'.

Disease activity

Disease activity was determined in light of two distinct purposes: (i) to examine whether bNfL and bGFAP levels could be used to predict the stage of rapid disease progression at diagnosis; and (ii) to examine whether higher bNfL and bGFAP levels were associated with any disease progression over time. For the first purpose, the validated Gross Motor Function Classification in MLD (GMFC-MLD) was used to distinguish between patients with 'rapid disease progression' and 'slow disease progression' at the moment of diagnosis based on the time interval between symptom onset or, in the case of presymptomatic patients, diagnosis and entering GMFC-MLD level 2 (loss of independent walking).[27] Rapid disease progression and slow disease progression at diagnosis were defined as a time interval $\leqslant$27 months and >27 months, respectively, in line with previous publications.[2,5] For the second purpose, a validated measure of brain demyelination, the MRI severity score,[28] was used to quantify disease progression at any available time point during follow-up. We analysed the correlation with the crude MRI severity score at blood sampling and with 'MRI deterioration'. For the latter, we calculated the difference in MRI severity score compared to the closest previous MRI and divided this by the time interval in-between to account for variability in follow-up duration.

Sample collection and processing

Samples of paediatric controls and patients with MLD consisted of leftover samples from venous and lumbar puncture performed in the context of routine clinical care. In addition, NfL measurements were already performed in seven CSF samples and 15 blood samples of patients ($n = 13$) during previous clinical assessments. When paired, blood samples were collected directly before or after lumbar puncture. CSF was collected in polypropylene collection tubes and stored at –80°C or –30°C until analyses according to previously published consensus guidelines.[29] Blood was collected in EDTA, heparin or serum separator tubes and centrifuged for 10 min at room temperature at 1800*g*. Leftover plasma and serum were aliquoted in 0.5-ml volumes and stored at –80°C or –30°C until analyses according to previously published consensus guidelines.[29] Long-term storage at –30°C is not thought to have affected

NfL and GFAP concentrations.[30,31] Collection and processing methods of adult control samples have been described previously.[15] Regarding adult controls, only serum measurements were available.

Measurement of NfL and GFAP

All measurements were performed in duplicate using Single Molecule Array technology (Quanterix) with the commercial NF-Light advantage Kit and GFAP Discovery Kit on a HD-X platform according to the manufacturer's instructions (www.quanterix.com/products-technology/assays). The inter-assay variation for NfL was 8%, based on three levels of internal quality control pools, measured in 45 runs. The inter-assay variation for GFAP was 15%, based on three levels of internal quality control pools, measured in 40 runs. The mean intra-assay coefficients of variation (duplicate measurements) were <10% for both NfL and GFAP. Therefore, samples with too low volume for a duplicate measurement ($n = 4$) were also included in the analysis. Based on an in-house quality study, NfL values in heparin plasma, GFAP values in serum and GFAP values in heparin plasma were adjusted by –29, –13 and –18%, respectively, to allow comparison between measurements in heparin plasma, serum and EDTA plasma.[31,32] Finally, the GFAP values of the adult reference cohort were multiplied by a factor of 1.3, based on internal quality control values, to correct for inter-batch variability. All measurements were performed by certified technicians at the Neurochemistry laboratory of the Amsterdam UMC location VUmc blinded to clinical information.

Statistical analyses

Categorical parameters were described by counts and percentages, and continuous and ordinal parameters by median and interquartile range (IQR). Univariable analyses were performed using a chi-squared test, Fisher's exact test, Mann–Whitney–Wilcoxon Test or Kruskal–Wallis multiple comparison with Benjamini–Hochberg stepwise adjustment as appropriate. Crude correlations and differences between NfL and GFAP in paired CSF and blood samples were assessed by either the Spearman's rank correlation method, repeated measures correlation or Wilcoxon signed-rank test.

In all regression analyses, NfL and GFAP levels and age were log-transformed to meet the assumptions of normality. However, for clarity, regression coefficients were back-transformed to the original scales reflecting multiplicative effects (β_{mult}) for associations with non-transformed predictive parameters and describing percentages (β_{perc}) for associations with log-transformed predictive parameters.

Comparisons of NfL and GFAP levels between controls and patients were statistically adjusted for age at sampling and sex based on previous publications.[23] Because of the low number of patients, only parameters that were significantly associated in univariable analyses were included in the multivariable analyses among patients at diagnosis. As FSIQ at diagnosis included many missing values ($n = 10$, 25%), was measured with 11 different tests, and was often not fully reliable due to the discrepancy between verbal and performance intelligence quotients in individual patients, we decided to exclude this parameter from multivariable analyses beforehand. One missing value in MRI severity score was imputed by predictive mean matching.[33]

To analyse NfL and GFAP levels over time in patients linear mixed models were fitted including predictive patient, disease and transplantation parameters. Samples ($n = 2$) of two patients treated with HSC-GT were excluded from the analyses. Detailed model characteristics can be found in the Supplementary material.

All statistical tests were two-tailed and P-values <0.05 were considered statistically significant. The R project (RStudio: Integrated Development for R. RStudio, Inc., Boston, MA) for statistical computing v.4.0.3 with the packages 'lme4', 'rmcorr' and 'ggplot2' was used for all analyses and for the creation of the figures.

Data availability

Unpublished anonymized data within this article are available on reasonable request from a qualified investigator.

Results

Demographics of patients and controls

The patient dataset consisted of 40 patients with MLD between the age of 0 and 42 years, including four sibling pairs. Twenty patients (50%) were male. Seven patients (17.5%) had a late-infantile clinical phenotype, seven patients (17.5%) had an early-juvenile clinical phenotype, 18 patients (45%) had a late-juvenile clinical phenotype and eight patients (20%) had an adult clinical phenotype. Eleven patients were presymptomatic at diagnosis (28%) of whom 10 proceeded to treatment. In total, 19 patients (48%) proceeded to treatment with HSCT ($n = 17$) or HSC-GT ($n = 2$). Moderate to severe peripheral neuropathy at diagnosis was significantly more often observed in symptomatic patients (69%) than in presymptomatic patients (36%, $P < 0.001$), and in patients with a motor phenotype (100%) than in those with a mixed or cognitive phenotype (50 and 25%, respectively, $P < 0.001$). By contrast, median MRI severity score at diagnosis was significantly lower in patients with a motor phenotype [11 (IQR 5–18)] than in patients with a mixed [19 (IQR 17–21), $P = 0.020$] or cognitive phenotype [19 (IQR 18–20), $P = 0.020$]. In addition, patients with a motor phenotype were on average younger [median age 5.1 (IQR 2.3–6.7) years] than patients with a mixed [median age 14.8 (IQR 11.3–19.6) years, $P = 0.006$] or cognitive phenotype [median age 21.4 (IQR 13.8–26.7) years, $P = 0.002$]. Detailed patient demographics and disease characteristics, including differences between untreated and treated patients, are presented in Table 1 (irrespective of clinical phenotype) and Supplementary Table 1 (per clinical phenotype).

The control dataset included 38 unrelated neurologically healthy children (aged 0–17 years) and 38 unrelated adult healthy volunteers (aged 18–45 years). Twenty-four (63%) children were male [median age 2.8 (IQR 2.1–4.1) years] and 14 (37%) children were female [median age 3.1 (IQR 0.9–5.3) years]. Sixteen (42%) adults were male [median age 39.0 (IQR 28.5–43.3) years] and 22 (58%) adults were female [median age 37.0 (IQR 29.3–41.8) years]. One 17-year-old female paediatric control was included in the adult group (age category: >16 years) for statistical comparison of bNfL and bGFAP levels between patients and controls (Supplementary Table 1).

