Original research

Neurofilament light chain predicts future dementia risk in cerebral small vessel disease

Marco Egle , Laurence Loubiere , Aleksandra Maceski , Jens Kuhle , Nils Peters and Hugh S Markus

ABSTRACT

Objectives Serum neurofilament light chain (NfL) has been proposed as prognostic markers in neurogenenerative disease. A cross-sectional study in cerebral small vessel disease (SVD) reported an association with cognition and disability. If NfL is to be used to predict outcome, studies are required to demonstrate baseline NfL predicts future dementia risk. Furthermore, if it is to be used as a surrogate marker in clinical trials, change in NfL over time periods typical of a clinical trial must be linked to clinical progression. In a longitudinal study of patients with lacunar stroke and confluent white matter hyperintensities, we determined whether both baseline, and change, in NfL levels were linked to changes in MRI markers, cognitive decline and dementia risk.

Methods Patients underwent MRI, cognitive testing and blood taking at baseline and annually for 3 years. Clinical and cognitive follow-up continued for 5 years.

Results NfL data were available for 113 subjects for baseline analysis, and 90 patients for the longitudinal analysis. Baseline NfL predicted cognitive decline (global cognition β=−0.335, SE=0.094, p=0.001) and risk of converting to dementia (HR=1.676 (95% CI 1.183 to 2.373), p=0.004). In contrast to imaging, there was no change in NfL values over the follow-up period.

Conclusions Baseline NfL predicts changes in MRI markers, cognitive decline and dementia rate over a 5 years follow-up period in SVD, suggesting NfL may be a useful prognostic marker. However, change in NfL values was not detected, and therefore NfL may not be a useful surrogate marker in clinical trials in SVD.

INTRODUCTION

Cerebral small vessel disease (SVD) is a highly prevalent condition, which causes lacunar stroke, vascular cognitive impairment and vascular dementia. Few treatments have been shown to delay disease progression. One challenge in testing therapies is that, while on a population basis MRI features of SVD predict dementia, predicting which individual patients will progress is difficult. This has led to calls for better markers to predict progression to dementia. Furthermore such markers may be useful as surrogate markers in disease studies. Due to the variable progression rate large sample sizes are required for clinical trials, and markers, which are sensitive to change would allow treatments to be tested in smaller phase 2 trials. MRI markers including white matter hyperintensities (WMH) and diffusion tensor imaging (DTI) have been suggested as surrogate markers.

Neurofilament light chain (NfL) is a blood marker sensitive to neuroaxonal damage and increased levels occur in neurodegenerative diseases. In a cross-sectional study in SVD, NfL levels were associated with cognition and disability. However, if NfL is to be used to predict outcome, studies are required to demonstrate baseline NfL predicts future dementia risk. Furthermore if it is to be used as a surrogate marker in clinical trials, a change in NfL level within a 2–3 year period typical of a clinical trial must be shown to be associated with clinical progression. In a longitudinal prospective study of patients with moderate-to-severe SVD, we determined whether both baseline, and change, in NfL levels predicted future cognitive decline and dementia risk.

METHODS

Subjects

Both baseline and follow-up data were used from the St George’s Cognition and Neuroimaging in Stroke (SCANS) prospective study. A total of 121 patients with symptomatic SVD, defined as a clinical lacunar stroke syndrome with MRI evidence of an anatomically corresponding lacunar infarct, and with confluent regions of WMH graded ≥2 on the modified Fazekas scale were enrolled from three stroke services in South London (St George’s Hospital, King’s College Hospital and St Thomas’ Hospital) at least 3 months post stroke. Exclusion criteria were any stroke not caused by SVD, other central nervous system diseases, major psychiatric disorders and any cause of white matter (WM) disease other than SVD. Written informed consent was obtained from all patients participating in the study. The study was granted ethical approval by Wandsworth Research Ethic Committee.

Subjects were followed up annually for 5 years with six incidences of strokes being recorded (four lacunar strokes, two intracerebral haemorrhages). MRI was performed at baseline and after 1, 2 and 3 years. Information on dementia incidence was acquired annually. All blood samples had been collected annually at the time of MRI imaging. Of the 121 subjects recruited, blood was available for 113, and in 90 patients blood samples were available from at least at two time points. Due to a small sample size, observations coming from follow-up time point 4 and 5 were removed.

