Diagnostic and prognostic value of serum NfL and p-Tau\textsubscript{181} in frontotemporal lobar degeneration

Alberto Benussi,1 Thomas K Karikari,2 Nicholas Ashton,2,3,4,5 Stefano Gazzina,6 Enrico Premi,7 Luisa Benussi,8 Roberta Ghidoni,8 Juan Lantero Rodriguez,2 Andreja Emeršič,2,9 Joel Simrén,10 Giuliano Binetti,11 Silvia Fostinelli,8 Marcello Giunta,1 Roberto Gasparotti,12 Henrik Zetterberg,2,10,13,14 Kaj Blennow,2,10 Barbara Borroni1

ABSTRACT

Objective To assess the diagnostic and prognostic value of serum neurofilament light (NfL) and serum phospho-Tau\textsubscript{181} (p-Tau\textsubscript{181}) in a large cohort of patients with frontotemporal lobar degeneration (FTLD).

Methods In this retrospective study, performed on 417 participants, we analysed serum NfL and p-Tau\textsubscript{181} concentrations with an ultrasensitive single molecule array (Simoa) approach. We assessed the diagnostic values of serum biomarkers in the differential diagnosis between FTLD, Alzheimer’s disease (AD) and healthy ageing; their role as markers of disease severity assessing the correlation with clinical variables, cross-sectional brain imaging and neurophysiological data; their role as prognostic markers, considering their ability to predict survival probability in FTLD.

Results We observed significantly higher levels of serum NfL in patients with FTLD syndromes, compared with healthy controls, and lower levels of p-Tau\textsubscript{181} compared with patients with AD. Serum NfL concentrations showed a high accuracy in discriminating between FTLD and healthy controls (area under the curve (AUC): 0.86, p<0.001), while serum p-Tau\textsubscript{181} showed high accuracy in differentiating FTLD from patients with AD (AUC: 0.93, p<0.001). In FTLD, serum NfL levels correlated with measures of cognitive function, disease severity and behavioural disturbances and were associated with frontotemporal atrophy and indirect measures of GABAergic deficit. Moreover, serum NfL concentrations were identified as the best predictors of survival probability.

Conclusions The assessment of serum NfL and p-Tau\textsubscript{181} may provide a comprehensive view of FTLD, aiding in the differential diagnosis, in staging disease severity and in defining survival probability.

INTRODUCTION

Frontotemporal lobar degeneration (FTLD) encompasses a series of early onset progressive neurodegenerative conditions for which, in the last decade, the diagnostic workup has substantially changed with the publication of revised clinical criteria.1 2 The careful characterisation of clinical features of the behavioural variant frontotemporal dementia (bvFTD), the agrammatic or the semantic variant of primary progressive aphasia (avPPA and svPPA), and the spectrum of FTLD with extrapyramidal symptoms, such as corticobasal syndrome (CBS) and progressive supranuclear palsy (PSP), has enabled a better understanding of the heterogeneity of FTLD phenotypes.

The pattern of brain atrophy and hypometabolism,3 and the results of new positron emission tomography tracers,4 have assisted in increasing the diagnostic accuracy of FTLD, while AB\textsubscript{42} or tau measurements in cerebrospinal fluid (CSF) have been proven to be key in ruling out Alzheimer’s disease (AD). Furthermore, the identification of monogenic FTLD, due to pathogenetic mutations within the granulin (GRN), chromosome 9 open reading frame 72 (C9orf72) or microtubule-associated protein tau (MAPT), has undoubtedly contributed to the diagnostic workup.5

Considering the possible drawbacks of these supportive biomarkers due to invasiveness, availability or expensiveness, there is an urgent need to identify robust and accessible screening tests to be used even in the earliest disease stages, in a disorder that is much more frequent than previously thought.6

Along with recently proposed neurophysiological markers, measuring FTLD-related neurotransmitter deficits by transcranial magnetic stimulation (TMS),7 a giant step forward towards potentially useful biomarkers for AD-related pathologies has been made with the new ultrasensitive single molecule array (Simoa) approach and the discovery of potentially useful blood-based biomarkers. It has been reported that concentrations of neurofilament light (NFL), a marker of axonal damage which is measurable in CSF, plasma or serum, are increased in FTLD and may be related to parameters of disease severity and prognosis.8–12

Furthermore, a Meso-Scale Discovery assay for plasma phosphorylated tau (p-Tau\textsubscript{181}) developed by Lilly Research Laboratories was found to differentiate AD from healthy controls (HC), suggesting its ability to identify mixed 3R/4R tau pathology.13 Recent studies have further highlighted the usefulness of this biomarker assay in FTLD, in differentiating FTLD from AD, and in monitoring disease progression.14–16 A study employing a single molecule array (Simoa) assay developed at the University of Gothenburg, also found a marked increase in plasma p-Tau\textsubscript{181} in AD, correlating with tau PET ligand retention, while levels were normal in other tauopathies including FTD and PSP.17

© Authors (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.
This retrospective study aimed at confirming and extending previous literature data, comprehensively assessing the clinical value of serum NfL and serum p-Tau_{181} in a large cohort of patients with FTLD. We discuss when either serum NfL or serum p-Tau_{181} should be considered on clinical grounds on the basis of specific clinical questions and defined outcomes. We analysed three main aims: a) the role of serum NfL as diagnostic markers, evaluating their accuracy in detecting FTLD and excluding AD, and, most importantly, their usefulness in the earliest disease stages; b) their role as markers of disease severity, assessing the correlation with clinical variables, cross-sectional brain imaging and neuropsychological data; c) their role as prognostic markers, considering their ability to predict survival probability in FTLD.