NfL and GFAP levels in paediatric controls

We present the results of all paediatric controls ($n = 38$). NfL and GFAP levels of adult controls were previously described.[15]

The paediatric control samples consisted of 33 paired (CSF and blood) and five unpaired CSF ($n = 4$) and blood ($n = 1$) samples. The blood samples consisted of heparin plasma ($n = 32$) and EDTA plasma ($n = 2$). Baseline cNfL and bNfL levels did not differ between males and females after correcting for age ($P = 0.240$ and $P = 0.144$, respectively), nor did cGFAP and bGFAP levels ($P = 0.768$ and $P = 0.087$, respectively). NfL and GFAP levels in CSF and blood were negatively associated with age (all $P < 0.001$). These associations

Table 1 Patient demographics and disease characteristics

Variable[a]	All (*n* = 40)	Treated (*n* = 19)	Untreated (*n* = 21)	*P* treated versus untreated[b]
Age at diagnosis, y	11.1 (4.6–19.2)	13.8 (3.2–17.4)	9.0 (5.6–20.2)	**< 0.001**
Male	20 (50.0)	8 (42.1)	12 (57.1)	0.527
Clinical phenotype, *n* (L-I/E-J/L-J/A)	7/7/18/8	3/3/8/5	4/4/10/3	0.860
Symptomatic at diagnosis	29 (72.5)	9 (47.4)	20 (95.2)	**0.001**
Age of onset, y	7.5 (3–14)	12 (3–12)	7 (4–14)	**< 0.001**
Type of first symptoms, *n* (mp/cp/mxp/p-s)	14/7/8/11	5/2/2/10	9/5/6/1	**< 0.001**
FSIQ at diagnosis[c]	88 (70–99)	93 (88–105)	65 (57–77)	**< 0.001**
GMFC-MLD score at diagnosis	1 (0–2)	0 (0–1)	1 (0–5)	**0.008**
MRI severity score at diagnosis	12 (2–19)	9 (3–12)	18 (17–20)	**< 0.001**
Moderate to severe peripheral neuropathy	24 (60.0)	10 (52.6)	14 (66.7)	0.561
Treatment, *n* (no, HSCT, HSC-GT)	21/17/2	0/17/2	21/0/0	—
Age at treatment, y	14.3 (4.7–17.8)	14.3 (4.7–17.8)	—	—
Rapid disease progression	15 (37.5)	4 (21.1)	11 (52.4)	0.055

A = adult; cp = cognitive phenotype; E-J = early-juvenile; GMFC-MLD = Gross Motor Function Classification in metachromatic leukodystrophy; L-J = late-juvenile; L-I = late-infantile; MLD = metachromatic leukodystrophy; mp = motor phenotype; mxp = mixed phenotype; p-s = presymptomatic; y = years.

[a]Values indicate median (IQR) or *n* (%) unless otherwise stated.

[b]P-values were obtained with a chi-squared test, Fisher's exact test, or Mann–Whitney–Wilcoxon Test as appropriate. Significant P-values are indicated in bold.

[c]Only 30 patients with available FSIQ measurements were included.

were not linear but showed a steeper decrease in NfL and GFAP levels during the first years of life. Thus, in healthy individuals NfL and GFAP levels first decrease in CSF and blood before increasing again with advancing age (Fig. 1A and B).

NfL levels in blood at metachromatic leukodystrophy diagnosis

NfL was quantified in 11 CSF and 92 blood samples (heparin plasma $n = 1$, EDTA plasma $n = 76$, serum $n = 15$), of which six and 27 were obtained at diagnosis, respectively. Median bNfL levels in patients at diagnosis were 136 (IQR 40.8–445) pg/ml for symptomatic patients and 14.7 (IQR 10.6–56.7) pg/ml for presymptomatic patients (Table 2). These levels were significantly increased in both patient groups compared to controls ($\beta_{mult} = 22.45$, $P < 0.001$ and $\beta_{mult} = 3.23$, $P < 0.001$, respectively), with a larger absolute and relative increase in patients with an earlier-onset clinical phenotype (Fig. 2A). Median bNfL levels were 456 (IQR 201–854) pg/ml, 291 (IQR 104–445) pg/ml, 38.6 (IQR 23.0–89.2) pg/ml and 33.8 (IQR 32.4–50.6) pg/ml for patients with late-infantile, early-juvenile, late-juvenile and adult MLD, respectively (Table 2).

Univariable analyses including all patients showed that bNfL levels at diagnosis were positively associated with a late-infantile or early-juvenile clinical phenotype (reference: adult phenotype, $P = 0.007$ and $P = 0.021$, respectively), presence of symptoms ($P = 0.003$), severe to moderate peripheral neuropathy at diagnosis ($P = 0.004$) and rapid disease progression ($P < 0.001$), while patients who were selected for treatment had lower bNfL levels ($P < 0.001$). A late-infantile or early-juvenile clinical phenotype and proceeding to treatment were confirmed to be independently associated with bNfL levels in a multivariable model (Table 2). Analyses including only symptomatic patients showed comparable results (Supplementary Table 2). Notably, MRI severity score at diagnosis was not associated with bNfL level in both analyses (Supplementary Fig. 1).

GFAP levels in blood at metachromatic leukodystrophy diagnosis

GFAP was quantified in four CSF and 88 blood samples (heparin plasma $n = 1$, EDTA plasma $n = 81$, serum $n = 6$) of which only blood samples ($n = 27$) were obtained at diagnosis. Median bGFAP levels in patients at diagnosis were 591 (IQR 224–1150) pg/ml for symptomatic patients and 403 (280–440) pg/ml for presymptomatic patients (Table 3). Levels of bGFAP were significantly increased in symptomatic patients compared to controls ($\beta_{mult} = 2.48$, $P < 0.001$), but not in presymptomatic patients ($\beta_{mult} = 1.44$, $P = 0.103$). The absolute GFAP increase was most pronounced in patients with an earlier-onset clinical phenotype, although its relative increase was comparable among the clinical phenotypes (Fig. 2B). Median bGFAP levels were 1069 (IQR 781–1414) pg/ml, 1150 (IQR 580–1348) pg/ml, 338 (IQR 227–458) pg/ml and 192 (IQR 179–221) pg/ml for patients with late-infantile, early-juvenile, late-juvenile and adult MLD, respectively (Table 3).

Levels of bGFAP at diagnosis were positively associated with age at sampling ($P < 0.001$), a late-infantile or early-juvenile clinical phenotype (reference: adult phenotype, $P < 0.001$ and $P < 0.001$, respectively), severe to moderate peripheral neuropathy at diagnosis ($P = 0.004$) and rapid disease progression ($P < 0.001$), and negatively associated with proceeding to treatment ($P = 0.023$) in univariable analyses. Only a late-infantile or early-juvenile clinical phenotype was confirmed to be independently associated with bGFAP levels in a multivariable model (Table 3). Analyses including only symptomatic patients showed similar results and an additional positive association between GFAP level and motor phenotype compared to cognitive phenotype ($P = 0.049$, Supplementary Table 3). MRI severity score at diagnosis was not associated with bGFAP level (Supplementary Fig. 1).

Association between NfL and GFAP levels and rapid disease progression at diagnosis

The associations between bNfL and bGFAP levels and rapid disease progression at diagnosis were analysed within all patients (statistically adjusted for clinical phenotype and presence of symptoms), and within each clinical phenotype (statistically adjusted for the presence of symptoms). Although median bNfL level was higher in patients with rapid disease progression than in patients with slow disease progression, this difference was only statistically significant in patients with late-infantile or early-juvenile MLD ($P = 0.006$ and $P = 0.048$, respectively; Fig. 2C and D). Interestingly, median bGFAP levels were also higher in patients with late-infantile and early-juvenile MLD with rapid disease progression than in those with slow disease progression ($P = 0.051$ and $P = 0.039$,