Serum NfL

All assays were performed at the same time in 2019–2020. The analysis was the same for all
samples using the same single-molecule array instrument (Simoa HD-1, Quanterix, Lexington, Massachusetts, USA) in Basel. We used the capture monoclonal antibody (mAB) 47:3 and the biotinylated detector mAB 2:1 (UmanDiagnostics, Umeå, Sweden),9 transferred onto the Simoa platform. Bovine lyophilised NfL was acquired from UmanDiagnostics. The range of the calibrators was between 0 and 2000 pg/mL. All samples had signals above the assay’s analytical sensitivity. The NfL analysis was blind to the dementia outcome.

MRI acquisition
Images were obtained using a 1.5 T General Electric Signa HDxt MRI system6 employing the same image acquisition protocol at every time point. The following scan sequences referring to whole head coverage were acquired:

1. Axial Fluid Attenuated Inversion Recovery (FLAIR): repetition time (TR)=9000 ms, echo to time (TE)=130 ms, inversion time=2200 ms, 28 slices of 5 mm thickness without slice gap, field of view=240×240 mm², matrix=256×192
2. Coronal T1-weighted spoiled gradient recalled echo: TR=11.5 ms, TE=5 ms, 176 slices, field of view=240×240 mm², matrix=256×192, flip angle=18° and characterised by 1.1 mm³ isotropic voxels
3. Axial single shot spin echo planar diffusion-weighted imaging: TR=15600 ms, TE=93.4 ms, 35 slices marked by isotropic voxels of 2.5 mm³, field of view=240×240 mm², 8 non-diffusion-weighted-images (b=0 smm −2). This was succeeded by diffusion-weighted volumes with diffusion gradients applied (b=1000 smm −2) in 25 non-collinear directions and the negative of these.

Conventional MRI markers
Normalised brain volume (NVB) and WMH lesion volume were computed as previously described in detail.8 10 NVB was computed from native T1 images as an estimate of size of the brain relative to the skull size with SIENAX as part of the FMRIB software library. To obtain the WMH measure, warped T1-weighted and FLAIR images were used to create population-specific tissue probability maps (TPMs). TPMs were employed to segment native images creating tissue classes of WM, WMH, grey matter (GM) and cerebrospinal fluid (CSF). WMH lesion volume was computed by binarising the segmentations at a manually determined threshold. WMH was calculated by taking the ratio of WMH volume to the total cerebral volume, which is composed of the sum of GM, WM and WMH at a tissue probability threshold ≥0.2. Lacune count (lacune), defined as 3–15 mm in diameter, and cerebral microbleed count (CMB), defined as focal spots up to 10 mm in diameter, were detected by an independent trained rater in line with agreed neuroimaging standards.11 12

DTI analysis
Two methods to assess WM ultrastructure on DTI were used:

All WM histogram analysis measure
The analysis has been described previously.8 In brief, DTI scans were preprocessed and eddy-corrected using FMRIB’s Diffusion Toolbox in FSL software. Susceptibility distortions were unwarped by normalising the images to the T1 images in the phase-encoding direction. DTIFIT was then applied to create mean diffusion maps. FLAIR to T1W and T1W to b0 registrations were performed and the affine transformation matrices were concatenated to produce a FLAIR-to-DTI transformation.13 The TPMs were registered into DTI space employing these transformations. A hard segmentation method was applied to generate maps of tissue classes. Histogram analysis was conducted on the MD map in all WM regions. The summary histogram measures were derived from normalised histograms with 1000 bins. The primary outcome measure was mean diffusivity peak height (MDPH) in the WM, as this has previously been shown to be the most sensitive parameter to change in the SCANs study.8

Peak width of skeletonised mean diffusivity (PSMD)
The PSMD is a fully automatically computed imaging marker being publicly available (http://psmd-marker.com) and has been fully described previously.14 The computation is divided into four main modules: (1) DTI sequence on MRI, (2) WM tract skeletonisation using tract-based statistics, (3) excluding CSF-prone regions employing a custom mask and (4) histogram values of mean diffusion on the skeleton. The PSMD measure refers to the difference between the 5th and 95th percentiles of the histogram distribution.