MATERIALS AND METHODS

Subjects
This retrospective study included 417 participants from two independent cohorts, 307 from the Centre for Neurodegenerative Disorders, University of Brescia, Italy and 110 from the IRCCS Istituto San Giovanni di Dio Fatebenefratelli, Brescia, Italy.

The cohort consisted of 291 patients meeting probable clinical criteria for a syndrome in the FTLD spectrum, namely 134 bvFTD, 48 avPPA, 27 svPPA, 51 CBS and 31 PSP. Moreover, 63 patients fulfilling clinical criteria for AD, and 63 HC, recruited among spouses or caregivers, were included as well.

Each patient with FTLD underwent a neurological evaluation, routine laboratory examination and a neuropsychological and behavioural assessment. In all cases, the diagnosis was supported by brain structural imaging, while CSF concentrations of tau, p-Tau_{181} and A\textsubscript{β}_{1-42} were measured in a subset of cases (n=133, 45.7%), to rule out AD, as previously reported. Furthermore, in familial cases (based on the presence of at least one dementia degree relatives) and early onset sporadic cases, genetic screening for GRN, C9orf72 and MAPT P301L mutations was performed; given the low frequency of MAPT mutations in Italy, we considered only the P301L mutation and we sequenced the entire MAPT gene only in selected cases.

Each participant underwent blood collection for measurements of serum NfL and p-Tau_{181} biomarkers, and a subset of FTLD patients underwent standardised brain MRI at baseline (n=132) to evaluate the correlation between serum biomarkers and imaging data. Moreover, a subgroup of patients underwent TMS protocols (n=113) to assess the correlation between serum biomarkers and neurophysiological data. For the purpose of the present study, we considered TMS measures that partially and indirectly reflect the activity of several neurotransmitters, including GABA\textsubscript{A} by short interval intracortical inhibition (SICI), glutamate by intracortical facilitation (ICF), GABA\textsubscript{B} by long interval intracortical inhibition (LICI) and acetylcholine by short latency afferent inhibition (SAI).

Clinical evaluation
At baseline patients underwent a standardised neuropsychological battery which included the Mini-Mental State Examination (MMSE), the short story recall test, the Rey complex figure (copy and recall), phonemic and semantic fluencies, the token test, the clock-drawing test and trail-making test (part A and part B). Disease severity was assessed with the FTLD modified clinical dementia rating (FTLD-CDR) sum of boxes scale, while the level of functional independence was assessed with the basic activities of daily living (BADL) and the instrumental activities of daily living (IADL) questionnaires. Furthermore, neuropsychiatric and behavioural disturbances were evaluated with the frontal behaviour inventory (FBI).

HCs underwent a brief standardised neuropsychological assessment (MMSE ≥27/30); psychiatric or other neurological illnesses were considered exclusion criteria.

Serum biomarkers
Serum was collected by venipuncture, processed and stored in aliquots at −80°C according to standardised procedures. Serum NfL was measured using a commercial NF-light assay (Quanterix, Lexington, Massachusetts, USA) according to the manufacturer’s instructions. Serum p-Tau_{181} was measured using an in-house Simoa assay developed at the University of Gothenburg. In brief, the capture antibody (AT270, Invitrogen) which is specific for the threonine-181 phosphorylation site was coupled to paramagnetic beads while the detector antibody (Tau12, BioLegend) was raised against the N-terminal epitope amino acid 6-QEFEVMEDHAGT-18 on human tau protein. Detailed analytical procedures and assay validation have been previously described. The lower limits of quantitation for serum NfL and p-Tau_{181} were 0.174 and 1 pg/mL, respectively. Measurements were carried out using an HD-X analyser (Quanterix, Boston, Massachusetts, USA) at the same study site on consecutive days, using the same batch of reagents, and the operators were blinded to all clinical information. Quality control samples had a mean intra-assay and inter-assay coefficient of variation of <8% and <20%, respectively.