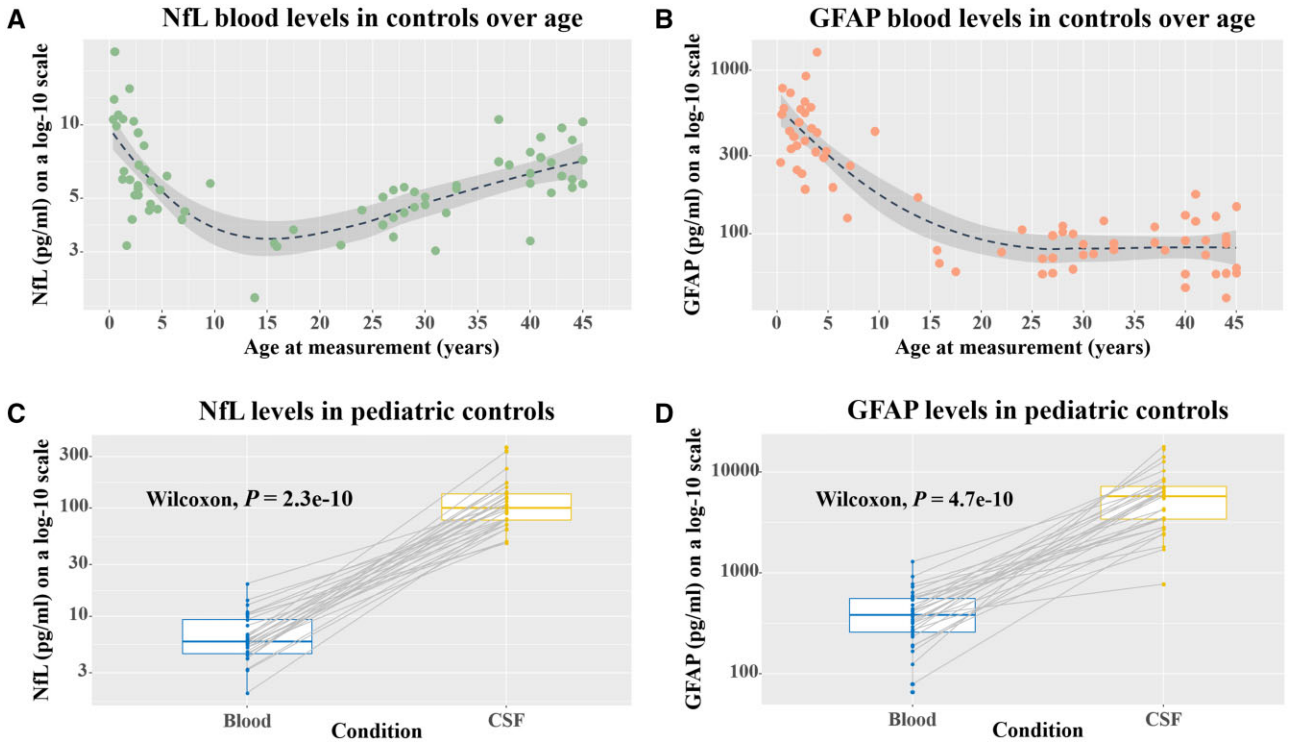

Figure 1 NfL and GFAP levels in controls expressed on a $\log_{10}$ scale. (**A**) NfL and (**B**) GFAP levels in blood are visualized over age with estimated LOESS regression curves (locally weighted scatter-plot smoother) and their 95% CIs (shadows). NfL and GFAP levels in blood show a non-linear decrease during childhood with a steeper slope in the first years of life, before increasing with advancing age in adulthood. NfL and GFAP levels in CSF show similar trends (data not shown because individual adult CSF control values were unavailable). (**C**) Median NfL level was significantly higher in CSF (100 pg/ml) compared to blood (5.8 pg/ml) and (**D**) median GFAP level was significantly higher in CSF (5564 pg/ml) compared to blood (370 pg/ml) in paired samples from paediatric controls (aged 0–17 years). Boxes depict median and IQR, grey lines indicate paired measurements, and upper/lower whiskers extend from the hinge towards the largest/smallest values but no farther than 1.5 times IQR from the hinge. The *P*-values were obtained with the non-parametric Wilcoxon signed-rank test.

respectively), while patients with late-juvenile and adult MLD with and without rapid disease progression had similar median bGFAP levels (Fig. 2E and F).

NfL levels in blood over time

NfL levels in blood in untreated and treated patients remained elevated over time compared to age-matched controls (Fig. 3A), even 20 years after diagnosis and treatment (Fig. 3B). Median bNfL level over time was lower in treated than in untreated patients ($P < 0.001$). In addition, bNfL decreased with increasing age ($P < 0.001$) and follow-up duration ($P < 0.001$; Table 4). However, in treated patients an increase (up to 6.4 times higher in the first 3 months) followed by a relatively steep decrease was observed within the first year after treatment with HSCT. Individual plots indicated that the increase and decrease were larger in younger than in older treated patients (Fig. 3A). Parameters that were associated with bNfL levels in treated patients separately are shown in Supplementary Table 4. Importantly, patients who were symptomatic at the moment of treatment had higher bNfL levels over time than patients who were treated presymptomatically (motor phenotype: $P = 0.019$, mixed phenotype: $P = 0.008$ and cognitive phenotype: $P = 0.015$, respectively).

GFAP levels in blood over time

Median GFAP levels in blood in untreated and treated patients remained elevated over time compared to age-matched controls, although in a few young individuals bGFAP levels decreased to normal values during the years following treatment with HSCT (Fig. 3C and D). Late-infantile, early-juvenile and late-juvenile clinical phenotypes were associated with lower bGFAP levels over time than an adult clinical phenotype ($P = 0.015$, $P = 0.025$ and $P = 0.045$, respectively). In addition, MRI severity score at diagnosis was positively associated with bGFAP levels over time ($P < 0.001$). Finally, bGFAP decreased with increasing age ($P < 0.001$), while follow-up duration was not a statically significant parameter (Table 5). Individual plots indicated also increase (up to 5.9 times higher in the first five months) and decrease in bGFAP level within the first year after treatment with HSCT, but this was comparable between younger and older treated patients (Fig. 3C). Parameters that were associated with bGFAP levels in treated patients separately are shown in Supplementary Table 5.

Blood NfL and blood GFAP levels and disease progression on brain MRI over time

Patients with a late-infantile clinical phenotype had overall lower crude MRI severity scores during follow-up than those with an early-juvenile, late-juvenile or adult clinical phenotype ($P = 0.076$, $P < 0.001$ and $P = 0.002$, respectively), but showed a higher increase in MRI severity score indicating faster MRI deterioration over time ($P = 0.028$, $P = 0.045$ and $P = 0.047$, respectively). Higher crude MRI severity scores during follow-up were correlated with both higher bNfL and bGFAP levels ($r = 0.31$; $P = 0.040$ and $r = 0.44$; $P = 0.003$, respectively), and these associations were retained after correcting for clinical phenotype ($\beta_{mult} = 1.096$; $P < 0.001$ and $\beta_{mult} = 1.062$; $P < 0.001$, respectively). MRI deterioration correlated, however, only with higher bNfL levels ($r = 0.45$; $P = 0.012$), and not with

