Clinical Value of Longitudinal Serum Neurofilament Light Chain in Prodromal Genetic Frontotemporal Dementia

Lucia A.A. Giannini, MD, Harro Seelaar, MD, PhD, Emma L. van der Ende, MD, PhD, Jackie M. Poos, PhD, Lize C. Jiskoot, PhD, Elise G.P. Dopper, MD, PhD, Yolande A.L. Pijnenburg, MD, PhD, Eline A.J. Willemsse, PhD, Lisa Vermunt, MD, PhD, Charlotte E. Teunissen, PhD, John C. van Swieten, MD, PhD, and Lieke H. Meeter, MD, PhD

Neurology® 2023;101:e1069-e1082. doi:10.1212/WNL.0000000000207581

Abstract

Background and Objectives
Elevated serum neurofilament light chain (NfL) is used to identify carriers of genetic frontotemporal dementia (FTD) pathogenic variants approaching prodromal conversion. Yet, the magnitude and timeline of NfL increase are still unclear. Here, we investigated the predictive and early diagnostic value of longitudinal serum NfL for the prodromal conversion in genetic FTD.

Methods
In a longitudinal observational cohort study of genetic FTD pathogenic variant carriers, we examined the diagnostic accuracy and conversion risk associated with cross-sectional and longitudinal NfL. Time periods relative to prodromal conversion (>3, 3–1.5, 1.5–0 years before; 0–1.5 years after) were compared with values of participants who did not convert. Next, we modeled longitudinal NfL and MRI volume trajectories to determine their timeline.

Results
We included 21 participants who converted (5 chromosome 9 open-reading frame 72 [C9orf72], 10 progranulin [GRN], 5 microtubule-associated protein tau [MAPT], and 1 TAR DNA-binding protein [TARDBP]) and 61 who did not (20 C9orf72, 30 GRN, and 11 MAPT). Participants who converted had higher NfL levels at all examined periods before prodromal conversion (median values 14.0–18.2 pg/mL; betas = 0.4–0.7, standard error [SE] = 0.1, p < 0.046) than those who did not (6.5 pg/mL) and showed further increase 0–1.5 years after conversion (28.4 pg/mL; beta = 1.0, SE = 0.1, p < 0.001). Annualized longitudinal NfL change was only significantly higher in participants who converted (vs. participants who did not) 0–1.5 years after conversion (beta = 1.2, SE = 0.3, p = 0.001). Diagnostic accuracy of cross-sectional NfL for prodromal conversion (vs. nonconversion) was good-to-excellent at time periods before conversion (area under the curve range: 0.72–0.92), improved 0–1.5 years after conversion (0.94–0.97), and outperformed annualized longitudinal change (0.76–0.84). NfL increase in participants who converted occurred earlier than frontotemporal MRI volume change and differed by genetic group and clinical phenotypes. Higher NfL corresponded to increased conversion risk (hazard ratio: cross-sectional = 6.7 [95% CI 3.3–13.7]; longitudinal = 13.0 [95% CI 4.0–42.8]; p < 0.001), but conversion-free follow-up time varied greatly across participants.

Discussion
NfL increase discriminates individuals who convert to prodromal FTD from those who do not, preceding significant frontotemporal MRI volume loss. However, NfL alone is limited in predicting the exact timing of prodromal conversion. NfL levels also vary depending on underlying variant-carrying genes and clinical phenotypes. These findings help to guide participant recruitment for clinical trials targeting prodromal genetic FTD.
Introduction

Frontotemporal dementia (FTD) is the second most common form of young-onset dementia.1 Clinically, it presents with prominent behavioral deficits (behavioral variant FTD [bvFTD])2 or with progressive language impairment (primary progressive aphasia [PPA]),3 and it overlaps clinically with amyotrophic lateral sclerosis (ALS). FTD has a genetic autosomal dominant etiology in approximately 30% of the cases, most often associated with genetic defects in chromosome 9 open-reading frame 72 (C9orf72), progranulin (GRN), or microtubule-associated protein tau (MAPT) genes, and rarely in the TAR DNA-binding protein (TARDBP) gene.4 Carriers of genetic FTD pathogenic variants are bound to develop the disease, but there is a considerable variability in age at onset within the same genetic group and even within the same family,5 and initial symptoms can be subtle and aspecific. Hence, reliable measures to predict and pinpoint disease onset are necessary.

Disease-onset biomarkers can be helpful in identifying individuals who are converting (or about to convert) from the presymptomatic to the prodromal stage.6 In genetic FTD, this is relevant for participant selection for clinical trials targeting specific gene defects that aim at participant inclusion in the prodromal stage, defined as the occurrence of mild cognitive and/or behavioral and/or motor impairment and corresponding to a Clinical Dementia Rating scale - frontotemporal lobar degeneration [FTLD-CDR] of 0.5,7,8 as opposed to the presymptomatic stage with no clinically overt FTD symptoms (FTLD-CDR = 0). Individuals in this stage may benefit the most from therapeutic interventions because early disease manifestations are present but with limited brain damage, so further neurodegeneration can still be prevented.9