MRI acquisition, processing and analysis
Brain images were collected using 1.5 T (Siemens Symphony and Avanto, Erlangen, Germany) or 3 T scanner (Siemens Skyra, Erlangen, Germany) equipped with a circularly polarised transmit-receive coil to obtain three-dimensional magnetisation-prepared rapid gradient echo T1-weighted scans. At 1.5 T, sequences were acquired with the following parameters: repetition time 2100–2050 ms, echo time 2.95–2.56 ms, inversion time 1100 ms, slice thickness 1 mm, voxel size 1×1×1 mm, in-plane field of view 236 mm, flip angle =15°. At 3 T, sequences were acquired with the following parameters: repetition time 2000 ms, echo time 2.92 ms, inversion time 850 ms, slice thickness 1.1 mm, voxel size 1.1×1.1×1.1 mm, field of view 282 mm, flip angle 8°.

T1 scans were visually inspected and excluded from subsequent analyses if excessive motion blurring or artefacts were present. Then, images were processed and analysed with the fully automated surface-based morphometry pipeline in the Computational Anatomy Toolbox (CAT12.6) (http://www.neuro.uni-jena.de/cat/) for Statistical Parametric Mapping (SPM12 V.7771) (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), running on MATLAB 9.2 (The MathWorks, Natick, Massachusetts, USA). Cortical meshes were resampled to the Human Connectome Project mesh and smoothed with a 15 mm filter.

Smoothed cortical thickness meshes were included in a multiple regression model, in which serum NfL and serum p-Tau_{181} values represented the independent variables. Age, sex, clinical phenotype and MRI scanner type were considered as confounding factors. The statistical threshold was set at 0.05 and corrected for multiple comparisons using false discovery rate (FDR) at whole-brain level.

Transcranial magnetic stimulation
A TMS figure-of-eight coil (each loop diameter 70 mm—D70\textsuperscript{2} coil) connected to a monophasic Magstim Bistim\textsuperscript{2} system
(Magstim, Oxford, UK) was employed for all TMS paradigms, as previously reported.27 Electromyographic (EMG) recordings were performed from the first dorsal interosseus muscle using 9 mm diameter, Ag-AgCl surface-cup electrodes. The active electrode was placed over the muscle belly and the reference electrode over the metacarpophalangeal joint of the index finger. Responses were amplified and filtered at 20 Hz and 2 kHz with a sampling rate of 5 kHz.

Resting motor threshold (RMT) was determined on the left motor cortex as the minimum intensity of the stimulator required to elicit motor evoked potentials (MEPs) with a 50 μV amplitude in 50% of 10 consecutive trails, recorded during full muscle relaxation.

SICI-ICF, LICI and SAI were studied using a paired-pulse technique, employing a conditioning-test design. For all paradigms, the test stimulus (TS) was adjusted to evoke a MEP of approximately 1 mV amplitude.

For SICI and ICF, the conditioning stimulus (CS) was adjusted at 70% of the RMT, employing multiple interstimulus intervals (ISIs), including 1, 2, 3 ms for SICI and 7, 10, 15 ms for ICF.28 LICI was investigated by implementing two suprathereshold stimuli, with the CS adjusted at 130% of the RMT, employing ISIs of 50, 100 and 150 ms.30 SAI was evaluated employing a CS of single pulses (200 μs) of electrical stimulation delivered to the right median nerve at the wrist, using a bipolar electrode with the cathode positioned proximally, at an intensity sufficient to evoke a visible twitch of the thenar muscles.31

Different ISIs were implemented (0, +4 ms), which were fixed relative to the N20 component latency of the somatosensory evoked potential of the median nerve.

For each ISI and for each protocol, 10 different paired CS-TS stimuli and 14 control TS stimuli were delivered in all participants in a pseudo-randomised sequence, with an intertrial interval of 5 s (±10%).

The conditioned MEP amplitude, evoked after delivering a paired CS-TS stimulus, was expressed as percentage of the average control MEP amplitude. Average values for SICI (1, 2, 3 ms ISI), ICF (7, 10, 15 ms ISI), LICI (50, 100, 150 ms ISI) and SAI (0, +4 ms ISI) were used for analysis.

Stimulation protocols were conducted in a randomised order. Audio-visual feedback was provided to ensure muscle relaxation during the entire experiment and trials were discarded if EMG activity exceeded 100 μV in the 250 ms prior to TMS stimulus delivery. Less than 5% of trials were discarded for each protocol. All of the participants were capable of following instructions and reaching complete muscle relaxation; if, however the data were corrupted by patient movement, the protocol was restarted and the initial recording was rejected.

Statistical analysis

Linear regression and stepwise multiple regression analysis (including all variables with a p<0.100 at univariate analysis) were used to characterise the relationship between serum biomarkers and demographic characteristics (age, age at onset, sex and mutation status).

Differences in clinical variables and biomarker concentrations were assessed with one-way analysis of covariance (ANCOVA), corrected for age, sex and/or mutation status, with Bonferroni multiple comparisons correction. Pearson’s correlations were used to assess associations between serum biomarkers, age-corrected and education-corrected clinical variables and TMS measures. There were linear relationships between variables, as assessed by scatterplots and partial regression plots. There was univariate normality, as assessed using histograms and normal Q-Q plots.