Cognition and dementia diagnosis
Standardised tasks sensitive in capturing the impaired cognitive function in SVD were carried out annually. Full descriptions of the tasks have been published previously.15 Cognitive index scores were created by grouping the standardised measures into subdomain scores15: global cognition score (Global), executive function (EF) and processing speed (PS).
Dementia was diagnosed using the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-5) definition of major neurocognitive disorder, as previously described, and was made under one of the following criteria:

1. Dementia was diagnosed in a memory clinic or in a similar clinical service.
2. A neurologist and a clinical neuropsychologist who were both blind to all imaging and risk factor information reviewed medical records as well as cognitive data and together agreed that a patient met the DSM-5 criteria.
3. Significant cognitive impairment and reduced capabilities in daily living indicated by a consistently reduced score of <24 in the mini-mental state examination (MMSE) and of ≤7 on the instrumental activities of daily living.

Disability was assessed using the modified Rankin Scale, a 7-point scale assessing disability during stroke.

The statistical analysis was carried out using R (version 3.6.3). NfL, WMH lesion load, lacune and CMB count were log 10 transformed due to the skewness of their data distribution. One outlier observation at baseline was excluded as being nearly 300 times higher than any of the patient’s follow-up NfL values. This patient did not show any clinical aspects that could explain such a high value, e.g., having a high blood pressure. The exclusion was justified due to the skewness of their distribution. The patient was a 72-year-old woman with a history of hypertension and diabetes who had been treated with antihypertensive medication for 6 years and had a clinical stroke 5 years ago. The patient had a high blood pressure of 160/100 mmHg, which was controlled with antihypertensive medication. A high blood pressure was also observed in a small number of patients with diabetes, and having suffered a clinical stroke 5 years ago.

Table 2 Cross-sectional regression between DTI and/or NfL and cognition or disability

<table>
<thead>
<tr>
<th>Regression on cognition with single marker</th>
<th>Regression on cognition with multiple markers</th>
<th>Regression on disability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global</strong></td>
<td>EF</td>
<td>PS</td>
</tr>
<tr>
<td>Serum NfL (pg/mL log 10)</td>
<td>−0.201 (0.072), p=0.006</td>
<td>0.459</td>
</tr>
<tr>
<td>MDPH (mm²/s)</td>
<td>0.290 (0.071), p=0.001</td>
<td>0.503</td>
</tr>
<tr>
<td>PSMD (mm²/s)</td>
<td>−0.277 (0.067), p=0.723e-05</td>
<td>0.506</td>
</tr>
</tbody>
</table>

Significant at p<0.05.
Values show standardised regression coefficients β and SEs for predictor variables in regression models. Adj R², adjusted explained variance; DTI, diffusion tensor imaging; EF, executive function; Global, global cognition; HL R², Hosmer and Lemeshow’s R²; MDPH, mean diffusivity normalised peak height; mRS score, modified rankin scale; NfL, neurofilament light chain; PS, processing speed; PSMD, peak width of skeletonised mean diffusivity.

Figure 1 Higher levels of NfL were associated with a higher SVD marker on imaging. Spearman’s rank correlation shows the cross-sectional associations between MRI markers and NfL. Blue and red quartiles show significant positive and negative correlations. White tiles show no significant association. CMB, cerebral microbleed; Lac, lacune; MDPH, mean diffusivity normalised peak height; NfL, neurofilament light chain; NBV, normalised brain volume; PSMD, peak width of skeletonised mean diffusivity; SVD, small vessel disease; WMH, white matter hyperintensity.

Table 2 Cross-sectional regression between DTI and/or NfL and cognition or disability

<table>
<thead>
<tr>
<th>Regression on cognition with single marker</th>
<th>Regression on cognition with multiple markers</th>
<th>Regression on disability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global</strong></td>
<td>EF</td>
<td>PS</td>
</tr>
<tr>
<td>Serum NfL (pg/mL log 10)</td>
<td>−0.201 (0.072), p=0.006</td>
<td>0.459</td>
</tr>
<tr>
<td>MDPH (mm²/s)</td>
<td>0.290 (0.071), p=0.001</td>
<td>0.503</td>
</tr>
<tr>
<td>PSMD (mm²/s)</td>
<td>−0.277 (0.067), p=0.723e-05</td>
<td>0.506</td>
</tr>
</tbody>
</table>

Significant at p<0.05.
Values show standardised regression coefficients β and SEs for predictor variables in regression models. Adj R², adjusted explained variance; DTI, diffusion tensor imaging; EF, executive function; Global, global cognition; HL R², Hosmer and Lemeshow’s R²; MDPH, mean diffusivity normalised peak height; mRS score, modified rankin scale; NfL, neurofilament light chain; PS, processing speed; PSMD, peak width of skeletonised mean diffusivity.
model the association between NfL and disability was tested while controlling for baseline age.