Neurofilament light chain (NfL) has been identified as a useful disease-onset biomarker in genetic FTD.10-14 Levels of NfL, a neuroaxonal cytoskeleton protein, increase in blood and CSF along with the occurrence of neuronal damage.15 As increases in serum NfL occur in many neurologic disorders, including neurodegenerative, neurovascular, and inflammatory disorders,16 NfL as a biomarker is nonspecific. In genetic FTD, serum NfL increases before conversion to symptomatic disease.11 Therefore, NfL measurements are increasingly used in the cognitive clinic as a diagnostic tool to substantiate the early diagnosis of FTD (as opposed to non-neurological causes of cognitive symptoms) and for clinical trials to include participants in the prodromal stage or approaching this stage.7,8

However, because the prodromal stage is an emerging concept in FTD research, previous studies have not examined NfL levels in this phase specifically or included only a few prodromal individuals.15 In addition, the timeline of NfL increase in relation to this stage is still unclear. One recent study suggested that NfL increase precedes symptom onset by 15 years.12 This window is, however, too broad for clinical trials because these aim to recruit participants in closer proximity of symptom onset. Another large recent study found genotype-specific trajectories of biomarkers changes, including NfL, relative to estimated prodromal onset,13 yet it had limited resolution on the clinical implications of these changes, such as their predictive and diagnostic value for the prodromal stage in the individual patient.

Here, we aimed to investigate the clinical value of serum NfL for the prediction and early diagnosis of the prodromal stage in a well-characterized cohort of genetic FTD pathogenic variant carriers. We studied cross-sectional and longitudinal NfL levels and MRI volumes along the presymptomatic-to-prodromal transition. By highlighting the timeline and magnitude of NfL changes in participants who converted to prodromal FTD, we mirror real-life clinical decision-making and provide novel insights into the utility of NfL for candidate selection in clinical trials.

Methods

Participant

Participants were included from our longitudinal at-risk cohort for genetic FTD at the Erasmus University Medical Center, which longitudinally follows at-risk individuals from families with genetic FTD yearly or biyearly (FTD-RisC cohort; see eMethods, links.lww.com/WNL/C981).19 We included all carriers of FTD-related pathogenic variants (MAPT, C9orf72, GRN, and TARDBP) who were either pre-symptomatic or prodromal at the baseline visit and had longitudinal data available (at least 2 distinct time points), collected between June 2012 and December 2021. Only one included participant was prodromal at baseline (i.e., baseline visit corresponded to prodromal conversion) while the rest was presymptomatic. As we examined serum NfL across the presymptomatic to prodromal stage, we did not include...
participants who were fully symptomatic at baseline. We excluded 2 participants with additional comorbidities (1 substantial brain vascular damage and 1 alcohol abuse) potentially confounding clinical assessment and/or serum NFL levels.

Clinical data from included participants (n = 82) were reviewed systematically to determine whether conversion to prodromal stage occurred between the first and last available time points. Prodromal conversion was defined as follows: for cognitive prodromal participants, (1) cognitive symptoms/signs on clinical history or clinical evaluation (global CDR plus NACC FTLD20 [FTLD-CDR] of 0.5 with consistent scores and no fluctuations back to 0) and (2) impairment (i.e., ≥1.5 SD below age-specific, sex-specific, and education-specific means) or decline relative to a prior measurement in at least one domain on neuropsychological assessment; for motor prodromal participants, (1) mild motor symptoms/signs consistent with ALS or parkinsonism and (2) evidence of subsequent progression in time. The date of conversion to prodromal stage occurred between the first and last available time points.

In case of discrepancies between clinical information and neuropsychological test results for cognitive prodromal/symptomatic participants, clinical data were reviewed by 2 or more clinicians (H.S., L.A.A.G., L.H.M., and J.C.S.), and consensus was reached regarding conversion.

Serum NFL Measurement

Blood was collected and processed, and serum was stored according to a standardized protocol as previously described.11 NFL measurement in serum samples was performed in one laboratory by experienced laboratory analysts blinded to clinical and genetic information. Serum NFL was measured in 2 batches using the Simoa NF-Light Advantage Kit (Quanterix; Billerica, MA) on a Simoa HD-1 Analyzer instrument for the first batch (January 2018) or a Simoa HD-X for the second batch (March 2022), according to the manufacturer’s instructions and using identical protocols. In each batch, longitudinal samples from the same participant were included in the same run. To account for batch effects in NFL measurements, a subset of the samples (n = 109 from 54 participants) was measured in duplicate in both batches (eFigure 1, links.lww.com/WNL/C981), and linear regression transformation factors were derived and applied to batch 1 data (eFigure 2), obtaining optimal transformation outcomes (eTable 1). See eMethods for details on the transformation. All presented data are transformed data. The median number of NFL measurements per participant was 4 (interquartile range [IQR] 3–6), with a median time interval between samples of 1.2 years (IQR 1.0–2.0).

MRI Acquisition and (Pre)Processing

A total of 316 T1-weighted MRI scans were available from 79 participants (19 who converted and 60 who did not) within 6 months before or after NFL blood sample (mean interval sample-MRI = 0.0 ± 0.1 years). Most participants (17 who converted and 59 who did not) had longitudinal scans available (median number of scans = 4, IQR 3–6). Preprocessing steps were performed using SPM12 (v7771) and CAT12.8.1 (v1975) as described.20,23 Additional details on acquisition and preprocessing can be found in the eMethods (links.lww.com/WNL/C981). Gray matter volumes were obtained using the Hammers region-of-interest (ROIs) atlas24 and summed to obtain total bilateral volumes of cortical (frontal, temporal, parietal, occipital, insular, and cingulate cortices) and subcortical structures (amygdala, hippocampus, thalamus, basal ganglia, and cerebellum). Relative volumes for these structures were estimated as percentages of the total intracranial volume. Finally, w-scores were obtained as described previously (w-score = actual volume - expected volume for a given age/standard deviation of residuals in healthy controls)25 for each ROI based on a cohort of 1:1 age-matched and sex-matched healthy controls from the FTD-RisC cohort (32 male participants, 47 female participants, and mean age 48.7 ± 12.0 years). In addition, the last available MRI scan of each participant was graded for deep white matter lesions using the Fazekas scale26 based on FLAIR or T2-weighted sequences.