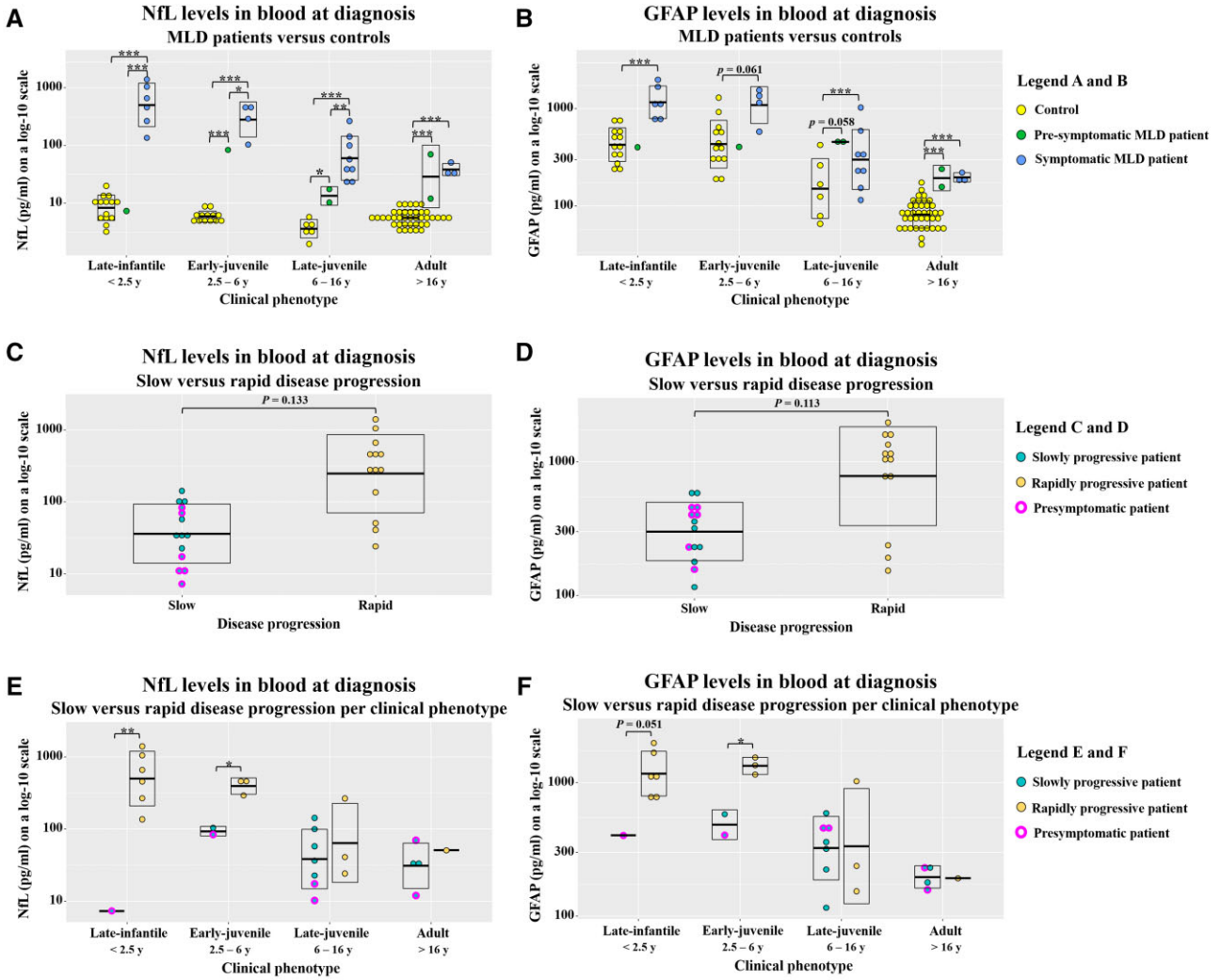

Figure 2 NfL and GFAP levels in blood in patients at MLD diagnosis. (**A**) Levels of bNfL and (**B**) bGFAP expressed on a $\log_{10}$ scale in patients with MLD compared to similar aged controls grouped by clinical phenotype and corresponding age. Boxes depict median and IQR within a group, and dots mark individual measurements. The *P*-values were obtained with a linear regression model adjusted for age at sampling and sex. (**C**) NfL levels and (**D**) GFAP levels in blood expressed on a $\log_{10}$ scale in patients with slow versus rapid disease progression, and separately for the four clinical phenotypes (**E** and **F**). Boxes depict median and IQR within a group, and dots mark individual measurements for patients. Magenta-contoured dots indicate presymptomatic patients. The *P*-values were obtained with a linear regression model adjusted for clinical phenotype and presence of symptoms. $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$. y = years.

bGFAP levels ($r = 0.250$; $P = 0.192$). This association retained after correcting for clinical phenotype ($\beta_{mult} = 1.171$; $P < 0.001$).

Correlations between NfL and GFAP levels in CSF and blood

Analysis of paired samples of paediatric controls revealed that median NfL level in CSF [101 (IQR 77.6–135) pg/ml] was 17.3-fold higher than in blood [5.8 (IQR 4.5–9.0) pg/ml; Fig. 1C]. Median GFAP level in CSF [5564 (IQR 3401–7169) pg/ml] was 15-fold higher than in blood [370 (IQR 246–550) pg/ml, Fig. 1D]. NfL and GFAP levels in CSF showed moderate correlations with NfL and GFAP levels in blood ($r = 0.52$, $P = 0.002$ and $r = 0.50$, $P = 0.004$, respectively), but these associations were not retained after correcting for age and sex ($P = 0.142$ and $P = 0.466$, respectively).

Analysis of paired samples of patients revealed that median NfL level in CSF [7955 (IQR 2544–10 291) pg/ml] was 56-fold higher than in blood [142.0 (IQR 53.9–280) pg/ml], far beyond the multiplication factor in controls. In addition, cNfL levels showed a strong correlation with bNfL levels ($r = 0.82$, $P = 0.034$), and this association retained after correcting for age and sex, with a 10% increase in cNfL leading to a 10.6% increase in bNfL ($\beta_{perc} = 1.06$, $P < 0.001$). Only one patient had paired CSF and blood results for GFAP. The cGFAP level (44 005 pg/ml) was 26.4-fold higher than bGFAP level (1664 pg/ml) in this patient.

Discussion

There is a need for objective and easy to measure biomarkers to predict MLD disease activity in patients in research and clinical settings. To determine the value of NfL and GFAP levels in CSF and blood as biomarkers for MLD activity, we measured the levels in untreated and transplanted patients over time and compared them to two combined reference cohorts of paediatric and adult controls. We found that NfL and GFAP levels in blood were significantly increased in patients with MLD compared to controls, and that higher levels in patients at diagnosis were associated with a more severe clinical phenotype and higher disease activity,

Table 2 Univariable and multivariable associations between NfL levels in blood at diagnosis and clinical parameters in patients with MLD

Variable (*n* patients = 27)		Univariable model			Multivariable model		
	bNfL, pg/ml[a]	β[b]	95% CI	P[c]	β[b]	95% CI	P[c]
Age at sampling, y	—	–0.368	–0.845–0.110	0.126	NA	NA	NA
Sex							
Female (ref)	29.0 (18.7–116)	—	—	—	—	—	—
Male	104 (50.6–291)	2.447	0.766–7.819	0.125	NA	NA	NA
Clinical phenotype							
Adult (ref)	33.8 (32.4–50.6)	—	—	—	—	—	—
Late-juvenile	38.6 (23.0–89.2)	1.306	0.338–5.052	0.687	1.140	0.403–3.231	0.795
Early-juvenile	291 (104–445)	6.480	1.359–30.890	**0.021**	3.471	1.046–11.519	**0.043**
Late-infantile	456 (201–854)	8.030	1.891–34.093	**0.007**	3.779	1.153–12.386	**0.030**
Symptomatic at diagnosis							
No (ref)	14.7 (10.6–56.7)	—	—	—	—	—	—
Yes	136 (40.8–445)	6.558	2.005–21.448	**0.003**	1.881	0.593–5.61	0.266
FSIQ at diagnosis[d]	—	1.000	0.970–1.032	0.981	NA	NA	NA
MRI severity score at diagnosis	—	1.021	0.936–1.114	0.625	NA	NA	NA
Peripheral neuropathy							
No to mild (ref)	35.1 (16.0–87.7)	—	—	—	—	—	—
Moderate to severe	266 (60.2–464)	4.710	1.736–12.773	**0.004**	1.370	0.429–4.380	0.577
Rapid disease progression							
No (ref)	35.1 (18.7–80.2)	—	—	—	—	—	—
Yes	291 (136–472)	6.814	2.845–16.319	**<0.001**	1.506	0.453–5.002	0.484
Proceeding to treatment							
No (ref)	291 (57.4–472)	—	—	—	—	—	—
Yes	36.6 (18.7–95.8)	0.177	0.069–0.452	**0.023**	0.353	0.142–0.879	**0.027**

NA = not assessed; ref = reference group; y = years.
[a]Values indicate median (IQR).
[b]The β estimates represent multiplicative effects of parameter, holding everything else constant. However, since age at sampling was also log-transformed, this β estimate indicates the average percentage of decrease in NfL level for every percent increase in age, holding everything else constant.
[c]P-values were obtained with univariable and multivariable linear regression models. Because of the low number of patients, only parameters that were significantly associated with bNfL level in univariable analysis were included in the multivariable analysis. These were 'clinical phenotype', 'symptomatic at diagnosis', 'peripheral neuropathy', 'rapid disease progression' and 'proceeding to treatment'. Significant *P*-values are indicated in bold.
[d]Only 22 patients with available FSIQ measurements at diagnosis were included.

thereby predicting (imminent) clinical progression. In addition, we observed that bNfL level declined over time since diagnosis but remained elevated in all transplanted patients, while bGFAP level declined to normal level in some transplanted individuals with an early- or late-juvenile clinical phenotype. In the current study, bNfL outperformed bGFAP as a biomarker for disease activity in MLD, because only bNfL level was associated with clinical eligibility for treatment at diagnosis and active demyelination on brain MRI during follow-up. Therefore, the question rises whether bGFAP level provides value as a biomarker of MLD disease activity in addition to bNfL level. Considering the exploring goal of the current study, this question still needs to be answered. It will be interesting to see whether, e.g. in clinical trial evaluation of new treatment modalities, a pattern of putative therapeutic effects across the two individual biomarkers might emerge, underlining the benefit of measuring both. The use of both biomarkers requires age-specific, paediatric reference values, since, in contrast to findings in adult controls, NfL and GFAP levels in CSF and blood in paediatric controls were negatively associated with age, showing a steep decline during the first years of life.