The effect of time on change in cognition for Global and PS over the five follow-up years was estimated using a linear mixed model (lme4). Fixed effect variation was explained by the study’s duration, and random effect variation accounted for the residual inter-individual differences. The linear trajectory of each patient’s intercept and slope were let to vary with both fixed and random effects. The average fixed effects slopes of time are the average annualised change rate for a given cognition. Using maximum likelihood estimation to fit the model, the missing slope for one participant was estimated by taking information from all other patients included in the model.

To determine the predictive association between baseline NfL or baseline DTI and cognitive decline, a linear regression model was employed. Baseline cognition together with the clinical markers age, gender and NART were added as confounders. Employing a logistic regression, it was furthermore tested whether baseline NfL predicted an increase vs no increase in mRS as a binary outcome over 5 years while controlling for age and its baseline disability score.

| Table 3 | Annualised cognitive and imaging change rate and over 5 years and 3 years, respectively |
|-----------------|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **SVD markers** | **Cognition** (N=90) | **Estimated mean baseline value (CI)** | **Estimated mean annual change (CI)** | **t- value** | **P value** | **Imaging characteristics** (n=90) | **Estimated mean baseline value (CI)** | **Estimated mean annual change rate (CI)** | **t- value** | **P value** |
| **Global** | | | | | | **Serum NfL (pg/mL)** (log 10) | 1.330 (1.279 to 1.381) | 0.0011 (-0.017 to 0.019) | 0.12 | 0.903 |
| | | | | | | **MDPH (mm²/s)** | 0.015 (0.002 to 0.016) | -0.0004 (-0.0005 to -0.0003) | -10.34 | <0.001 |
| | | | | | | **PSMD (mm²/s)** | 3.744e-04 (3.544e-04 to 3.943e-04) | 1.316e-05 (9.467e-06 to 1.686e-05) | 6.981 | <0.001 |
| | | | | | | **WMH (% brain)** (log 10) | 0.438 (0.370 to 0.506) | 0.084 (0.075 to 0.093) | 17.95 | <0.001 |
| | | | | | | **NBV (mL)** | 1294.21 (1277.007 to 1311.410) | -9.017 (-10.764 to -7.270) | -10.12 | <0.001 |
| | Lacune (count) | 2 (0 to 26) | | | | | | | | |
| | CMB (count) | 0 (0 to 41) | | | | | | | | |
| | **Estimated annual rates of change are defined as the mean estimates of the fixed effects from the linear mixed effect models with their 95% CI. Significant at p value <0.05.** |
| | **CMB, cerebral microbleeds count; Global, global cognition; Lacune, lacune count; MDPH, mean diffusivity normalised peak height; NBV, normalised brain volume; NfL, neurofilament light chain; PS, processing speed; PSMD, peak width of skeletonised mean diffusivity; SVD, small vessel disease; WMH, white matter hyperintensity lesion load.** |

| Table 4 | Longitudinal analysis |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Baseline marker prediction** | **Dementia conversion (n=19)** | | | | | | |
| | **β (SE)** | **P value** | **HR (95% CI)** | **AUC** | **β (SE)** | **P value** | **Adj. R** | **β (SE)** | **P value** | **Adj. R** |
| **Serum NfL (pg/mL) (log 10)** | 0.516 (0.178) | **0.004** | 1.676 (1.183 to 2.373) | 0.775 | -0.335 (0.094) | **0.001** | 0.353 | -0.277 (0.114) | **0.017** | 0.050 |
| **MDPH (mm²/s)** | -0.848 (0.265) | **0.001** | 0.428 (0.253 to 0.719) | 0.791 | 0.148 (0.106) | 0.169 | 0.265 | 0.139 (0.123) | 0.262 | -0.006 |
| **PSMD (mm²/s)** | 0.507 (0.155) | **0.001** | 1.661 (1.225 to 2.252) | 0.758 | -0.215 (0.098) | **0.032** | 0.290 | -0.143 (0.116) | 0.224 | -0.003 |
| **Serum NfL (pg/mL) (log10) indep. MDPH (mm²/s)** | 0.369 (0.179) | **0.040** | 1.446 (1.017 to 2.055) | 0.804 | | | | | | |
| **Serum NfL (pg/mL) (log10) indep. PSMD (mm²/s)** | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | **DTI and NfL baseline markers predicting cognitive decline and dementia conversion.** |
| | **Values show standardised regression coefficients: β and SEs for predictor variables in regression models of dementia conversion and decline in cognition. Significant at p value <0.05.** |
| | **Adj R², adjusted explained variance; AUC, area under the curve; DTI, diffusion tensor imaging; Global, global cognition; MDPH, mean diffusivity normalised peak height; NfL, neurofilament light chain; PS, processing speed; PSMD, peak width of skeletonised mean diffusivity.** |
A Cox regression model was used to test whether the baseline NfL or imaging marker predicted dementia conversion. The proportional hazards assumption was met on the basis of Schoenfeld residuals. To assess how well the predictive models’ discriminatory ability classifies patients converting to dementia and those with no dementia at varying threshold, ROC curves were computed and the area under the curve (AUC) was estimated. Differences in clinical characteristics and estimated changes in NfL and change in MRI markers over 3 years were estimated using a linear mixed model and baseline age. The AUC was computed to assess diagnostic discriminatory ability in correctly classifying the binary outcome measures. The association between baseline and change in NfL and change in NBV and in WMH was tested using a linear regression model controlling for the imaging marker’s baseline and the permutation Welch’s t-test for variables with parametric distribution and the Wilcoxon rank sum test for variables with a non-parametric distribution.