Receiver operating characteristics (ROC) curve analyses were used to determine the ability of serum NfL and p-Tau181 to differentiate between diagnostic groups. The area under the curve (AUC) including 95% CI values are reported, with cut-off points set to achieve highest levels of sensitivity and specificity (Youden’s index).

Survival was calculated as time from symptom onset to time of death from any cause (outcome=0) or censoring date (outcome=1). Survival analysis was carried out by the Kaplan-Meier method with log rank post hoc testing and by means of univariate and multivariate stepwise Cox proportional-hazard regression analysis; HRs are provided with their respective 95% CIs.

A two-sided p value <0.05 was considered significant and corrected for multiple comparisons using FDR when appropriate. Statistical analyses were performed using SPSS (V.24; SPSS, IBM).

Data availability

All study data, including raw and analysed data, and materials will be available from the corresponding author, BB, on reasonable request.

RESULTS

Participant characteristics

Baseline demographics, clinical variables and fluid biomarker levels are reported in table 1.

In the FTLD group, serum NfL concentrations did not correlate with age (β = −0.07, p=0.272), age at onset (β=−0.03, p=0.614) or sex (β=−0.08, p=0.193), but correlated with the presence of a pathogenic mutation in both the linear regression (β=−0.48, p<0.001) and in the stepwise multiple regression model (β=−0.48, p<0.001). Serum NfL concentrations were higher in patients with a pathogenic mutation (mean±SEM, GRN mutations n=30, 86.2±5.0 pg/mL; MAPT mutations n=3, 43.0±15.9 pg/mL) compared with patients without a pathogenic mutation (no mutation/unknown n=258, 36.0±17.9 pg/mL), p<0.001.

Serum p-Tau181 concentrations also did not correlate with age (β =0.07, p=0.209), age at onset (β=0.08, p=0.159) or sex (β=−0.08, p=0.200), but correlated inversely with the presence of a pathogenic mutation in both the linear regression (β=−0.16, p=0.006) and in the stepwise multiple regression model (β=−0.13, p<0.021). Serum p-Tau181 concentrations were lower in patients with a pathogenic mutation (GRN mutations, 0.6±1.1 pg/mL; MAPT mutations, 2.3±3.6 pg/mL) compared with patients without a pathogenic mutation (no mutation/unknown, 3.9±0.4 pg/mL), p=0.006.

Serum NfL and serum p-Tau181 concentrations in FTLD subgroups

Serum NfL concentrations were significantly increased in most FTLD subgroups (age-corrected and sex-corrected ANCOVA, F(6,408) =11.97, p<0.001, η²=0.15). In Bonferroni-corrected post hoc tests, we observed a significant increase in serum NfL levels in bvFTD, avPPA and CBS, and in patients with AD compared with HC. Patients with avPPA had significantly higher levels of serum NfL compared with svPPA, CBS, PSP and AD (table 1). These results were confirmed also after correcting for disease duration, excluding HC.
After correcting also for mutation status, considering the unbalanced distribution of pathogenic mutations across FTLD subgroups (table 1) and the increased NIL concentrations in mutation carriers, we observed a significant increase in NIL levels in all the FTD variants (bvFTD, avPPA and svPPA) compared with HC (age-corrected, sex-corrected and mutation-corrected ANCOVA, $F(6,408)=7.00$, $p<0.001$, $\eta^2=0.24$) without significant differences between avPPA and the other subgroups (figure 1A).

Serum $p$-Tau$_{181}$ concentrations were significantly reduced in all FTLD subgroups compared with AD (age-corrected and sex-corrected ANCOVA, $F(6,408)=21.35$, $p<0.001$, $\eta^2=0.24$) (table 1 and figure 1B). No significant differences between FTLD subgroups were found except for higher values in CBS compared with bvFTD (see figure 1B). Serum $p$-Tau$_{181}$ was also significantly increased in patients with AD compared with HC (figure 1B). No significant differences between FTLD subgroups and HC were found (figure 1B). Comparable results were observed also after adjusting for age, sex and mutation status (ANCOVA, $F(6,408)=20.21$, $p<0.001$, $\eta^2=0.23$).

Considering the overall sample of 417 subjects, 97 (23.3%) showed $p$-Tau$_{181}$ levels below the lower limit of quantitation (60 bvFTD, 6 svPPA, 24 avPPA, 3 CBS, 1 PSP, 0 AD and 3 HC).

**Diagnostic accuracy of serum NIL and serum p-Tau$_{181}$**

To differentiate patients with FTLD from HC, we applied a ROC curve analysis on serum NIL concentrations, observing an AUC of 0.862 ($p<0.001$, 95% CI 0.818 to 0.906); the serum NIL cut-off of 22.5 pg/mL differentiated FTLD from HC with a sensitivity of 71.5% and a specificity of 92.1% (figure 2A). In patients with a mild disease stage (FTLD-modified CDR $\leq$ 6.5), a serum NIL cut-off of 18.0 pg/mL differentiated mild FTLD from HC with a sensitivity of 74.8% and specificity of 74.2%, with an AUC of 0.808 ($p<0.001$, 95% CI 0.747 to 0.869) (figure 2B).