Value of blood NfL as a biomarker in metachromatic leukodystrophy

The data in this study indicate that especially bNfL might be useful as an enrichment and prognostic biomarker to support treatment eligibility decisions, indicate rapid disease progression at diagnosis and to monitor disease progression during long-term follow-up. The NfL level in blood was strongly correlated with NfL level in CSF in patients and has therefore the potential of reflecting disease activity in the CNS and PNS without the need for invasive testing. Although validation of this finding is needed in another cohort of patients with MLD, our results are in accordance with previous studies on cNfL and bNfL levels in paediatric neurological diseases, including juvenile neuronal ceroid lipofuscinosis, multiple sclerosis and acute disseminated encephalomyelitis.[24,25] In addition, NfL levels were associated with clinical phenotype and biomarkers for disease activity and treatment response in such diseases, e.g. disease phenotype in patients with ataxia-telangiectasia; magnetic resonance spectroscopy markers and the Unified Batten Disease Rating Scale score in juvenile neuronal ceroid lipofuscinosis; lesion load on brain MRI, disease progression and time to relapse in paediatric multiple sclerosis; improvement of motor function after treatment in spinal muscular atrophy type 1; and treatment in paediatric opsoclonus–myoclonus syndrome.[24,25,34–37]

Another important advantage of bNfL level compared to other currently used biomarkers in MLD care and research, including MRI severity scores, cognitive functioning and the GMFC-MLD score, is that bNfL, due to its rapid increase on neuroaxonal injury and its half-life of only a few weeks to 2 months, reflects recent or ongoing disease activity instead of earlier accumulated disease damage.[38] This could explain why we could not establish a close correlation between bNfL and those biomarkers at diagnosis, and that patients with lower MRI severity scores already had relatively

Table 3 Univariable and multivariable associations between GFAP levels in blood at diagnosis and clinical parameters in patients with MLD

Variable (*n* patients = 27)		Univariable model			Multivariable model		
	bGFAP, pg/ml[a]	β[b]	95% CI	*P*[c]	β[b]	95% CI	*P*[c]
Age at sampling, y	—	–0.433	–0.659–0.207	**<0.001**	NA	NA	NA
Sex							
Female (ref)	405 (190–680)	—	—	—	—	—	—
Male	580 (240–1150)	1.342	0.674–2.674	0.388	NA	NA	NA
Clinical phenotype							
Adult (ref)	192 (179–221)	—	—	—	—	—	—
Late-juvenile	338 (227–458)	1.668	0.894–3.110	0.103	1.581	0.885–2.822	0.115
Early-juvenile	1150 (580–1348)	4.558	2.219–9.361	**<0.001**	3.875	1.949–7.704	**<0.001**
Late-infantile	1069 (781–1414)	5.120	2.630–9.968	**<0.001**	4.034	2.042–7.968	**<0.001**
Symptomatic at diagnosis							
No (ref)	403 (280–440)	—	—	—	—	—	—
Yes	591 (224–1150)	1.614	0.733–3.551	0.223	NA	NA	NA
FSIQ at diagnosis[d]	—	1.011	0.993–1.029	0.212	NA	NA	NA
MRI severity score at diagnosis	—	0.976	0.929–1.025	0.318	NA	NA	NA
Peripheral neuropathy							
No to mild (ref)	380 (210–454)	—	—	—	—	—	—
Moderate to severe	809 (238–1251)	1.945	1.044–3.625	**0.037**	0.850	0.452–1.598	0.596
Rapid disease progression							
No (ref)	338 (222–440)	—	—	—	—	—	—
Yes	1069 (753–1348)	2.604	1.502–4.512	**<0.001**	1.455	0.734–2.881	0.267
Proceeding to treatment							
No (ref)	1069 (358–1348)	—	—	—	—	—	—
Yes	360 (227–458)	0.489	0.266–0.898	**0.023**	0.678	0.428–1.074	0.093

NA = not assessed; ref = reference group; y = years.

[a]Values indicate median (IQR).

[b]The β estimates represent multiplicative effects of parameter, holding everything else constant. However, since age at sampling was also log-transformed, this β estimate indicates the average percentage of decrease in GFAP level for every percent increase in age, holding everything else constant.

[c]P-values were obtained with univariable and multivariable linear regression models. Because of the low number of patients, only parameters that were significantly associated with bNfL level in univariable analysis were included in the multivariable analysis. These were 'clinical phenotype', 'peripheral neuropathy', 'rapid disease progression' and 'proceeding to treatment'. 'Age at sampling' was removed from the model due to strong multicollinearity. Significant P-values are indicated in bold.

[d]Only 22 patients with available FSIQ measurements at diagnosis were included.

high bNfL levels and vice versa, depending on the combination of clinical phenotype and stage of the disease. In more detail, NfL blood level was increased in presymptomatic patients compared to similar aged controls, suggesting that the disease was already causing neuroaxonal injury but not clinical symptoms yet. Preclinical neuroaxonal injury is also observed in other neurodegenerative diseases, including familial amyotrophic lateral sclerosis, familial Alzheimer's disease and familial frontotemporal dementia.[39–41] Interestingly, of the two presymptomatic patients with normal brain MRI, cognitive functioning and nerve conduction study results at moment of sampling in our cohort, one, a late-juvenile patient, had increased bNfL level at age 6 years (expected age at symptom onset: 10 years), while the other, a late-infantile patient, had still normal bNfL level at age 2 months (expected age at symptom onset: 1.5–2 years). If a certain threshold of neuroaxonal and white matter damage has to be reached before symptoms and MRI abnormalities appear, NfL might be used to monitor disease activity in these patients in case immediate HSCT or HSC-GT is undesired or not possible. More research is, however, needed to examine whether this hypothesis is true, and to establish the added value of bNfL besides the currently used biomarkers.

Potential implications of blood NfL and blood GFAP levels in metachromatic leukodystrophy pathophysiology

An important finding is that bNfL level declined over time from diagnosis in all patients, including untreated patients. In untreated patients, this decline might reflect transition from the stage of rapid disease progression to the stage of slow deterioration or clinical stabilization, parallel to neuroaxonal loss, which is substantial at the beginning, but burnt out in later stages, resulting in less residual tissue to release NfL into CSF and blood. Consequently, NfL levels should always be interpreted in light of disease stage in addition to age and clinical phenotype, and comparison between patients in different disease stages should be avoided. In addition, a decline in NfL level in treated patients should be inferred as a treatment response only when it is steeper than the decline observed in untreated patients.

Contrary to bNfL, overall bGFAP levels were not increased in presymptomatic patients, and bGFAP levels continued to increase after diagnosis in most untreated patients. In addition, higher bGFAP levels over time were associated with higher MRI severity scores and an adult clinical phenotype. The mechanisms of bGFAP elevation are not yet understood. One possible explanation is that significant astrogliosis may occur only in a later symptomatic stage, especially in patients with an early-onset clinical phenotype. Pathological or imaging evidence for differences in severity of the astrogliosis between MLD clinical phenotypes and during disease evolution has however not been reported yet. A similar relation has been observed in multiple sclerosis, where bGFAP correlated with disease severity scores particularly in progressive multiple sclerosis, reflecting increasing astrogliosis in later disease stages.[14] Alternative explanations, including differences in the mechanisms of release, turnover and kinetics of GFAP and NfL in the CNS and blood compartment, should also be considered.

Table 4 Multivariable associations between NfL levels in blood over time and clinical parameters in untreated and treated patients with MLD

Variable (*n* patients = 35, *n* measurements = 90)	Full model			Final model		
	β[a]	95% CI	P[b]	β[a]	95% CI	P[b]
Follow-up since diagnosis, y	See Fig. 3A (rcs)		**<0.01**	See Fig. 3A (rcs)		**<0.001**
Age at sampling, y	–0.663	–1.519–0.194	0.127	–0.854	–1.097 to –0.612	**<0.001**
Sex						
Female (ref)	—	—	—	—	—	—
Male	1.248	0.713–2.184	0.421	1.459	0.979–2.176	0.063
Clinical phenotype						
Adult (ref)	—	—	—	—	—	—
Late-juvenile	1.003	0.366–2.753	0.995	NA	NA	NA
Early-juvenile	0.936	0.175–5.005	0.936	NA	NA	NA
Late-infantile	1.235	0.098–15.525	0.865	NA	NA	NA
Type of first symptoms						
Presymptomatic (ref)	—	—	—	—	—	—
Motor phenotype	1.641	0.535–5.030	0.370	NA	NA	NA
Mixed phenotype	1.759	0.549–5.563	0.898	NA	NA	NA
Cognitive phenotype	1.062	0.409–2.753	0.898	NA	NA	NA
MRI severity score at diagnosis	0.989	0.902–1.083	0.801	NA	NA	NA
Peripheral neuropathy						
No to mild (ref)	—	—	—	—	—	—
Moderate to severe	1.047	0.581–1.886	0.875	NA	NA	NA
Proceeding to treatment						
No (ref)	—	—	—	—	—	—
Yes	0.409	0.179–0.933	**0.035**	0.388	0.256–0.588	**<0.001**

NA = not applicable; rcs = restricted cubic splines; ref = reference group; y = years.