The effect of time on change in the imaging marker and in NfL over the three follow-up years was estimated using a linear mixed model as previously described (lme4). Using maximum likelihood estimation to fit the model, the missing DTI slope for the observations after the dementia diagnosis were omitted.

To determine the predictive association between the change in the markers and cognitive decline or dementia conversion, a linear regression or a Cox regression model adjusted by age, gender and premorbid IQ was employed. The proportional hazards assumption was met on the basis of Schoenfeld residuals and the AUC was estimated. Differences in clinical characteristics and estimated changes in NfL and in the imaging markers between patients converting to dementia and those not converting to dementia/ censored over 5 years were tested using the permutation Welch’s t-test.

Finally it was also tested whether baseline and change in NfL predicted change in lacunes, CBM employing a logistic regression model controlling for the imaging marker’s baseline and age. The AUC was computed to assess diagnostic discriminatory ability in correctly classifying the binary outcome measures. The association between baseline and change in NfL and change in NBV and in WMH was tested using a linear regression model while controlling for the imaging marker’s intercept value coming from the linear mixed model and baseline age.

**Data availability**
Cohort data can be shared on reasonable request for scientific purpose by contacting the corresponding author.

**RESULTS**

**Cross-sectional analysis**
Study began in January 2008 and was completed in October 2013 with a consecutive case series design. Recruitment started in December 2007 and finished in August 2010. NfL data were available for 113 subjects. Clinical characteristics, cognitive measures and SVD markers at baseline are shown in table 1. NfL levels had a mean of 36.505 pg/mL, a median of 23.20 pg/mL, an IQR of 16.90 pg/mL and a minimum and maximum of 5.7 pg/mL and 708.9 pg/mL, respectively. The normal range for NfL levels using this assay has a mean of 34.69 (SD 13.09) pg/mL for age 60–70, and a mean of 45.85 (SD 15.31) pg/mL above age 70.

NfL levels inversely associated with global cognitive function, the executive function and processing speed subdomain scores (table 2). Serum NfL also negatively correlated with disability (table 2). Higher NfL levels were positively associated with lacune count, CMB count, WMH and with DTI measures (higher PSDM and lower MDPH), and negatively with NBV (figure 1). We also determined whether association with cognition persisted after controlling for DTI (MDPH) and for the clinical markers; the association remained significant for PS but not for Global or EF (table 2).

**Longitudinal analysis**

**Prediction by baseline NfL**
There was a significant decline in both Global and PS over the 5 years follow-up (table 3). Higher NfL at baseline predicted lower function in Global (β = –0.335 (SE = 0.094), p = 0.001) independently of the clinical markers and baseline cognition (table 4). Higher NfL...
Cerebrovascular disease also predicted lower function in Global ($\beta=-0.303$ (SE=0.095), $p=0.002$), when controlling DTI to the existing regression model. Comparing patients showing more versus less cognitive decline further showed that there were significant baseline differences between the two groups for NART, MMSE score, NfL levels, NBV, lacune count and DTI measures but not for age, education, WMH or CMB (online supplemental table 1). Baseline NfL did not predict increases in mRS score ($\beta=0.407$ (SE=0.864), $p=0.638$).

