ARTICLE   OPEN

Plasma neurofilament light as a potential biomarker for cognitive decline in a longitudinal study of middle-aged urban adults

May A. Beydoun [1✉], Nicole Noren Hooten [1], Hind A. Beydoun[2], Ana I. Maldonado [1,3], Jordan Weiss[4], Michele K. Evans[1] and Alan B. Zonderman [1]

Plasma neurofilament light (NfL) is a marker for neurodegenerative diseases. Few studies have examined the association of NfL with middle-aged changes in cognitive performance, and no studies have examined differential NfL effects by race. Using data from the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study ($n = 625$, Agev1: 30–66 y, 41.6% male, 56.3% African American, 27.8% below poverty), we investigated the associations of initial NfL levels and annualized change with cognitive performance over time in global mental status, verbal and visual memory, fluency, attention, and executive function. We used ordinary least squares and mixed-effects regressions stratified by race, while exploring differential associations by age group, sex, and poverty status. Over a mean follow-up of 4.3 years, we found initial NfL level was associated with a faster decline on normalized mental status scores in Whites only and in those >50 years old. Annualized increase in NfL was associated with a greater decline in verbal fluency in men. In other exploratory analyses, annualized increase in NfL was associated with a slower decline in verbal memory among individuals living above poverty; in the older group (>50 years), first-visit NfL was linked with better performance at baseline in global mental status and verbal memory. In summary, first-visit NfL was primarily associated with the global mental status decline among Whites, while exhibiting inconsistent relationships in some exploratory analyses. Plasma NfL levels can be detected and quantified in non-demented middle-aged adults and changes can be analyzed over time. More longitudinal studies are needed to address the clinical utility of this biomarker for early cognitive defects.

INTRODUCTION
When axons become damaged, cytoskeletal proteins known as neurofilaments are released into the extracellular space, followed by the cerebrospinal fluid (CSF), with marked transmigration into the blood at a lower concentration [1]. Notably, among biomarkers for neurodegenerative disease, there is a need for minimally invasive, readily available, cost-effective biomarkers as current methods rely on measures derived from CSF and neuroimaging. Recently, sensitive methods were developed to measure blood-levels of neurofilament light (NfL) [2]. This methodological development for assaying plasma NfL has stimulated potential opportunities for large-scale applications in clinical practice and in randomized clinical trials as a method for identifying patients at risk for dementias, including Alzheimer's disease (AD) [3]. Thus far, NfL reflects sub-cortical large-caliber axonal degeneration [4, 5]. Plasma NfL levels correlate strongly with CSF NfL levels [3, 6], adding to its clinical utility in differential diagnoses for dementias. While most studies have focused on plasma NfL's positive association with AD, including at earlier stages [7–10], as well as other neurodegenerative diseases [11–14]. Thus, plasma NfL is a marker of non-specific neurodegeneration.

To date, only few studies have been conducted thus far reporting its predictive value for future cognitive decline and brain aging [15–22], and none have tested associations differentially across racial groups. Furthermore, few studies have examined how longitudinal changes in plasma NfL are related to change in cognition over time (e.g., [21]). Thus, our study (i) examined baseline NfL in relation to baseline and change in cognitive performance over time; (ii) examined change in NfL in relation to cognitive performance over time; (iii) examined baseline and change in NfL in relation to follow-up cognitive performance; and, (iv) tested racial differentials in those main associations; as well as exploring those associations across sex, age group, and poverty status.

MATERIALS AND METHODS
Database
We selected a sample from the Healthy Aging in Neighborhoods of Diversity across the Life Span (HANDLS) study. Since 2004, HANDLS is an on-going prospective cohort study of socioeconomically diverse White and African American adult women and men residing in Baltimore, MD. Initial

[1]Laboratory of Epidemiology and Population Sciences, NIA/NIH/IRP, Baltimore, MD, USA. [2]Department of Research Programs, Fort Belvoir Community Hospital, Fort Belvoir, VA, USA. [3]Department of Psychology, University of Maryland, Baltimore County, Catonsville, MD, USA. [4]Department of Demography, University of California, Berkeley, Berkeley, CA, USA. ✉email: baydounm@mail.nih.gov

data (visit 1) were collected between 2004 and 2009, in two phases. Phase I consisted of a home visit, with information collected for screening, recruitment, and a household in-person interview that included the first 24 h dietary recall of that visit. Phase II (v_1) was performed as an in-person complete physical health examination including a cognitive test battery inside Medical Research Vehicles (MRV) and included a second 24 h dietary recall. Participants were invited for follow-up in-person visits (v_2) between 2009 and 2013, which applied a similar protocol as v1 (phase II). Fasting blood samples were obtained from consenting participants in both in-person examinations. All participants provided written informed consent. The Institutional Review Board of the National Institutes of Health, National Institute of Environmental Health Sciences approved the HANDLS study protocol.