[a]The β estimates represent multiplicative effects of parameter, holding everything else constant. However, since age at sampling was also log-transformed, this β estimate indicates the average percentage of decrease in NfL level for every percent increase in age, holding everything else constant.

[b]P-values were obtained with a multivariable mixed-effect model including all parameters (full model) and only those selected by maximum likelihood estimation (final model). Significant P-values are indicated in bold.

NfL and GFAP blood levels in relation to transplantation

The increase in bNfL and bGFAP levels within the first months after transplantation suggests that neuroaxonal damage and astrogliosis are more prominent after HSCT, in line with the previously described progression of white matter abnormalities and atrophy on brain MRI during this period.[10,11] Individual plots indicated that the increase in bNfL and bGFAP level was larger in younger than in older patients, although the longer time intervals between HSCT and blood sampling in the older patients might (partially) have caused this difference. An increase in bNfL and bGFAP levels is likewise reported in rats and humans treated with myeloablative HSCT, suggesting that these increases are, at least to some degree, the result of (transient) neurotoxicity caused by chemotherapy.[42,43] Accordingly, Thebault *et al.*[43] found that a higher total busulfan dose correlated with a larger increase in bNfL and bGFAP level at 3 months post-HSCT and with a greater degree of grey and white matter volume loss in patients with multiple sclerosis. However, the average increase in bNfL and bGFAP levels in that study was much lower (bNfL: 7.0 pm/ml, representing a 32.1% increase, and bGFAP: 80.3 pm/ml, representing a 74.8% increase) than the average increase observed at 3 months post-HSCT in our study (bNfL: 182 pm/ml, representing a 609% increase, and bGFAP: 245 pm/ml, representing a 187% increase).[43] Nevertheless, since busulfan is often used as a chemotherapeutic agent in HSCT and HSC-GT for MLD, more knowledge on the increase in bNfL and bGFAP level due to busulfan neurotoxicity is needed before these levels can serve as biomarkers to monitor short-term treatment response.[10,12]

In addition, we found evidence that neuroaxonal damage and astrogliosis continue in patients treated with HSCT, despite stabilization or even improvement of white matter abnormalities on brain MRI. This might reflect progressive peripheral neuropathy after HSCT,[44] or, more likely, the fact that enzymatic cross-correction of neurons and neuroglia after HSCT is limited if present at all, leading to a suboptimal treatment effect.[8,45,46] Nevertheless, in accordance with previous findings of better treatment outcomes,[9,11,47–49] patients who underwent transplantation at a presymptomatic stage exhibited lower bNfL levels over time than patients at a symptomatic stage. This finding again underlines the importance of early treatment in MLD.

NfL and GFAP levels in CSF and blood require age-specific, paediatric reference values

We found that NfL and GFAP levels in both CSF and blood in paediatric controls were negatively associated with age, with a steep decline during the first years of life. This is in accordance with another large study on NfL and GFAP levels in CSF of paediatric controls, although the decline during the first years of life was not explicitly mentioned there.[50] There is no good hypothesis yet to explain this decline. NfL and GFAP turnover in infants and young children might be slower than in older children.[50] Myelination, mainly occurring in the first 2 years of life, might also contribute to decreasing axonal neurofilament loss with older age.[52,53] Other previous studies on NfL and GFAP levels in CSF or blood of paediatric controls did not find an association with age.[24,25,36,37] However, these studies were conducted in a small number of children,[24,25,37] did not include controls younger than 3 years,[24,25] or used patients with other neurological diseases as a control group.[36] Importantly, these physiological changes may complicate the implementation

Table 5 Multivariable associations between GFAP levels in blood over time and clinical parameters in untreated and treated patients with MLD

Variable (*n* patients = 33, *n* measurements = 86)	Full model			Final model		
	β[a]	95% CI	P[b]	β[a]	95% CI	P[b]
Follow-up since diagnosis, y	See Fig. 3B (rcs)		NS	See Fig. 3B (rcs)		NS
Age at sampling, y	−1.583	−2.206 to −0.956	**<0.001**	−1.469	−1.963 to −1.000	**<0.001**
Sex						
Female (ref)	—	—	—	—	—	—
Male	0.854	0.604–1.207	0.354	NA	NA	NA
Clinical phenotype						
Adult (ref)	—	—	—	—	—	—
Late-juvenile	0.481	0.226–1.027	0.058	0.581	0.342–0.987	**0.045**
Early-juvenile	0.290	0.084–1.000	0.050	0.358	0.147–0.870	**0.025**
Late-infantile	0.147	0.023–0.957	**0.045**	0.208	0.060–0.719	**0.015**
Type of first symptoms						
Presymptomatic (ref)	—	—	—	—	—	—
Motor phenotype	1.369	0.651–2.878	0.390	NA	NA	NA
Mixed phenotype	1.162	0.562–1.429	0.673	NA	NA	NA
Cognitive phenotype	0.778	0.424–1.429	0.401	NA	NA	NA
MRI severity score at diagnosis	1.076	1.005–1.151	**0.036**	1.060	1.030–1.091	**<0.001**
Peripheral neuropathy						
No to mild (ref)	—	—	—	—	—	—
Moderate to severe	0.758	0.521–1.102	0.139	NA	NA	NA
Proceeding to treatment						
No (ref)	—	—	—	—	—	—
Yes	1.113	0.588–2.108	0.731	NA	NA	NA

NA = not applicable; rcs = restricted cubic splines; ref = reference group; y = years.

[a]The β estimates represent multiplicative effects of parameter, holding everything else constant. However, since age at sampling was also log-transformed, this β estimate indicates the average percentage of decrease in GFAP level for every percent increase in age, holding everything else constant.

[b]P-values were obtained with a multivariable mixed-effect model including all parameters (full model) and only those selected by maximum likelihood estimation (final model). Significant *P*-values are indicated in bold.