A total of 107 patients had complete baseline DTI and NfL data of whom 19 converted to dementia over time. Higher baseline NfL predicted the risk of converting to dementia while accounting for the clinical markers (HR=1.676 (95% CI=1.183–2.373), $p=0.004$, table 4). MRI parameters also predicted both cognitive decline and dementia (table 4). The AUC for prediction for dementia was 0.775 for NfL, 0.791 for MDPH and 0.758 for PSMD. To determine whether NfL contributed additional information above that provided by DTI, we additionally included baseline MDPH in the model resulting an increase in AUC of 0.804. Baseline MRI images for one patient aged 70 with a high baseline NfL level, 73.4 pg/mL, and who converted to dementia can be found in online supplemental file 1 A,B.

Baseline NfL predicted changes during follow-up in the number of lacunes and CMB, and also NBV independent of their initial MRI baseline values and the patient’s age (table 5). The model’s AUC was 0.773 and 0.822 for lacunes and CMB, respectively.

Figure 2  NfL measures and incidences of new lacune over time per patient. The grey line refers to NfL levels over time in patients without dementia conversion. The dotted black line refers to NfL levels over time in patients with dementia conversion. The red (left), green (middle) and orange (right) dotted vertical lines represent an incidence of a new lacune in a patient between baseline and TP 1, TP 1 and 2, TP 2 and 3, respectively. NfL, neurofilament light chain; TP, time point.

Prediction by change in NfL there was no significant overall change in NfL over the 3 years follow-up period (table 3). In view of this lack of change in NfL, there was no association with dementia conversion ($n=15$) or change in any conventional MRI measure (tables 5 and 6). The individual NfL over time per patient together with incidences in lacune are found in figure 2A–C. In contrast, DTI significantly changed over time (both assessed by MDPH and PSMD, table 3) that predicted dementia conversion (table 6). Detailed information regarding serial change of NfL, MRI markers and cognition for patients converting and not converting to dementia are found in online supplemental...
Cerebrovascular disease

There were significant differences between the two patient groups for change in MDPH but not PSMD or NfL.

DISCUSSION
In this longitudinal prospective study in patients with moderate-to-severe SVD, NfL levels at baseline predicted both cognitive decline and conversion to dementia. The level of prediction was similar to that found for DTI, which has been previously suggested as the most sensitive marker to WM damage in SVD, and has been shown to predict future dementia risk. These findings suggest that NfL may be a useful marker to identify patients with SVD who are at higher risk of dementia. Previous evidence has shown that baseline NfL predicted disease severity and survival in CADASIL.

In contrast, change in NfL measured on repeat testing at yearly intervals did not predict cognitive decline or dementia. This was because no significant change in NfL levels could be detected during the follow-up period. Previous evidence has shown that serum NfL proved to be a reliable marker. In contrast, change in DTI, with sampling over the same time periods as NfL levels, predicted dementia risk. This implies that while NfL levels may provide a simple marker to predict risk, they may not be useful in assessing outcome in clinical trials with a duration of a few years, where change in the surrogate marker needs to predict change in a clinical marker such as conversion to dementia.

A previous cross-sectional study has shown that NfL levels were associated with the markers of disease burden, cognition and disability in SVD. Cross-sectional analysis of our baseline data confirmed the results with NfL levels associated with the degree of cognitive impairment, of disability and markers of disease burden.

We also showed that NfL at baseline predicts changes in imaging markers of SVD burden such as NBV, CMB and lacunes but not WMH in the longitudinal analysis. This evidence is partly in line with previous findings in a subgroup (<60 years of age) of the community-based stroke-free cohort showing that baseline but also change in NfL levels were the two main determinants for explaining brain atrophy. It is also broadly consistent with recent evidence in mild SVD showing baseline NfL predicts risk of future lacunes.

Why there was no significant change in NfL over 3 years in SCANS is uncertain and may be explained by dynamic changes in NfL post stroke. Findings from patients with small subcortical infarcts demonstrated that NfL was elevated but eventually returned to normal levels after 15 months. In SCANS blood sampling was unfortunately only taken annually not allowing us to test monthly changes in NfL levels. An alternative explanation for the lack of significant change is that the rate of change in NfL over 3 years with a sample size of 90 patients is too low in this chronic disease for any association to be detected.