To differentiate FTLD from patients with AD, we applied a ROC curve analysis on serum phospho-Tau$_{181}$ concentrations, observing an AUC of 0.930 ($p<0.001$, 95% CI 0.903 to 0.956); a serum p-Tau$_{181}$ cut-off of 5.88 pg/mL differentiated FTLD from AD with a sensitivity of 81.4% and a specificity of 93.3% (figure 2C).

In patients with a mild disease stage (FTLD with an FTLD-modified CDR $\leq$ 6.5 and AD with a MMSE $\geq$20), the serum p-Tau$_{181}$ cut-off of 6.66 pg/mL differentiated FTLD from AD with a sensitivity of 89.3% and specificity of 82.0%, with an AUC of 0.909 ($p<0.001$, 95% CI 0.863 to 0.954) (figure 2D).

**Serum NIL and serum p-Tau$_{181}$ associations with cognitive function and disease severity in FTLD**

**Cognitive and behavioural assessment.** Serum NIL concentrations showed significant associations with baseline BADL ($r=0.23$, $p<0.001$), IADL ($r=0.23$, $p<0.001$) and FTLD-modified CDR sum of boxes ($r=0.28$, $p<0.001$); the higher the serum NIL levels, the greater impairment in functional activities and disease severity. Significant correlations were observed between serum NIL concentrations and MMSE scores ($r=-0.30$, $p<0.001$), phonemic ($r=-0.24$, $p=0.001$) and semantic fluencies ($r=-0.24$, $p=0.001$), clock drawing ($r=-0.24$, $p=0.001$),

---

**Table 1** Demographic and clinical characteristics of patients with FTLD and controls

<table>
<thead>
<tr>
<th>Variable</th>
<th>bvFTD</th>
<th>avPPA</th>
<th>svPPA</th>
<th>CBS</th>
<th>PSP</th>
<th>AD</th>
<th>HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>134</td>
<td>48</td>
<td>27</td>
<td>51</td>
<td>31</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>Age, years</td>
<td>64.5±8.0</td>
<td>67.7±8.8</td>
<td>64.0±8.2</td>
<td>65.8±7.6</td>
<td>72.9±7.4</td>
<td>75.5±8.1</td>
<td>65.4±12.1</td>
</tr>
<tr>
<td>Sex, female %</td>
<td>58.2</td>
<td>43.8</td>
<td>59.3</td>
<td>52.9</td>
<td>51.6</td>
<td>31.7</td>
<td>20.6</td>
</tr>
<tr>
<td>Age at onset, years</td>
<td>61.5±7.8</td>
<td>64.9±8.6</td>
<td>60.5±8.0</td>
<td>63.2±7.5</td>
<td>68.8±7.3</td>
<td>74.0±8.3</td>
<td>–</td>
</tr>
<tr>
<td>Disease duration, years</td>
<td>2.9±2.8</td>
<td>2.8±2.6</td>
<td>3.3±2.2</td>
<td>2.5±1.8</td>
<td>4.1±2.8</td>
<td>1.5±1.7</td>
<td>–</td>
</tr>
<tr>
<td>Sex, female %</td>
<td>20 (14.9)</td>
<td>12 (25.0)</td>
<td>0 (0.0)</td>
<td>1 (2.0)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>–</td>
</tr>
<tr>
<td>FTLD-CDR</td>
<td>7.9±5.3</td>
<td>6.2±5.1</td>
<td>5.7±3.6</td>
<td>4.3±5.3</td>
<td>4.2±5.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>MMSE</td>
<td>21.0±7.65</td>
<td>18.7±9.9</td>
<td>21.7±8.3</td>
<td>24.7±4.6</td>
<td>24.1±7.0</td>
<td>19.5±4.7</td>
<td>–</td>
</tr>
<tr>
<td>Serum NfL (pg/mL)</td>
<td>mean±SEM 43.0±2.4</td>
<td>54.6±3.9</td>
<td>33.3±5.2</td>
<td>36.3±3.8</td>
<td>30.4±4.9</td>
<td>32.7±3.6</td>
<td>14.2±3.5</td>
</tr>
<tr>
<td>Serum p-Tau$_{181}$ (pg/mL)</td>
<td>mean±SEM 2.5±0.7</td>
<td>3.3±1.1</td>
<td>3.8±1.5</td>
<td>7.1±1.1</td>
<td>3.9±1.4</td>
<td>16.4±1.1</td>
<td>5.4±1.0</td>
</tr>
</tbody>
</table>
| Results are expressed as mean±SD, unless otherwise specified. Monogenic disease: all GRN mutations, but three MAPT mutations (two bvFTD and one CBS).