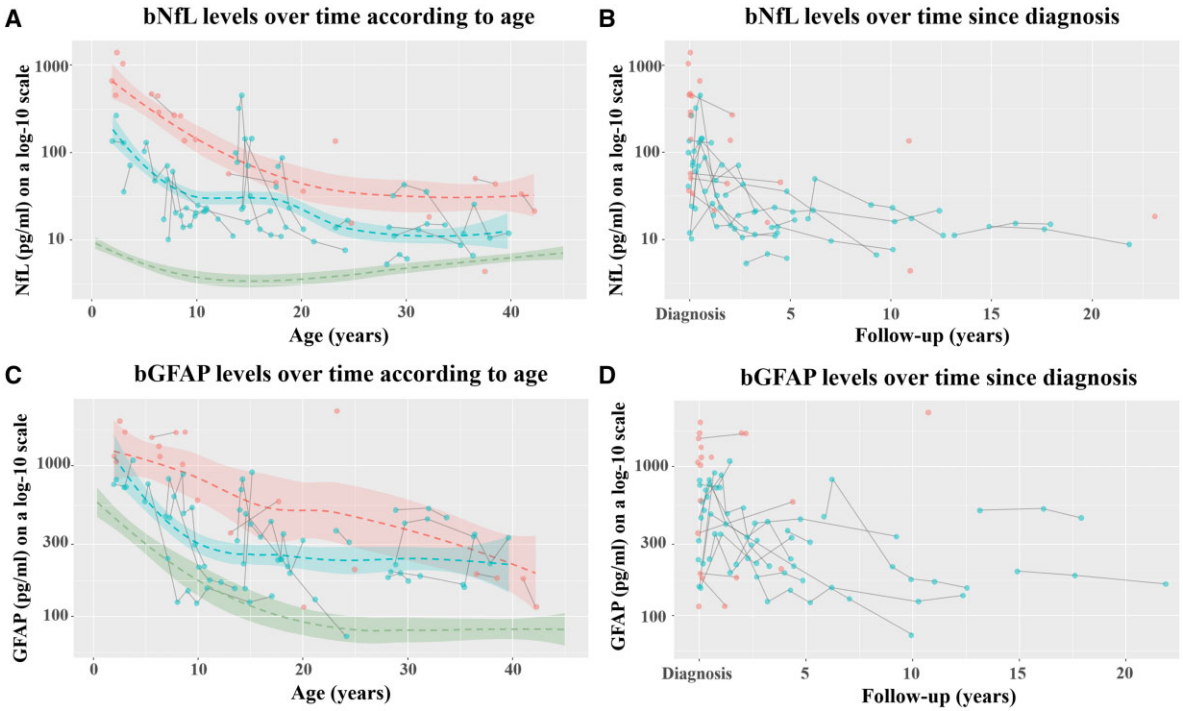

Figure 3 Longitudinal comparison of untreated and treated patients with MLD according to age (A and C) or since diagnosis (B and D). The *left* panels show the course of bNfL level (**A**) and bGFAP level (**C**) over time according to age as estimated by a linear mixed model for untreated (peach line) and treated patients (blue line). The 95% CIs are shown as shadows in a corresponding colour. Reference values are visualized in green with an estimated LOESS regression curve and 95% CI (shadow). Coloured dots (measurements) and shaded lines (course over time) reflect the individual patient data. The *right* panels show the same individual patient data at bNfL level (**B**) and bGFAP level (**D**), but over time since diagnosis.

of NfL or GFAP as biomarker in childhood leukodystrophies such as MLD and require age-specific reference values.

Strengths and limitations of the study

Among the strengths of this study are the combination of extensive cross-sectional and longitudinal phenotype data in a relatively large MLD cohort, and the inclusion of a paediatric reference cohort with paired CSF and blood samples. However, there are also some limitations to be addressed. First, our results have yet to be validated in an external cohort of patients. Second, the number of paediatric controls aged 6–16 years was relatively low. Increasing the number of control samples from this age group would be of special importance since the current data indicate that NfL and GFAP levels are lowest within this age range. In addition, the number of patients limited the number of clinical parameters that could be included in our analyses regarding bNfL and bGFAP levels at diagnosis, and our study might not have had sufficient power to detect all significant associations. The groups in the analyses regarding rapid disease progression were particularly small, and these results should therefore be interpreted with caution. Finally, only a few bNfL and bGFAP measurements of untreated patients were available of during follow-up. These patients were overall older and had a slower disease course than most untreated patients.

Conclusion

To conclude, this study indicates that both NfL and GFAP levels in blood hold promise as relatively non-invasive biomarkers for MLD disease stage in both clinical and research setting, and that especially NfL could support neurologists to make informed treatment decisions and to monitor residual disease activity during clinical follow-up. However, bNfL and bGFAP thresholds and their added value with respect to current biomarkers in MLD remain to be validated in a future prospective study. Currently, data are not sufficient to consider NfL and GFAP treatment response biomarkers. Finally, age-specific reference values in children are required.

Acknowledgements

We thank Lynn Boonkamp and Patty Hoede for their help with the GFAP analyses. We are also grateful for the cooperation of all referring hospitals and to Dr Marc Engelen for sharing his NfL and GFAP data of adult controls.

Funding

This study was partially funded by the Dutch charity organization Metakids (Metakids: 2017–073 to N.I.W. and J.J.B.). In addition, research by C.E.T. was supported by the European Commission [Marie Curie International Training Network, grant agreement no, 860197 (MIRIADE) and JPND], Health Holland, the Dutch Research Council (ZonMW), Alzheimer Drug Discovery Foundation, The Selfridges Group Foundation, Alzheimer Netherlands and the Alzheimer Association. The funding sources had no role in the design, analyses, reporting of the study or in the decision to submit the manuscript for publication. The authors of this publication affiliated with the Amsterdam Leukodystrophy Center are members of the European Reference Network for Rare Neurological Diseases—Project ID no. 739510.

Competing interests

C.A.L. is adviser for trials in Metachromatic Leukodystrophy (Orchard). M.S.v.d.K. is consultant and co-investigator to Ionis (trial in Alexander disease); and consultant to Calico, Denali and Evotec, all without personal payment. She has a patent P112686CA00, therapeutic effects of Guanabenz treatment in vanishing white matter, pending to VU University Medical Center. C.T. has a collaboration contract with ADx Neurosciences and Quanterix, and has performed contract research or received grants from AC-Immune, Axon Neurosciences, Biogen, Brainstorm Therapeutics, Celgene, EIP Pharma, Eisai, PeopleBio, Roche, Toyama and Vivoryon. N.I.W. is adviser and/or co-investigator for trials in Metachromatic Leukodystrophy (Shire/Takeda, Orchard, Evidera) and other leukodystrophies (Ionis, PassageBio, Vigil Neuro), without personal payment. All other authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Von Figura K, Gieselmann V, Jaeken J. Metachromatic leukodystrophy. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The metabolic and molecular bases of inherited disease*. McGraw-Hill;2001:3695–3724.
2. Groeschel S, Kehrer C, Engel C, et al. Metachromatic leukodystrophy: Natural course of cerebral MRI changes in relation to clinical course. *J Inherit Metab Dis*. 2011;34(5):1095–1102.
3. Elgün S, Waibel J, Kehrer C, et al. Phenotypic variation between siblings with Metachromatic Leukodystrophy. *Orphanet J Rare Dis*. 2019;14(1):136.
4. Dali C, Barton NW, Farah MH, et al. Sulfatide levels correlate with severity of neuropathy in metachromatic leukodystrophy. *Ann Clin Transl Neurol*. 2015;2(5):518–533.
5. Strölin M, Krägeloh-Mann I, Kehrer C, Wilke M, Groeschel S. Demyelination load as predictor for disease progression in juvenile metachromatic leukodystrophy. *Ann Clin Transl Neurol*. 2017;4(6):403–410.
6. Kehrer C, Blumenstock G, Gieselmann V, Krageloh-Mann I, German L; on behalf of the German Leukonet. The natural course of gross motor deterioration in metachromatic leukodystrophy. *Dev Med Child Neurol*. 2011;53(9):850–855.
7. Kehrer C, Elgun S, Raabe C, et al. Association of age at onset and first symptoms with disease progression in patients with metachromatic leukodystrophy. *Neurology*. 2020;97(2):e255–e266.
8. Wolf NI, Breur M, Plug B, et al. Metachromatic leukodystrophy and transplantation: Remyelination, no cross-correction. *Ann Clin Transl Neurol*. 2020;7(2):169–180.
9. Martin HR, Poe MD, Provenzale JM, Kurtzberg J, Mendizabal A, Escolar ML. Neurodevelopmental outcomes of umbilical cord blood transplantation in metachromatic leukodystrophy. *Biol Blood Marrow Transplant*. 2013;19(4):616–624.
10. van Rappard DF, Boelens JJ, van Egmond ME, et al. Efficacy of hematopoietic cell transplantation in metachromatic leukodystrophy: The Dutch experience. *Blood*. 2016;127(24):3098–3101.
11. Beschle J, Doring M, Kehrer C, et al. Early clinical course after hematopoietic stem cell transplantation in children with juvenile metachromatic leukodystrophy. *Mol Cell Pediatr*. 2020;7(1):12.
12. Sessa M, Lorioli L, Fumagalli F, et al. Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic

leukodystrophy: An ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet*. 2016;388(10043):476–487.
13. Khalil M, Teunissen CE, Otto M, et al. Neurofilaments as biomarkers in neurological disorders. *Nat Rev Neurol*. 2018;14(10):577–589.
14. 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.
15. Ballegoij WJC, Stadt SIW, Huffnagel IC, et al. Plasma NfL and GFAP as biomarkers of spinal cord degeneration in adrenoleukodystrophy. *Ann Clin Transl Neurol*. 2020;7(11):2127–2136.
16. Disanto G, Barro C, Benkert P, et al.; the Swiss Multiple Sclerosis Cohort Study Group. Serum neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann Neurol*. 2017;81(6):857–870.
17. Thebault S, R. Tessier D, Lee H, et al. High serum neurofilament light chain normalizes after hematopoietic stem cell transplantation for MS. *Neurol Neuroimmunol Neuroinflamm*. 2019;6(5):e598.
18. van Lieverloo GGA, Wieske L, Verhamme C, et al. Serum neurofilament light chain in chronic inflammatory demyelinating polyneuropathy. *J Peripher Nerv Syst*. 2019;24(2):187–194.
19. Sandelius A, Zetterberg H, Blennow K, et al. Plasma neurofilament light chain concentration in the inherited peripheral neuropathies. *Neurology*. 2018;90(6):e518–e524.
20. Petzold A, Hinds N, Murray NMF, et al. CSF neurofilament levels: A potential prognostic marker in Guillain-Barré syndrome. *Neurology*. 2006;67(6):1071–1073.
21. Mariotto S, Farinazzo A, Magliozzi R, Alberti D, Monaco S, Ferrari S. Serum and cerebrospinal neurofilament light chain levels in patients with acquired peripheral neuropathies. *J Peripher Nerv Syst*. 2018;23(3):174–177.
22. Hayer SN, Krey I, Barro C, et al. NfL is a biomarker for adult-onset leukoencephalopathy with axonal spheroids and pigmented glia. *Neurology*. 2018;91(16):755–757.
23. Bridel C, van Wieringen WN, Zetterberg H, et al.; and the NFL Group. Diagnostic value of cerebrospinal fluid neurofilament light protein in neurology: A systematic review and meta-analysis. *JAMA Neurol*. 2019;76(9):1035–1048.
24. Dang Do AN, Sinaii N, Masvekar RR, et al. Neurofilament light chain levels correlate with clinical measures in CLN3 disease. *Genet Med* 2020;23(4):751–757.
25. Wong YYM, Bruijstens AL, Barro C, et al. Serum neurofilament light chain in pediatric MS and other acquired demyelinating syndromes. *Neurology*. 2019;93(10):e968–e974.
26. Beerepoot S, van Dooren SJM, Salomons GS, et al. Metachromatic leukodystrophy genotypes in The Netherlands reveal novel pathogenic ARSA variants in non-Caucasian patients. *Neurogenetics*. 2020;21(4):289–299.
27. Kehrer C, Blumenstock G, Raabe C, Krägeloh-Mann I. Development and reliability of a classification system for gross motor function in children with metachromatic leucodystrophy. *Dev Med Child Neurol*. 2011;53(2):156–160.
28. Eichler F, Grodd W, Grant E, et al. Metachromatic leukodystrophy: A scoring system for brain MR imaging observations. *AJNR Am J Neuroradiol*. 2009;30(10):1893–1897.
29. Teunissen CE, Tumani H, Engelborghs S, Mollenhauer B. Biobanking of CSF: International standardization to optimize biomarker development. *Clin Biochem*. 2014;47(4-5):288–292.
30. Koel-Simmelink MJ, Vennegoor A, Killestein J, et al. The impact of pre-analytical variables on the stability of neurofilament proteins in CSF, determined by a novel validated SinglePlex Luminex assay and ELISA. *J Immunol Methods*. 2014;402(1-2):43–49.
31. Verberk IMW, Misdorp EO, Koelewijn JMW et al. Characterization of pre-analytical sample handling effects on a panel of Alzheimer's disease-related blood-based biomarkers: Results from the Standardization of Alzheimer's Blood Biomarkers (SABB) working group. *Alzheimers Dement*. Published online 29 November 2021. doi:10.1002/alz.12510.
32. Verberk IMW, Misdorp EO, Koelewijn JMW, et al. A biorepository for the in-depth validation of pre-analytical sample handling effects on novel blood-based biomarkers for Alzheimer's disease: The first results. *Alzheimer's & Dementia*. 2020;16(S5):e045763.
33. Kleinke K. Multiple imputation under violated distributional assumptions: A systematic evaluation of the assumed robustness of predictive mean matching. *J Educ Behav Stat*. 2017;42(4):371–404.
34. van der Vuurst de Vries RM, Wong YYM, Mescheriakova JY, et al. High neurofilament levels are associated with clinically definite multiple sclerosis in children and adults with clinically isolated syndrome. *Mult Scler*. 2019;25(7):958–967.
35. Olsson B, Alberg L, Cullen NC, et al. NFL is a marker of treatment response in children with SMA treated with nusinersen. *J Neurol*. 2019;266(9):2129–2136.
36. Pranzatelli MR, Tate ED, McGee NR, Verhulst SJ. CSF neurofilament light chain is elevated in OMS (decreasing with immunotherapy) and other pediatric neuroinflammatory disorders. *J Neuroimmunol*. 2014;266(1-2):75–81.
37. Veenhuis SJG, Gupta AS, de Gusmão CM, et al. Neurofilament light chain: A novel blood biomarker in patients with ataxia telangiectasia. *Eur J Paediatr Neurol*. 2021;32:93–97.
38. Al Nimer F, Thelin E, Nystrom H, et al. Comparative assessment of the prognostic value of biomarkers in traumatic brain injury reveals an independent role for serum levels of neurofilament light. *PLoS One*. 2015;10(7):e0132177.
39. van der Ende EL, Meeter LH, Poos JM, et al. Serum neurofilament light chain in genetic frontotemporal dementia: A longitudinal, multicentre cohort study. *Lancet Neurol*. 2019;18(12):1103–1111.
40. Weydt P, Oeckl P, Huss A, et al. Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis. *Ann Neurol*. 2016;79(1):152–158.
41. Preische O, Schultz SA, Apel A, et al.; Dominantly Inherited Alzheimer Network. Serum neurofilament dynamics predicts neurodegeneration and clinical progression in presymptomatic Alzheimer's disease. *Nat Med*. 2019;25(2):277–283.
42. Meregalli C, Fumagalli G, Alberti P, et al. Neurofilament light chain: A specific serum biomarker of axonal damage severity in rat models of chemotherapy-induced peripheral neurotoxicity. *Arch Toxicol*. 2020;94(7):2517–2522.
43. Thebault S, Lee H, Bose G, et al. Neurotoxicity after hematopoietic stem cell transplant in multiple sclerosis. *Ann Clin Transl Neurol*. 2020;7(5):767–775.
44. Beerepoot S, Nierkens S, Boelens JJ, Lindemans C, Bugiani M, Wolf NI. Peripheral neuropathy in metachromatic leukodystrophy: Current status and future perspective. *Orphanet J Rare Dis*. 2019;14(1):240.
45. Weinstock NI, Shin D, Dhimal N, et al. Macrophages expressing GALC improve peripheral Krabbe disease by a mechanism independent of cross-correction. *Neuron*. 2020;107(1):65–81.e9.
46. Kaminski D, Yaghootfam C, Matthes F, et al. Brain cell type-specific endocytosis of arylsulfatase A identifies limitations of enzyme-based therapies for metachromatic leukodystrophy. *Hum Mol Genet*. 2021;29(23):3807–3817.
47. van Rappard DF, Boelens JJ, Wolf NI. Metachromatic leukodystrophy: Disease spectrum and approaches for treatment. *Best Pract Res Clin Endocrinol Metab*. 2015;29(2):261–273.
48. Boucher AA, Miller W, Shanley R, et al. Long-term outcomes after allogeneic hematopoietic stem cell transplantation for

metachromatic leukodystrophy: The largest single-institution cohort report. *Orphanet J Rare Dis*. 2015;10(1):94.
49. Groeschel S, Kuhl JS, Bley AE, et al. Long-term outcome of allogeneic hematopoietic stem cell transplantation in patients with juvenile metachromatic leukodystrophy compared with nontransplanted control patients. *JAMA Neurology*. 2016;73(9):1133–1140.
50. Shahim P, Darin N, Andreasson U, et al. Cerebrospinal fluid brain injury biomarkers in children: A multicenter study. *Pediatr Neurol*. 2013;49(1):31–39.e2.
51. Lepinoux-Chambaud C, Eyer J. Review on intermediate filaments of the nervous system and their pathological alterations. *Histochem Cell Biol*. 2013;140(1):13–22.
52. Wolf NI, Ffrench-Constant C, van der Knaap MS. Hypomyelinating leukodystrophies - unravelling myelin biology. *Nat Rev Neurol*. 2021;17(2):88–103.
53. Bishop DL, Misgeld T, Walsh MK, Gan WB, Lichtman JW. Axon branch removal at developing synapses by axosome shedding. *Neuron*. 2004;44(4):651–661.