Statistical Analysis

All analyses were performed using R statistical software 4.2.2. Demographic and clinical data at baseline and at last available time point were compared between participants who converted and those who did not using Mann-Whitney U analysis or Chi-Square tests.

We examined both NFL raw levels and age-corrected z-scores. Serum NFL z-scores were obtained using normalization formulas derived from an independent cohort of reference age-matched controls (n = 1,698 healthy individuals aged 19–85 years) whose samples were measured at the same laboratory using identical protocols and instrumentation.27 z-score normalization formulas were as follows: (a) younger than 50 years, z-score = [log2(NFL) – (1.622+(age*0.023))]/0.699; (b) for 50 years and older: z-score = [log2(NFL) – (0.445 + (age * 0.047))] / 0.706.

We assessed cross-sectional NFL raw levels and age-adjusted z-scores at baseline and at the following clinically relevant periods relative to conversion: (1) >3 years before conversion (n = 10); (2) 3–1.5 years before conversion [-1.5; 0) (n = 16), (3) 1.5–0 years before conversion [-1.5; 0) (n = 11), and (4) [-1.5; 0) years after conversion [0; 1.5] (n = 19) and compared this with baseline NFL levels of participants who did not convert (n = 61). Annualized longitudinal change in NFL levels (i.e., difference between NFL raw levels at 2 subsequent time points divided by the time) was analyzed at the same time periods. Annualized change of participants who converted were compared with first available annualized change (i.e., delta between second and first time point) of participants who did not.

Neurology.org/N

Neurology | Volume 101, Number 10 | September 5, 2023 e1071
Mean cross-sectional raw NfL levels, z-scores, and annualized longitudinal change at different time periods were assessed using linear mixed modeling, which is suited for the analysis of unbalanced longitudinal data sets, accounting for missing data. We included age at sample collection and sex as covariates and a random intercept for each individual. The overall effect of the main predictor (i.e., time period relative to conversion) was assessed using type III ANOVA with the Satterthwaite method and post hoc pairwise comparisons using LME-derived least-square means with Tukey correction for multiple comparisons. Raw NfL levels underwent logarithmic transformation to obtain normally distributed data, whereas annualized longitudinal change underwent cubic root transformation because of negative values.

Diagnostic accuracy for conversion to prodromal FTD was evaluated with receiver operating curve (ROC) analyses, and optimal cutoff levels were determined using the Youden index. We compared ROC curves within the same time period using the Venkatraman and Begg method, and ROC curves between time periods using the DeLong method from the pROC package in R. Multimodal diagnostic accuracy, combining NfL and MRI volumes, was assessed on predicted estimates derived from a logistic regression model with clinical group (participants who converted vs participants who did not) as outcome and NfL z-scores and MRI data as predictors.

For all subsequent analyses, z-scores were used to account for the effect of age on NfL levels. Correlations between NfL z-scores and MRI volumes (across all time points) and longitudinal modeling of biomarker trajectories were performed using linear mixed modeling to account for repeated measurements from each participant (see model details in the eMethods, links.lww.com/WNL/C981).

The risk of prodromal conversion was analyzed using Cox regression models with prodromal conversion as event, follow-up time as time-to-event, baseline NfL z-scores or first-available annualized change as main predictor, and age and sex as covariates. In addition, a joint model was used to test the risk of prodromal conversion in relation to longitudinal NfL measurements (as time-dependent predictor). Kaplan-Meier curves were plotted to display conversion-free time in relation to NfL.

### Standard Protocol Approvals and Patient Consents

The study was approved by the Medical Ethics Review Committee of the Erasmus University Medical Center, and written informed consent was obtained from all participants.

### Data Availability

Data used for this article will be made available by the corresponding author in anonymized form on reasonable request.

### Results

#### Cohort Characterization

Our cohort (n = 82) consisted of 21 participants who converted to the prodromal stage (FTLD-CDR = 0.5: 5 C9orf72, 10 GRN, 5 MAPT, and 1 TARDBP) and 61 who did not (FTLD-CDR = 0: 20 C9orf72, 30 GRN, and 11 MAPT) (Table 1). Of the participants who converted, 13 (2 C9orf72, 6 GRN, and 5 MAPT) progressed to fully symptomatic stage...
during follow-up; 9 met criteria for bvFTD (4 GRN and 5 MAPT), 2 GRN for nonfluent variant PPA, 1 C9orf72 for ALS, and 1 C9orf72 for FTD-ALS. The remaining 8 participants who converted were still prodromal at last follow-up. Participants who converted had higher baseline age (median = 55.7 years) than those who did not convert (median = 45.8 years; \( p = 0.013; \) Mann-Whitney \( U \)) but did not differ in other baseline features and follow-up time. At last follow-up, participants who converted progressed to a median global FTLD-CDR of 0.5 (IQR 0.5–1.0) and had a median disease duration of 2.0 years (IQR 1.2–3.1).