Study sample

In our present study, up to two repeats on cognitive tests were available from v_1 or v_2. Exposure data on plasma NfL concentrations were available at both visits for a sub-sample of Whites and African Americans after excluding participants who did not survive within a year of follow-up or who did not have NfL data at v_2. As shown in the study design flowchart (Fig. 1), among 3,720 initially recruited HANDLS participants, $N = 674$ had complete v_1 and v_2 data on plasma NfL. Of those participants, $N = 625$ had data on v_1 or v_2 for all 11 cognitive test scores, with an average number of observations/participant $k = 1.9-2.0$, indicating 0–5% missingness on cognitive test performance outcomes. A sub-set of those participants had complete and credible v_2 cognitive performance data, with somewhat variable sample sizes. This sub-set was also analyzed, thus excluding those with unavailable or non-credible v_2 cognitive performance on each test. Mean ± SD follow-up time for the final analytic sample ($n = 625$ participants) was $4.30 ± 0.95$ y. Method S1 shows a detailed description for sample selection with respect to the NfL exposure. Compared to the initial sample with incomplete data for our analysis, the final sample had a lower proportion of individuals living below poverty (27.8% vs. 43.9%, $p < 0.001$, χ^2 test), and a reduced proportion of men (41.6% vs. 45.9%, $p = 0.048$, χ^2 test). A similar pattern was observed when the sample with v_1 NfL ($N = 674$) was compared with the sample without this data, notwithstanding other exclusions.

Cognitive assessment

HANDLS clinical staff examined cognitive performance with a battery of tests which included the Mini-Mental State Examination (MMSE), the California Verbal Learning Test (CVLT) immediate (List A) and Delayed Free Recall (DFR), the Benton Visual Retention Test (BVRT, # of errors), Brief Test of Attention (BTA), Animal Fluency test (AF), the Digit Span Forward and Backwards tests (DS-F and DS-B), the Clock Drawing Test (CDT), Trailmaking test parts A and B (TRAILS A and B, in seconds), (described in detail in Method S2). Cognitive domains spanned global mental status, verbal memory, verbal fluency, attention, visual memory, visuo-spatial abilities, and executive function, which includes working memory. A total of 11 cognitive test scores were computed from these tests. Total MMSE was normalized using previously described methods [23]; while Trails A and B scores (in seconds) were Log_e transformed to achieve pseudo-normality. With the exception of BVRT, Trails A and B, all test scores were in the direction of higher values reflecting better performance at v_1 or over time.
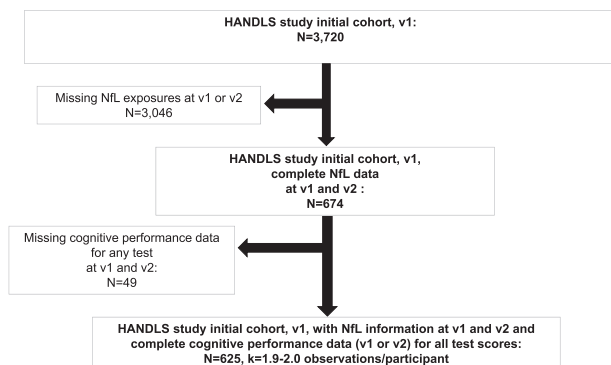
Plasma neurofilament light (NfL)