AD, Alzheimer’s disease; avPPA, agrammatic variant of primary progressive aphasia; bvFTD, behavioural variant frontotemporal dementia; CBS, corticobasal syndrome; FTLD, frontotemporal lobar degeneration; FTLD-CDR, frontotemporal lobar degeneration-modified clinical dementia rating scale; HC, healthy controls; MMSE, Mini-Mental State Examination; NIL, neurofilament light chain; PSP, progressive supranuclear palsy; svPPA, semantic variant of primary progressive aphasia.
short story ($r = -0.25$, $p=0.002$), trail-making part B ($r = -0.22$, $p=0.011$), digit symbol ($r = -0.16$, $p=0.027$) and token test ($r = -0.17$, $p=0.035$), with higher levels of serum NfL correlating with poorer scores. No significant correlations were observed for the Rey figure copy ($r = -0.10$, $p=0.155$) and recall ($r = -0.09$, $p=0.222$), and trail-making test part A ($r = 0.11$, $p=0.117$). Neuropsychiatric and behavioural disturbances, evaluated with the FMI, significantly correlated with serum NfL levels ($r = 0.18$, $p=0.007$). All tests were age-corrected and education-corrected; FDR-adjusted $p$ values for multiple comparisons are reported for each test.

No significant correlations were observed between serum p-Tau$_{181}$ concentration and FTLD-CDR sum of boxes score or other neuropsychological, behavioural or functional measures.

Brain imaging. Brain imaging analysis was performed in 132 patients (57 bvFTD, 28 aPPA, 16 sPPA, 19 CBS and 12 PSP). As reported in figure 3, serum NfL concentration was inversely correlated with cortical thinning of the prefrontal (dorsolateral, mesial and orbitofrontal), temporal and parietal regions. The highest correlation was found for the dorsolateral prefrontal cortex, mainly involving the left side ($p<0.05$ whole-brain FDR-corrected, cluster threshold = 200).

There was no statistically significant association between serum p-Tau$_{181}$ and cortical thickness in patients with FTLD.

TMS measures. TMS analysis was performed in 113 patients (47 bvFTD, 15 aPPA, 14 sPPA, 9 CBS 4 PSP, 12 AD and 12 HC). TMS measures were performed to evaluate average SICI, ICF, LICI and SAI. In the FTLD group ($n=89$), serum NfL levels were significantly associated with SICI ($r = 0.46$, $p<0.001$) and LICI ($r = 0.55$, $p<0.001$) (figure 4A and B), but not with ICF or SAI. No associations were observed between serum p-Tau$_{181}$ and TMS measures.

Interestingly, in the AD group ($n=12$), we observed a significant association between serum p-Tau$_{181}$ and average SAI ($r = 0.72$, $p=0.048$) (figure 4C). We did not observe any significant associations between serum NfL and TMS measures.

Reported $p$-values are FDR-adjusted for multiple comparisons.

Serum NfL and serum p-Tau$_{181}$ associations with prognosis in FTLD
Serum NfL concentration significantly predicted the survival rate in patients with FTLD.

The univariate and stepwise multivariate Cox regression analysis showed a significant association between survival and serum NfL levels (HR 1.01, 95% CI 1.00 to 1.02, $p=0.005$), but not with p-Tau$_{181}$, age, age at onset or mutation status (figure 5A). Patients with high serum NfL levels (upper than median values) had significantly shorter survival than those with low serum NfL levels (lower than median values) at the Kaplan-Meier survival curves ($p=0.034$) (figure 5B).

DISCUSSION
In this work, we confirmed and extended previous literature claiming for a different usefulness of serum NfL and serum p-Tau$_{181}$ measurements in clinical practice, depending on specific clinical questions. Serum NfL concentrations showed high accuracy in identifying FTLD from cognitively unimpaired elderly, as well as in assessing FTLD severity and prognosis, while serum p-Tau$_{181}$ concentrations showed high accuracy in discriminating FTLD from AD. Importantly, in this study we also further demonstrate high accuracy of these biomarkers even in the earliest disease stages.
Figure 4  Significant associations between serum biomarkers and neurophysiological measures. Association between serum NfL concentrations (pg/mL) and (A) average SICI (ISI 1, 2, 3 ms ISI), (B) average LICI (ISI 50, 100, 150 ms ISI); association between serum p-Tau181 concentrations (pg/mL) and (C) average SAI (0, +4 ms ISI). ISI, interstimulus interval; LICI, long-interval intracortical inhibition; NfL, neurofilament light; p-Tau181, phospho-Tau181; SICI, short-interval intracortical inhibition; SAI, short latency afferent inhibition.

The non-invasiveness and reliability of serum NfL and p-Tau181 measurements make these markers extremely useful in clinical practice for the diagnosis of FTLD, even in the early disease stages, compared with CSF biomarkers or more expensive brain imaging modalities.