**Cross-sectional NfL**

At all time periods before conversion (>3 years before, 3–1.5 years before, 1.5–0 years before; Figure 1), raw NfL levels were higher in participants who converted (>3 years before: median = 14.0 pg/mL, beta = 0.4, standard error [SE] = 0.1, \( p = 0.045; \) 3–1.5 years before: median = 15.9 pg/mL, beta = 0.6, SE = 0.1, \( p < 0.001; \) 1.5–0 years before: median = 18.2 pg/mL, beta = 0.7, SE = 0.1, \( p < 0.001; \) Figure 2A) compared with baseline of participants who did not convert (median = 6.5 pg/mL). NfL levels in participants who converted were similar between >3 years, 3–1.5 years, and 1.5–0 years before conversion (\( p > 0.2 \)). NfL levels 0–1.5 years after conversion (median = 28.4 pg/mL) showed a significant elevation relative to prior time periods (beta = 0.4–0.7, SE = 0.1, \( p < 0.03 \)) and relative to participants who did not convert (beta = 1.0, SE = 0.1, \( p < 0.001 \)). NfL z-scores showed analogous results (Figure 2B).

**Annualized Longitudinal NfL Change**

The median annualized longitudinal change in NfL levels was 2.2 pg/mL/y >3 years before, 2.4 pg/mL/y 3–1.5 years before, and 2.5 pg/mL/y 1.5–0 years before conversion to prodromal
Figure 2 Cross-sectional NfL Raw Levels, z-Scores, and Annualized Longitudinal Change at Time Periods Relative to Conversion

Boxplots depict cross-sectional comparisons of NfL raw levels (A), z-scores (B) and annualized change (C) at time periods relative to conversion. Data points are color-coded by mutated FTD gene. For both NfL raw levels and z-scores, participants who converted at all time periods differed significantly from reference values of those who did not. For annualized change, participants who converted 0–1.5 years after conversion differed significantly from reference values of those who did not. ***p < 0.001, **p < 0.01, *p < 0.05; statistical outcomes from a linear mixed-effects model with individuals as random intercept, time period as main predictor, and age at sample and gender as covariates. Tukey correction for multiple comparisons was applied. FTD = frontotemporal dementia; NfL = neurofilament light chain.
Table 2 Diagnostic Accuracy of NFL and MRI Volumes at Time Periods Relative to Conversion