Fasting, morning plasma samples were collected into EDTA blood collection tubes. Tubes were centrifuged at $600 \times g$ for 15 min and the buffy coat was removed. These steps were repeated two times and the samples were visually examined for hemolysis. Plasma was aliquoted and stored at $-80\,°C$ until use. Plasma NfL levels were measured by Quanterix (Billerica, MA, USA) using the Simoa® NF-light Advantage Kit following the kit instruction. Longitudinal samples for each person were run on the same plate and the proportion of people in each demographic group (race/sex/poverty) were balanced across all plates. Plasma samples were diluted 1:4 and concentrations reflect the dilution correction. Pooled plasma samples from two individuals were run in duplicate on all plates. These duplicate pooled plasma samples were used to calculate both the within plate (intra-assay) and between plates (inter-assay) coefficient of variation (CV). The average intra-assay CV was 4.5% and the average inter-assay CV was 7%. The analytical limit of detection (LOD) was calculated as 2.5 standard deviations above the background (mean of calibrator blanks). For the analytical lower limit of quantification (LLOQ), triplicate measurements of serially diluted calibrator were run as unknowns and read on the standard calibration curve. The LLOQ was determined as the lowest dilution with a pooled CV ≤ 20% and a sample read back recovery between 80 and 120% of the expected concentration. The analytical upper limit of detection (ULOQ) was the highest concentration of the calibrator curve. Analytical LOD, LLOQ, and ULOD values were converted to functional values by multiplying by the dilution factor (4×) to enable direct comparison to the sample results. The functional LOD and the functional LLOQ were 0.152 and 0.696 pg/ml, respectively. The functional ULOD was 1872 pg/ml.

Covariates

Several covariates were considered in this study as potential confounders, given their previously shown association with cognitive performance or decline, which may also be associated with NfL exposures. These included v_1 age (continuous, years), sex (male, female), race (White, African American), poverty status (below vs. above 125% the federal poverty line), educational attainment (less than high school, high school, more than high school), and literacy (Wide Range Achievement Test, third edition [WRAT-3]). Age at v_2 was also used to compute time between v_1 and v_2, a measure relevant to our main models. Poverty status was operationalized using the 2004 US Census Bureau poverty thresholds [24] based on household income and total family size (including children <18 years). Furthermore, lifestyle and health-related factors were among those considered as potential confounders, given their potential impact on both exposures and outcomes. Those factors included current smoking status (0 = no vs. 1 = yes), illicit drug use (0 = no vs. 1 = yes, using any of marijuana, opiates, and cocaine), body mass index (BMI, weight/height^2, kg m^{-2}, continuous), self-rated health status categorized as 0=poor/average (referent), 1 = good and 2 = very good/excellent, the Healthy Eating Index 2010 (HEI-2010) [25], measuring overall diet quality based on food and macronutrient-related guidelines for Americans, total energy intake (kcal/d), and the 20-item CES-D total score for depressive symptoms. Moreover, an unweighted co-morbidity index was also accounted for. This index was composed of hypertension (0 = no, 1 = yes), diabetes (0 = diabetic, 1 = pre-diabetic, 2 = diabetic) and dyslipidemia (or statin use) (0 = no, 1 = yes), and self-reported history of any of several cardiovascular disease conditions (0 = no, 1 = yes). The latter component screened for the occurrence of several conditions, namely atrial fibrillation, angina, coronary artery disease, congestive heart failure, and myocardial infarction. Consequently, the co-morbidity index could potentially range between 0 and 5.

Statistical methods

Stata release 16 [26] was used to conduct all analyses. We first described the analytic sample's characteristics at baseline using means and proportions with bivariate linear, logistic, and multinomial logit models to examine racial differences in continuous, binary, and categorical multi-level covariates, respectively. We then adjusted those models for age, sex, and poverty status to determine whether racial differences remained statistically significant. Second, for testing our main hypotheses, a series of linear models were conducted (mixed-effects and ordinary least-square, OLS) (Method S3 for mixed-effects models). Separate analyses for 11 cognitive test scores were conducted, adjusting for two sets of covariates: Model 1: only socio-demographic variables: age at v1, sex, race, and poverty status; Model 2: socio-demographics + all other lifestyle and health-related covariates. To reduce missing data due to the addition of covariates into different models, given that each covariate had, individually

Fig. 1 Participant flowchart. Abbreviations: *HANDLS* Healthy Aging in Neighborhoods of Diversity across the Life Span; *k* = # of observations/participant; *NfL* neurofilament light; *v_1* visit 1; *v_2* visit 2.

Table 1. Study sample characteristics, overall and by race in the final analytic sample with imputed covariates ($N = 625$), HANDLS 2004–2013[a].