This study has a number of strengths. Repeated sampling of both MRI and blood was performed and data on dementia conversion at 5 years was available for all subjects. MRI included both conventional markers and DTI, allowing the comparative performance of prediction by NfL levels with those of MRI markers to be determined. It also has limitations. NfL levels at more than one time point were not available on all subjects. It has been shown that dropouts had worse cognitive function than those who remained in the study, which might result in an under-estimation of the magnitude of cognitive, MRI and NfL changes over time.

This prospective cohort study demonstrates that baseline NfL values predict cognitive decline and dementia rate over a 5 years follow-up duration in patients with severe SVD. NfL may be a useful prognostic marker in this disease. However, in contrast to DTI, change in NfL values was not detected over a 3 years follow-up period with annual sampling, suggesting NfL is unlikely to be a useful surrogate marker in a phase 2 clinical trial, although further larger studies are required to confirm this finding.

Twitter Hugh S Markus @Camstroke
Acknowledgements The authors thank Prof. Dr. med. Marco Durier in connecting the two research groups.

Contributors ME was responsible for the data acquisition and analysis, interpretation of data, drafting the manuscript and critical revision of the manuscript for important intellectual content. He has full access to all of the data in the study. LL was responsible for the data acquisition and analysis as well as critical revision of the manuscript for important intellectual content. She has full access to all of the data in the study. AM was responsible for the data analysis, the interpretation of the data and critical revision of the manuscript for important intellectual content. HSM was responsible for the study concept and design, the data acquisition and analysis, the interpretation of the data, drafting the manuscript and critical revision of the manuscript for important intellectual content and has full access to all of the data in the study.

Funding This study was supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 667 375 (CoSTREAM). ME is supported by a PhD studentship funded by a Priority Programme Grant from the Stroke Association (award reference PPA 2015/02). The SCANS study was supported by a Welcome Trust grant #081589, awarded to HSM. The study was also supported by the Neurology Research Pool, which is an internal funding within the Hospital for neurological research (University Hospital Basel). HSM is supported by an NIHR Senior Investigator award and his work receives infrastructural support from the Cambridge University Hospitals NIHR Biomedical Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval SCANS was approved by the London–Wandsworth ethics committee (ukctg.nihr.ac.uk; study ID: 4577).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Cohort data can be shared upon reasonable request for scientific purpose by contacting the corresponding author. Anonymised data will be made available to qualified investigators on reasonable request to the corresponding author.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs
Marco Egle http://orcid.org/0000-0002-5247-7068
Hugh S Markus http://orcid.org/0000-0002-9794-5996

REFERENCES
Cerebrovascular disease

**Supplementary Table 1**: Clinical, NfL and imaging characteristics between patients with less vs. more decline in the longitudinal sample. Differences were tested using two-sample Welch’s t-test for variables with a parametric distribution and the Wilcoxon rank sum test for variables with a non-parametric distribution. There was a significant difference for NART, MMSE, NfL, NBV, Lacunes, MDPH and PSMD.

<table>
<thead>
<tr>
<th></th>
<th>Less Decline (N=45)</th>
<th>More Decline (N=45)</th>
<th>Overall (N=90)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>68.4 (9.61)</td>
<td>70.9 (9.31)</td>
<td>69.7 (9.49)</td>
</tr>
<tr>
<td><strong>NART</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>105 (15.4)</td>
<td>97.4 (14.6)</td>
<td>101 (15.4)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.4 (3.70)</td>
<td>11.6 (3.53)</td>
<td>12.0 (3.61)</td>
</tr>
<tr>
<td><strong>MMSE</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28.9 (1.09)</td>
<td>27.2 (2.59)</td>
<td>28.0 (2.15)</td>
</tr>
<tr>
<td><strong>NfL</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.2 (9.11)</td>
<td>37.6 (34.3)</td>
<td>28.5 (26.7)</td>
</tr>
<tr>
<td><strong>NBV</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1320 (60.9)</td>
<td>1270 (97.3)</td>
<td>1290 (84.6)</td>
</tr>
<tr>
<td><strong>WMH</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.06 (2.22)</td>
<td>3.82 (2.61)</td>
<td>3.44 (2.44)</td>
</tr>
<tr>
<td><strong>Lacunes</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.16 (2.76)</td>
<td>5.69 (6.31)</td>
<td>3.92 (5.16)</td>
</tr>
<tr>
<td><strong>CMB</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.96 (4.26)</td>
<td>3.78 (8.74)</td>
<td>2.87 (6.90)</td>
</tr>
<tr>
<td><strong>MDPH</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0161 (0.00274)</td>
<td>0.0144 (0.00224)</td>
<td>0.0153 (0.00263)</td>
</tr>
<tr>
<td><strong>PSMD</strong></td>
<td>Mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.000347 (0.0000747)</td>
<td>0.000403 (0.000109)</td>
<td>0.000375 (0.0000968)</td>
</tr>
</tbody>
</table>