<table>
<thead>
<tr>
<th>Time Period</th>
<th>N conv/non-conv</th>
<th>AUC</th>
<th>AUC 95% CI</th>
<th>Cutoff</th>
<th>Sens</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&gt;3 y before conversion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL raw scores</td>
<td>10/61</td>
<td>0.86</td>
<td>0.75–0.98</td>
<td>8.9</td>
<td>80.0</td>
<td>77.1</td>
</tr>
<tr>
<td>NFL z-scores</td>
<td>10/61</td>
<td>0.72</td>
<td>0.52–0.91</td>
<td>0.3</td>
<td>70.0</td>
<td>70.5</td>
</tr>
<tr>
<td>NFL annualized change</td>
<td>5/61</td>
<td>0.78</td>
<td>0.51–1.00</td>
<td>0.7</td>
<td>80.0</td>
<td>73.8</td>
</tr>
<tr>
<td>MRI frontal w-scores*</td>
<td>9/58</td>
<td>0.28</td>
<td>0.12–0.44</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>MRI temporal w-scores</td>
<td>9/58</td>
<td>0.42</td>
<td>0.22–0.61</td>
<td>−0.6</td>
<td>66.7</td>
<td>39.7</td>
</tr>
<tr>
<td>MRI frontal w-scores &amp; NFL z-scores</td>
<td>9/58</td>
<td>0.79&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.65–0.93</td>
<td>100.0</td>
<td>51.7</td>
<td></td>
</tr>
<tr>
<td>MRI temporal w-scores &amp; NFL z-scores</td>
<td>9/58</td>
<td>0.75&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.56–0.94</td>
<td>55.6</td>
<td>91.4</td>
<td></td>
</tr>
<tr>
<td><strong>3–1.5 y before conversion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL raw scores</td>
<td>16/61</td>
<td>0.90</td>
<td>0.80–1.00</td>
<td>12.3</td>
<td>81.3</td>
<td>93.4</td>
</tr>
<tr>
<td>NFL z-scores</td>
<td>16/61</td>
<td>0.91</td>
<td>0.81–1.00</td>
<td>0.6</td>
<td>81.3</td>
<td>93.4</td>
</tr>
<tr>
<td>NFL annualized change</td>
<td>9/61</td>
<td>0.76</td>
<td>0.52–0.99</td>
<td>2.4</td>
<td>55.6</td>
<td>95.1</td>
</tr>
<tr>
<td>MRI frontal w-scores</td>
<td>13/58</td>
<td>0.51</td>
<td>0.32–0.70</td>
<td>−1.6</td>
<td>30.8</td>
<td>84.5</td>
</tr>
<tr>
<td>MRI temporal w-scores</td>
<td>13/58</td>
<td>0.65</td>
<td>0.51–0.79</td>
<td>−1.0</td>
<td>84.6</td>
<td>56.9</td>
</tr>
<tr>
<td>MRI frontal w-scores &amp; NFL z-scores</td>
<td>13/58</td>
<td>0.94&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.88–1.00</td>
<td>84.6</td>
<td>91.4</td>
<td></td>
</tr>
<tr>
<td>MRI temporal w-scores &amp; NFL z-scores</td>
<td>13/58</td>
<td>0.94&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.88–1.00</td>
<td>84.6</td>
<td>94.8</td>
<td></td>
</tr>
<tr>
<td><strong>1.5–0 y before conversion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL raw scores</td>
<td>11/61</td>
<td>0.92</td>
<td>0.81–1.00</td>
<td>12.6</td>
<td>90.9</td>
<td>95.1</td>
</tr>
<tr>
<td>NFL z-scores</td>
<td>11/61</td>
<td>0.86</td>
<td>0.72–1.00</td>
<td>0.8</td>
<td>72.7</td>
<td>96.7</td>
</tr>
<tr>
<td>NFL annualized change</td>
<td>9/61</td>
<td>0.79</td>
<td>0.57–1.00</td>
<td>2.4</td>
<td>66.7</td>
<td>95.1</td>
</tr>
<tr>
<td>MRI frontal w-scores</td>
<td>10/58</td>
<td>0.42</td>
<td>0.27–0.57</td>
<td>0.2</td>
<td>100.0</td>
<td>20.7</td>
</tr>
<tr>
<td>MRI temporal w-scores</td>
<td>10/58</td>
<td>0.66</td>
<td>0.52–0.81</td>
<td>−0.7</td>
<td>100.0</td>
<td>41.4</td>
</tr>
<tr>
<td>MRI frontal w-scores &amp; NFL z-scores</td>
<td>10/58</td>
<td>0.85&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.70–1.00</td>
<td>70.0</td>
<td>96.6</td>
<td></td>
</tr>
<tr>
<td>MRI temporal w-scores &amp; NFL z-scores</td>
<td>10/58</td>
<td>0.87&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.74–1.00</td>
<td>70.0</td>
<td>96.6</td>
<td></td>
</tr>
<tr>
<td><strong>0–1.5 y after conversion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL raw scores</td>
<td>19/61</td>
<td>0.97</td>
<td>0.92–1.00</td>
<td>14.2</td>
<td>94.7</td>
<td>98.4</td>
</tr>
<tr>
<td>NFL z-scores</td>
<td>19/61</td>
<td>0.94&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.87–1.00</td>
<td>0.7</td>
<td>84.2</td>
<td>95.1</td>
</tr>
<tr>
<td>NFL annualized change</td>
<td>18/61</td>
<td>0.84</td>
<td>0.70–0.99</td>
<td>1.4</td>
<td>77.8</td>
<td>88.5</td>
</tr>
<tr>
<td>MRI frontal w-scores</td>
<td>14/58</td>
<td>0.66</td>
<td>0.50–0.81</td>
<td>−0.9</td>
<td>71.4</td>
<td>60.3</td>
</tr>
<tr>
<td>MRI temporal w-scores</td>
<td>14/58</td>
<td>0.74</td>
<td>0.62–0.86</td>
<td>−1.1</td>
<td>92.9</td>
<td>62.1</td>
</tr>
<tr>
<td>MRI frontal w-scores &amp; NFL z-scores</td>
<td>14/58</td>
<td>0.92&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.82–1.00</td>
<td>78.6</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>MRI temporal w-scores &amp; NFL z-scores</td>
<td>14/58</td>
<td>0.94&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.87–1.00</td>
<td>85.7</td>
<td>93.1</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: AUC = area under the curve; conv = participants who converted; non-conv = participants who did not convert; Sens = sensitivity; Spec = specificity. The 95% CI was estimated using the DeLong method and computed with 2000 stratified bootstrap replicates. The optimal cutoff was determined as to maximize the Youden index.

* Because of the low discriminative ability (AUC = 0.28) of MRI frontal w-scores >3 y before conversion, the optimal cutoff and sensitivity/specificity values could not be estimated.

<sup>b</sup> ROC statistics of combined MRI frontal/temporal w-scores & NFL z-scores improved significantly compared with MRI frontal/temporal w-scores only (p < 0.05 with Venkatraman and Begg test).

<sup>c</sup> ROC statistics of NFL z-scores at time period 0–1.5 y after conversion improved significantly compared with NFL z-scores at time period >3 y before conversion (p = 0.040 with DeLong test).
FTD, similar in all these time periods \((p > 0.9)\). Annualized change at these time periods was not significantly higher compared with baseline of participants who did not convert \((\text{median} = 0.1 \text{ pg/mL/y}; \ p > 0.1)\). The median annualized change 0–1.5 years after conversion was 5.5 pg/mL/y, which did not differ significantly from previous time periods \((p > 0.9)\) but was significantly higher than reference annualized change in participants who did not convert \((\beta = 1.2, \ SE = 0.3, \ p = 0.001; \text{Figure 2C})\).