Serum NfL concentrations, as already demonstrated in other neurodegenerative disorders, were associated with measures of disease severity, and are helpful in assessing disease stage. In fact, higher serum NfL levels were significantly associated with more pronounced cognitive impairment and behavioural disturbances. We also observed an association with cortical thickness at brain imaging analysis. In particular, NfL concentrations were inversely correlated with cortical thickness values mainly in frontal, temporal and parietal regions, supporting the view that NfL is a neurodegeneration marker strongly related to FTLD.34 These findings were also consistent with previous studies in FTLD that reported a correlation between brain structure and NfL concentrations, with a predominant involvement of the left frontotemporal area.8 35 36 To further corroborate the role of serum NfL as a marker of disease severity, we evaluated the association between serum NfL concentrations and indirect measures of GABAergic neurotransmission, which have been demonstrated to be impaired in FTLD.37–39 We observed that the higher the serum NfL levels, the greater was the impairment in SICI and LICI, which are considered to reflect short-lasting postsynaptic...
inhibition mediated through the GABA<sub>A</sub> and GABA<sub>B</sub> receptors at the level of local interneurons, respectively.\textsuperscript{22, 24}

Altogether, these findings strongly support the notion that serum NfL concentrations may be useful to stage disease severity, in a disorder where there is urgent need to find diagnostic and prognostic markers, in light of the near onset of new pharmacological clinical trials.

Compared with AD, FTLD is clinically heterogeneous, with patients presenting a combination of behavioural disturbances, impairment of executive functions or language deficits. Available standardised neuropsychological and clinical assessments may not be ideal in detecting the effects of future treatments, particularly in the early disease stages and across different FTLD subtypes. A non-invasive and easy to perform peripheral biomarker may represent a practical and valuable choice to assess disease severity, to monitor outcomes and to categorise patients into disease subgroups.

Furthermore, these findings extend previous literature data, suggesting that serum NfL assessment may be considered as an additional tool to detect neurodegenerative disorders in the earliest disease stages, even though it is not sufficiently helpful in discriminating between FTLD clinical phenotypes. More interestingly, in the next future, serum NfL concentrations may be implemented to regularly screen subjects with subtle, or even subjective, cognitive disturbances, to identify the first stages of an ongoing neurodegenerative process.

Most importantly, this study has demonstrated that serum NfL concentrations are able to predict survival rates. Indeed, several studies have now shown the prognostic value of NfL in patients with FTLD; however, concentrations were evaluated in CSF, or in small group of patients or in patients with monogenic disease.\textsuperscript{40, 41} These confirmatory results observed using serum NfL concentrations in a large cohort of subjects with FTLD are key to clearly prove that patients with higher NfL levels show decreased survival. These findings further prove that NfL, a major component of neuronal cytoskeleton involved in axonal and dendritic growth, signalling and transport, reflect the ongoing neuronal loss also in FTLD.\textsuperscript{40}

Conversely, serum p-Tau<sub>181</sub> levels, besides being very accurate in discriminating AD from FTLD, were not helpful in monitoring disease severity or predicting prognosis in FTLD. Indeed, according to previous data, blood p-Tau<sub>181</sub> may detect mixed 3R/4R neuropathology, that is, AD, but not other tauopathies, such as 4R tauopathy (ie, Pick’s disease) or 3R tauopathy (ie, PSP or CBS).\textsuperscript{13, 17} Accordingly, a recent study has clearly demonstrated that plasma p-Tau<sub>181</sub> concentrations were increased by 3.5-fold in AD and differentiated AD from both clinically diagnosed (AUC of 0.894) and autopsy-confirmed FTLD (AUC of 0.878).\textsuperscript{15} However, all these data were generated using plasma samples. We have shown that even serum p-Tau<sub>181</sub> was highly increased in AD but not FTLD. For this reason, serum p-Tau<sub>181</sub> was not able to identify FTLD subtypes. The modest increase in p-Tau<sub>181</sub> concentrations observed in patients with CBS could be secondary to a concomitant AD neuropathology, which has been frequently observed in these patients.\textsuperscript{42} Accordingly, in patients carrying a MAPT P301L mutation, p-Tau<sub>181</sub> concentrations were not significantly higher than in other FTLD subtypes (data not shown), as they have a pure 4R tau pathology. It is however noteworthy that patients carrying GRN mutations, and consequently with FTLD-TDP43 pathology, showed decreased serum p-Tau<sub>181</sub> compared with patients without a pathogenetic mutation. The related pathological mechanism needs to be further explored. Finally, SAI, a TMS measure of cholinergic dysfunction widely associated with AD,\textsuperscript{43} correlated harmoniously with serum p-Tau<sub>181</sub> levels, further confirming the reliability of peripheral p-Tau<sub>181</sub> in detecting AD, as previously reported.\textsuperscript{47}

Major strengths of our study are the large series of patients with FTLD and the comprehensive approach in correlating clinical, imaging and neurophysiological data with fluid biomarkers, carried out at the same study site to minimise variability. A weakness of the study is the lack of autopsy confirmation, which prevented correlations between biomarkers and FTLD-related proteinopathies. Second, as compared with previous studies,\textsuperscript{13, 44} the relatively small sample size of patients with PSP prevented us to detect a significant increase in serum NfL. Finally, longitudinal serum NfL measurements were not available, and we were not able to draw conclusions on possible changes throughout disease progression.