* p < 0.05

NART- premorbid IQ score, MMSE - Mini–Mental State Examination, NfL- serum neurofilament light chain, NBV- normalised brain volume, WMH- white matter hyperintensity, CMB- cerebral microbleeds, MDPH- mean diffusivity normalised peak height, PSMD- peak width skeletonized mean diffusivity
**Supplementary Figure 1A-B.** MRI images of a patient aged 70 with a high NfL level (73.4 pg/mL) at baseline and who later converted to dementia. The MRI sequences (A) T1 and (B) FLAIR are shown.
**Supplementary Table 2.** Differences in clinical characteristics and estimated changes in NfL and in imaging markers between patients converting to dementia and those without dementia/censored over 5 years. Values post dementia were removed. Permutation Welch’s t-test showed that there were significant differences between the groups for Intercept NfL, Intercept MDPH, Slope MDPH.

<table>
<thead>
<tr>
<th></th>
<th>Censored, no dementia (N=75)</th>
<th>Developed dementia (N=15)</th>
<th>Overall (N=90)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>68.9 (9.55)</td>
<td>73.5 (8.47)</td>
<td>69.7 (9.49)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>70.6 [43.6, 88.8]</td>
<td>73.8 [53.8, 87.7]</td>
<td>70.8 [43.6, 88.8]</td>
</tr>
<tr>
<td><strong>NART</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>102 (15.1)</td>
<td>95.2 (16.1)</td>
<td>101 (15.4)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>103 [76.0, 127]</td>
<td>90.0 [75.0, 126]</td>
<td>103 [75.0, 127]</td>
</tr>
<tr>
<td><strong>Intercept NfL</strong>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.30 (0.168)</td>
<td>1.49 (0.268)</td>
<td>1.33 (0.200)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>1.30 [0.850, 1.73]</td>
<td>1.45 [1.16, 2.06]</td>
<td>1.32 [0.850, 2.06]</td>
</tr>
<tr>
<td><strong>Slope NfL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>-0.000208 (0.0315)</td>
<td>0.00638 (0.0186)</td>
<td>0.000890 (0.0297)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>-0.00124 [-0.119, 0.0867]</td>
<td>0.00431 [-0.0348, 0.0406]</td>
<td>0.000890 [-0.119, 0.0867]</td>
</tr>
<tr>
<td><strong>Intercept MDPH</strong>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.0156 (0.00261)</td>
<td>0.0136 (0.00204)</td>
<td>0.0153 (0.00263)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>0.0152 [0.00979, 0.0232]</td>
<td>0.0135 [0.00843, 0.0165]</td>
<td>0.0151 [0.00843, 0.0232]</td>
</tr>
<tr>
<td><strong>Slope MDPH</strong>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>-0.000377 (0.000100)</td>
<td>-0.000446 (0.000769)</td>
<td>-0.000388 (0.0000996)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>-0.000384 [-0.000739, -0.0000385]</td>
<td>-0.000408 [-0.000623, -0.0000373]</td>
<td>-0.000393 [-0.000739, -0.0000385]</td>
</tr>
<tr>
<td><strong>Intercept PSMD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.000366 (0.0000845)</td>
<td>0.000417 (0.000128)</td>
<td>0.000374 (0.0000944)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>0.000358 [0.000229, 0.000659]</td>
<td>0.000390 [0.000291, 0.000782]</td>
<td>0.000362 [0.000229, 0.000782]</td>
</tr>
<tr>
<td><strong>Slope PSMD</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.0000120 (0.0000113)</td>
<td>0.0000174 (0.0000177)</td>
<td>0.0000129 (0.0000127)</td>
</tr>
<tr>
<td>Median [Min, Max]</td>
<td>0.0000112 [-0.0000464, 0.0000340]</td>
<td>0.0000138 [-0.0000274, 0.0000507]</td>
<td>0.0000120 [-0.0000464, 0.0000507]</td>
</tr>
</tbody>
</table>

*P-values < 0.05