**Diagnostic Accuracy of NfL**

NfL raw levels, z-scores, and annualized change had good-to-excellent diagnostic accuracy for distinguishing participants who converted to prodromal FTD from those who did not convert at all time periods (Table 2). NfL raw levels had relatively higher diagnostic accuracy at all time periods compared with z-scores and annualized longitudinal change. Diagnostic accuracy increased while approaching prodromal conversion and was greatest 0–1.5 years after conversion for all 3 measures \((\text{area under the curves } [\text{AUCs}] 0.84–0.97)\). Particularly, sensitivity of NfL raw levels and z-scores for prodromal conversion was relatively lower \((70%–91\%)\) at time periods before conversion than after conversion \((84%–95\%)\). Sensitivity of annualized change remained below 80% even after conversion. Conversely, specificity of all 3 measures was consistently close to or above 90% across all time periods, except for the earliest time period \((>3\text{ years before})\). Diagnostic accuracy of NfL z-scores was significantly better in the period after conversion \((0–1.5\text{ years after})\) compared with the earliest time period \((>3\text{ years before}; \ p = 0.040)\). There were no other significant differences in diagnostic accuracy values between tested measures \((\text{raw levels vs z-scores vs annualized change})\) at the same time period nor between the same measure at different time periods.
Figure 4  Linear Mixed Modeling Effects Plot  Showing Predicted Longitudinal NfL Trajectories in Participants Who Converted of Different FTD Gene Groups

A

C9orf72

GRN

MAPT

Gene

C9orf72

GRN

MAPT

Effects plot from linear mixed modeling shows (A) the predicted longitudinal trajectories of NfL z-scores in different FTD gene groups in participants who converted along time to prodromal conversion (modeled nonlinearly) and (B) a subanalysis in the C9orf72 group excluding 2 participants with a motor phenotype (1 ALS, 1 FTD-ALS). The black dashed line indicates the time of prodromal conversion. NfL longitudinal trajectories of participants who converted showed a different course depending on FTD gene group (p = 0.001). FTD = frontotemporal dementia; NfL = neurofilament light chain.

MRI Volumes Related to NfL

Higher NfL z-scores correlated with smaller volume w-scores of frontal, temporal, parietal, insular, and cingulate cortices and of the amygdala, hippocampus, and basal ganglia (betas = −0.2 to −0.4, SE = 0.0–0.1, Bonferroni-corrected p < 0.001) while no significant correlation was found with occipital, thalamic, and cerebellar volumes (Bonferroni-corrected p > 0.05; eTable 2, links.lww.com/WNL/C981). All subsequent analyses were performed using frontal and temporal volumes only, which showed the strongest association with NfL.

By modeling longitudinal NfL z-scores and longitudinal MRI w-scores along participant age, we observed different trajectories between participants who converted and those who did not (significant interaction between clinical group and participant age; NfL: F(2,45) = 9.6; frontal: F(2,50) = 17.6; temporal: F(2,30) = 12.1; p < 0.001 for all; Figure 3, A–C). The initial increase in NfL levels could not be captured based on our data because NfL levels were already increased around 40 years of age in participants who converted (Figure 3A) while the relative decrease in MRI volumes occurred later, that is, between age 45–50 years for frontal and temporal volumes (Figure 3, B and C). As such, the relative increase in NfL levels between participants who converted and those who did not started earlier than the relative decrease in frontal and temporal volumes (Figure 3D). Accordingly, diagnostic accuracy of frontal and temporal w-scores was overall lower (AUCs 0.28–0.74; Table 2) than diagnostic accuracy of NfL levels and improved significantly when combining MRI volumes with NfL (AUCs 0.75–0.94, p < 0.05; Table 2).
NfL Longitudinal Trajectories Related to FTD Gene and Clinical Phenotype

Through exploratory modeling of NfL along time to disease onset, we found that NfL longitudinal trajectories of participants who converted had a variable course depending on FTD genetic group ($F(6,40) = 4.6, p = 0.001$; Figure 4A). In a subanalysis excluding C9orf72 participants with a motor phenotype (1 ALS and 1 FTD-ALS), we observed that the elevation in NfL levels in the C9orf72 group was less pronounced (Figure 4B).

Next, we examined individual longitudinal NfL trajectories (Figure 5). In 3 of 21 (14%) participants who converted (1 C9orf72, 2 GRN), NfL z-score shortly after conversion was lower than our ROC-defined cutoff of 0.7 (Table 2). Remarkably, these participants showed a relatively slow

---

Figure 5 Individual Longitudinal Trajectories of NfL z-Scores per Gene Group

Plots show longitudinal NfL z-scores trajectories in each gene group along time to prodromal conversion (years) for participants who converted and along follow-up time (years) for participants who did not. The black dashed line indicates the time of prodromal conversion in the group of participants who converted. The red dashed line indicates the 0.7 cutoff having optimal diagnostic accuracy 0–1.5 years after conversion to distinguish participants who converted from those who did not (AUC 0.94). The orange-colored area signals the time period of 0–1.5 years after conversion. Data points are shape-coded for clinical state (presymptomatic vs prodromal vs fully symptomatic) and color-coded for the profile of most prominent clinical symptomatology. NfL = neurofilament light chain.
prodromal presentation because they had been prodromal for longer than 1 year and had not progressed to fully symptomatic disease (FTLD-CDR = 1) yet. The C9orf72 participant was a 45-year-old woman (at the time of conversion) with mainly mild neuropsychiatric symptoms next to very mild behavioral features. The 2 GRN participants were (1) a 75-year-old man with mild behavioral features (i.e., irritable, childish, and inflexible behavior) and slightly increasing NfL at the subsequent follow-up from NfL z-score of 0.1 to a score of 0.6, i.e., still below cutoff and (2) a 64-year-old man with mild word-finding and short-term memory difficulties, who developed subtle behavioral changes (i.e., apathy, decreased initiative) at the subsequent follow-up while NfL remained relatively stable (NfL z-scores = 0.5–0.6).