In conclusion, our results show the usefulness of both peripheral NfL and p-Tau<sub>181</sub> assessment, with different and specific purposes in clinical practice. Assessing both blood-based biomarkers may provide a comprehensive view of FTLD, aiding in the differential diagnosis, in staging disease severity and in defining survival probability.

Author affiliations
\textsuperscript{1}Neurology Unit, Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
\textsuperscript{2}Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Mölndal, Sweden
\textsuperscript{3}Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Mölndal, Sweden
\textsuperscript{4}Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, London, United Kingdom
\textsuperscript{5}NIHR Biomedical Research Centre for Mental Health & Biomedical Research Unit for Dementia, South London and Maudsley NHS Foundation Trust, London, United Kingdom
\textsuperscript{6}Neurophysiology Unit, ASST Spedali Civili di Brescia, Brescia, Italy
\textsuperscript{7}Vascular Neurology Unit, ASST Spedali Civili di Brescia, Brescia, Italy
\textsuperscript{8}Molecular Markers Laboratory, IRCCS Istituto San Giovanni di Dio Fatebenefratelli, Brescia, Italy
\textsuperscript{9}Department of Neurology, University Medical Centre, Ljubljana, Slovenia
\textsuperscript{10}Clinical Neurochemistry Laboratory, Sahlgrenska University Hospital, Mölndal, Sweden
\textsuperscript{11}MAC Memory Clinic and Molecular Markers Laboratory, IRCCS Istituto San Giovanni di Dio Fatebenefratelli, Brescia, Italy
\textsuperscript{12}Neuroangiology Unit, University of Brescia, Brescia, Italy
\textsuperscript{13}Department of Neuroscience, University College London, London, United Kingdom
\textsuperscript{14}Department of Neurodegenerative Disease, Institute of Neurology, University College London, London, United Kingdom

Twitter Roberto Gasparotti @bobgas@giumbominx

Acknowledgements The authors would like to thank A. Alberici for patients’ recruitment, and E. Bonomi and C. Brattini for excellent practical work.

Contributors Conception and design of the study: AB and BB. Acquisition and analysis of data: AB, TKK, NA, SG, EP, LB, RG, JLR, AE, JM, GB, SF, MG, RG, HZ, KB and BB. Drafting the manuscript and figures: AB, MG, SG, EP and BB. Revising the manuscript for intellectual content: AB, TKK, NA, SG, EP, LB, RG, JLR, AE, J5, GB, SF, MG, RG, HZ, KB and BB.

Funding This study was partially funded by the Italian Ministry of Health (Ricerca Corrente). HZ is a Wallenberg Scholar supported by grants from the Swedish Research Council (#2018-02532), the European Research Council (8681712), the Swedish State Support for Clinical Research (#ALFGGB-720931), the Alzheimer Drug Discovery Foundation (ADDF), USA (#201809-2016862) and the UK Dementia Research Institute at UCL. KB is supported by the Swedish Research Council (#2017-00915), the Alzheimer Drug Discovery Foundation (ADDF), USA (#RDAPB-201809-2016615), the Swedish Alzheimer Foundation (#AF-742881), the Swedish Dementia Center, Sweden (#F2017-02434), the Swedish state under the agreement between the Swedish government and the County Councils, the ALF-agreement (#ALFGGB-715986) and European Union Joint Program for Neurodegenerative Disorders (JPND2019-466-236).

Competing interests HZ has served at scientific advisory boards for Denali, Roche Diagnostics, Wave, Sarnemed and CoqRx, has given lectures in symposia sponsored by Fujirebio, Alzecure and Biogen and is a co-founder of Brain Biomarker Solutions...
in Gothenburg AB, a GU Ventures-based platform company at the University of Gothenburg, all unrelated to the work presented in this paper. KB has served as a consultant or at advisory boards for Abcam, Axon, Biogen, Lilly, MagQus, Novartis and Roche Diagnostics, and is a co-founder of Brain Biomarker Solutions in Gothenburg AB, a GU Venture-based platform company at the University of Gothenburg, all unrelated to the work presented in this paper.

Patient consent for publication Not required.

Ethics approval Full written informed consent was obtained from all subjects according to the Declaration of Helsinki. The Brescia Ethics Committee approved the study protocol (NP 1965).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All study data, including raw and analysed data, and materials will be available from the corresponding author, BB, on reasonable request.

ORCID iDs
Alberto Benussi http://orcid.org/0000-0002-8703-6940
Barbara Borroni http://orcid.org/0000-0001-9340-9814

REFERENCES