The unique TARDBP participant who converted, not included in the longitudinal modeling analysis because of low sample size, showed an increasing NfL z-score around prodromal conversion, above the 0.7 z-score cutoff (Figure 5).

**Prediction of Prodromal Conversion**

Higher baseline NfL z-scores were associated with increased risk for prodromal conversion (hazard ratio [HR] 6.7 [95% CI 3.3–13.7], \( p < 0.001 \)), more strongly than higher first-available annualized change (HR 1.3 [95% CI 1.1–1.5], \( p < 0.001 \)). Of participants with a baseline NfL z-score \( \geq 0.7 \), 50% converted within 2.6 years (eFigure 3A, links.lww.com/WNL/C981) while 50% of those with first-available annualized change \( \geq 1.4 \) within 5.7 years (eFigure 3B). Longitudinal NfL z-score increase was also associated with increased risk for conversion (HR 13.0 [4.0–42.8], \( p < 0.001 \)). Of participants with a longitudinal slope \( \geq 0.6 \), 50% converted within 3.6 years (eFigure 3C). All 3 analyses showed broad confidence intervals for conversion-free follow-up time (eFigure 3), suggesting low precision in predicting the timing of conversion.

**Discussion**

This study explored the clinical value of longitudinal serum NfL in predicting and diagnosing the prodromal stage of genetic FTD. We found higher NfL levels in participants who converted relative to those who did not already \( > 3 \) years before prodromal conversion. These levels remain relatively stable in the years before conversion and undergo a further increase shortly after conversion. Accordingly, NfL has good-to-excellent diagnostic accuracy for prodromal conversion, highest after conversion. However, NfL measurement before conversion is limited in predicting the exact timing of prodromal conversion. Longitudinal modeling showed that NfL in individuals who convert becomes abnormal at an earlier age compared with MRI volumes. Finally, NfL trajectories of individuals who convert differ partly by genetic group and clinical phenotype.

NfL is higher in individuals who convert several years before prodromal conversion compared with those who do not. This is consistent with previous studies describing NfL increase in the years before clinical diagnosis \( ^{11,12,14} \) or elevated baseline NfL levels in individuals who convert compared with those who do not. \( ^{14,33} \) Moreover, longitudinal modeling suggests that NfL increases 15 years before disease onset, \( ^{13} \) or even 30 years in the C9orf72 group, compared with healthy controls. \( ^{18} \)

In contrast to these previous studies, we examined NfL in pathogenic variant carriers only and compared participants who converted with those who did not, reflecting clinical decision-making for clinical trial selection. In this group, NfL has an important role for both prediction and early diagnosis of prodromal FTD because the origin of subtle prodromal symptoms cannot always be ascertained through clinical assessment alone.

NfL has good-to-excellent diagnostic accuracy for prodromal genetic FTD before conversion (AUC 0.72–0.92) and excellent on conversion (AUC 0.94–0.97). Previous studies found lower diagnostic accuracy (AUC 0.68–0.78), \( ^{14,33} \) which may be due to differences in the definition of prodromal conversion. These studies classified participants based on FTLD-CDR scores, which are in part reliant on heterogeneous informant-based history and may show fluctuations in longitudinal visits. \( ^{7} \)

To account for these challenges, here we reviewed each participant and reached consensus on the clinical state, especially when discrepancies were present between informant-based history and clinical observations. Furthermore, diagnostic performance of raw NfL levels should be interpreted with caution because of the age-dependent increase in NfL. \( ^{17,27} \) The higher diagnostic accuracy of raw NfL levels (compared with z-scores) may partly stem from the older age of participants who converted compared with those who did not. Importantly, we observed that annualized longitudinal NfL change was less helpful than cross-sectional NfL in differentiating participants who converted from those who did not, especially before conversion. This observation is consistent with our finding that NfL levels remain relatively stable in this timeframe. Furthermore, although high levels of both cross-sectional and longitudinal NfL were associated with a greater risk of conversion, they were limited in predicting its exact timing, as demonstrated by the large variation in conversion-free follow-up time across participants.

NfL levels reflect MRI volume loss in FTD-associated regions and become abnormal at an earlier stage in individuals who convert compared with frontotemporal MRI volumes. Similar to previous reports, we found negative associations between NfL levels and gray matter volumes in FTD-associated cortical and subcortical regions. \( ^{10,11} \) We could not capture the initial increase in NfL levels in participants who converted, earlier than 40 years of age, while significant decrease in MRI frontotemporal volumes appeared only later. Moreover, MRI volumes had lower diagnostic accuracy for prodromal conversion, which may be due to relative volume deficits early on in some of the participants who did not convert, as reported before especially for C9orf72 and MAPT presymptomatic carriers. \( ^{18,34,35} \)

A previous study found that NfL and MRI...
advances occur closely to each other across all FTD genetic groups, by modeling cross-sectional data over participant age in carriers vs controls of the same age. Here, conversely, we aimed to specifically highlight the changes observable within the risk age range in carriers who convert compared with those further away from conversion. Direct comparisons of these groups across the whole genetic FTD spectrum may, however, be limited because of considerable gene-related heterogeneity. Variability in NfL longitudinal trajectories in individuals who convert appears related to the variant-carrying gene and partly to the clinical course. We found different trajectories of NfL levels in FTD genetic groups, characterized by a steep rapid increase after conversion in C9orf72, a steady increase in GRN, and a gradual increase followed by a plateau in MAPT, in line with previous research showing different NfL trajectories between genetic groups. In addition, clinical phenotypes influence NfL trajectories. The steep increase observed in C9orf72 was primarily driven by 2 (FTD-)ALS participants, that have been associated with highest NfL levels. Notably, 3 of 21 participants who converted had NfL levels below our established cutoff, suggesting that low NfL does not exclude prodromal conversion. Of these participants, one C9orf72 participant had low, relatively stable NfL levels, as reported earlier in association with a neuropsychiatric presentation. The other 2 participants, with GRN pathogenic variants, presented clinically with subtle and slowly progressive symptoms. Disease stage and progression rate may, therefore, account for some variability in NfL in the GRN group, as reported before too. Our observations on gene-related and phenotype-related differences in NfL trajectories should be confirmed and further explored in larger multicenter cohorts.

The main strength of this study is the well-characterized nature of our single-center cohort. The availability of ample clinical information enabled us to systematically review clinical data for the identification of prodromal participants through consensus. Despite the single-center design, the number of participants who converted (n = 21) was the largest described thus far in a longitudinal study and included substantial follow-up data (median 5.7 years) enabling to combine cross-sectional analyses at time periods of interest with longitudinal modeling of NfL and MRI data. Finally, our focus on prodromal genetic FTD is novel and important because this stage is relevant for clinical trials targeting the prodromal stage of genetic FTD, but with some limitations. NfL levels alone do not reliably predict the timing of prodromal conversion, and NfL trajectories have a variable course depending on the variant-carrying FTD gene and the clinical phenotype. These aspects should be taken into account for the clinical interpretation of NfL levels and for patient selection in clinical trials targeting the prodromal stage of genetic FTD.

Acknowledgments We thank all study participants and their families for contributing to this study. We also thank Hans Heijst (Amsterdam UMC, Vrije Universiteit) and Shamiram Melhem (Erasmus University Medical Center) for technical assistance with the collection and analysis of the samples.

Study Funding Memorabel grants from Deltraplan Dementie (ZonMw and Alzheimer Nederland; grant numbers 733050813, 733050103, 73050513), the Bluefield Project to Cure Frontotemporal Dementia, the Dioraphé foundation (grant number 1402 1300), and the European Joint Programme—Neurodegenerative Disease Research and the Netherlands Organisation for Health Research and Development (PreFrontALS: 733051042, RiModFTD: 733051024).

Disclosure The authors report no relevant disclosures. Go to Neurology.org/N for full disclosures.

Publication History Received by Neurology January 19, 2023. Accepted in final form May 10, 2023. Submitted and externally peer reviewed. The handling editor was Associate Editor Linda Hershey, MD, PhD, FAAN.
### Appendix (continued)

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>John C. van Swieten, MD, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data</td>
</tr>
<tr>
<td>Lieke H. Meeter, MD, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data</td>
</tr>
</tbody>
</table>

### References


---

**Appendix**

**Authors**

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucia A.A. Giannini, MD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data</td>
</tr>
<tr>
<td>Harro Seelaar, MD, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data</td>
</tr>
<tr>
<td>Emma L. van der Ende, MD, PhD</td>
<td>Amsterdam Neuroscience, Neurodegeneration; Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, The Netherlands</td>
<td>Analysis or interpretation of data</td>
</tr>
<tr>
<td>Jackie M. Poos, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Major role in the acquisition of data</td>
</tr>
<tr>
<td>Lize C. Jiskoot, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Elise G.P. Dopper, MD, PhD</td>
<td>Department of Neurology, Alzheimer Center Erasmus MC, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
<td>Major role in the acquisition of data</td>
</tr>
<tr>
<td>Yolande A.L. Pijnenburg, MD, PhD</td>
<td>Amsterdam Neuroscience, Neurodegeneration; Alzheimer Center Amsterdam, Amsterdam, Amsterdam, Amsterdam UMC location VUMc, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data</td>
</tr>
<tr>
<td>Elinne A.J. Willemse, PhD</td>
<td>Amsterdam Neuroscience, Neurodegeneration; Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, The Netherlands</td>
<td>Analysis or interpretation of data</td>
</tr>
<tr>
<td>Lisa Vermunt, MD, PhD</td>
<td>Amsterdam Neuroscience, Neurodegeneration; Neurochemistry Laboratory, Department of Clinical Chemistry, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, The Netherlands</td>
<td>Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data</td>
</tr>
</tbody>
</table>


Clinical Value of Longitudinal Serum Neurofilament Light Chain in Prodromal Genetic Frontotemporal Dementia
Lucia A.A. Giannini, Harro Seelaar, Emma L. van der Ende, et al.
Neurology 2023;101:e1069-e1082 Published Online before print July 25, 2023
DOI 10.1212/WNL.00000000000207581

This information is current as of July 25, 2023

Updated Information & Services
including high resolution figures, can be found at:
http://n.neurology.org/content/101/10/e1069.full

References
This article cites 42 articles, 13 of which you can access for free at:
http://n.neurology.org/content/101/10/e1069.full#ref-list-1

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Frontotemporal dementia
http://n.neurology.org/cgi/collection/frontotemporal_dementia

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
http://www.neurology.org/about/about_the_journal#permissions

Reprints
Information about ordering reprints can be found online:
http://n.neurology.org/subscribers/advertise