	Overall	Whites	African American
	($X \pm$ SE), %	($X \pm$ SE), %	($X \pm$ SE), %
	($N = 625$)	($N = 273$)	($N = 352$)
$X \pm$ SE or %±SE			
NfL at v_1, pg/mL			
Log_e transformed	+1.976 ± 0.020	+2.114 ± 0.029[****e]	+1.870 ± 0.026
Annualized rate of change in Log_e NfL between v_1 and v_2 δNfL	+0.044 ± 0.004	+0.038 ± 0.006	+0.050 ± 0.005
Baseline socio-demographic, SES and health-related variables			
Sex, % male	41.6 ± 2.0	40.3 ± 3.0	42.6 ± 2.6
Age at v_1, yrs.	47.9 ± 0.36	48.7 ± 0.51[**]	47.3 ± 0.51
African American, %	56.3 ± 2.0	0.00	100.0
Poverty status, % <125% of the 2004 federal poverty guidelines	27.8 ± 1.8	26.0 ± 2.7	29.3 ± 2.4
Education, Completed, %			
<HS	5.5 ± 0.9	7.9 ± 1.6[**e]	3.7 ± 1.0
HS	57.3 ± 2.0	57.3 ± 3.1	57.4 ± 2.6
>HS	37.1 ± 2.0	34.8 ± 3.0	38.9 ± 2.6
Literacy, WRAT-3 score	43.1 ± 0.3	44.9 ± 0.5[****e]	41.6 ± 0.4
Baseline drug and tobacco use			
Any drug, current user, %	16.3 ± 1.6	13.0 ± 2.1[*]	18.9 ± 2.2
Tobacco, current user, %	40.0 ± 2.0	39.2 ± 3.0	40.6 ± 2.6
BMI, kg/m^2	30.2 ± 0.3	30.1 ± 0.4	30.2 ± 0.4
Self-rated health, %			
Poor/Average,	19.4 ± 1.6	24.5 ± 2.6[***e]	15.3 ± 1.9
Good	41.3 ± 2.0	37.4 ± 2.9	44.3 ± 2.6
Very good/excellent	39.4 ± 2.0	38.1 ± 2.9	40.3 ± 2.6
HEI-2010 total score at v_1	42.1 ± 0.6	41.2 ± 0.8[*e]	42.8 ± 0.6
Total energy intake, kcal/day	1,986 ± 44	1,995 ± 64	1,978 ± 56.1
CES-D total score	14.1 ± 0.4	15.1 ± 0.70[**e]	13.3 ± 0.57
Hypertension[b], %	41.0 ± 2.0	36.2 ± 2.9[***e]	44.7 ± 2.7
Diabetes[b], %			
No	66.3 ± 2.0	62.2 ± 3.0	69.5 ± 2.6
Pre-diabetic	21.7 ± 1.6	24.5 ± 2.6[*]	19.4 ± 2.2
Diabetic	12.0 ± 1.4	13.3 ± 2.1	11.1 ± 1.7
Dyslipidemia[b], %	25.6 ± 1.8	29.4 ± 2.9[*]	22.7 ± 2.4
Cardiovascular disease[b], %	13.2 ± 1.4	11.6 ± 2.0	14.3 ± 2.0
Co-morbidity index[b]	3.26 ± 0.05	3.31 ± 0.08	3.23 ± 0.07
Cognitive performance at v_1, unadjusted[c]			
MMSE, normalized	76.9 ± 0.6	79.8 ± 0.9[****e]	74.6 ± 0.8
CVLT-List A	24.90 ± 0.28	26.1 ± 0.4[****e]	23.9 ± 0.4
CVLT-DFR	7.61 ± 0.13	8.30 ± 0.20[****e]	7.06 ± 0.17
BVRT	6.17 ± 0.20	5.93 ± 0.29	6.35 ± 0.27
BTA	6.80 ± 0.09	7.09 ± 0.14[***e]	6.52 ± 0.12
AF	19.07 ± 0.22	19.53 ± 0.34[*]	18.71 ± 0.28
DS-F	7.29 ± 0.09	7.58 ± 0.13[***e]	7.07 ± 0.11
DS-B	5.61 ± 0.08	6.03 ± 0.14[****e]	5.28 ± 0.10
CDT	8.77 ± 0.05	8.94 ± 0.07[***e]	8.64 ± 0.06
Log_e (TRAILS A)	3.44 ± 0.02	3.36 ± 0.02[****e]	3.51 ± 0.02
Log_e(TRAILS B)	4.57 ± 0.03	4.37 ± 0.04[****e]	4.72 ± 0.04
Annualized change in cognitive performance estimated between v_1 and v_2, unadjusted[c]			
MMSE, normalized	−0.06 ± 0.13	+0.02 ± 0.23	−0.05 ± 0.16
CVLT-List A	−1.25 ± 0.06[†]	−1.43 ± 0.10[†,**e]	−1.14 ± 0.06[†]
CVLT-DFR	−0.44 ± 0.03[†]	−0.47 ± 0.04[†]	−0.41 ± 0.03[†]

Table 1 continued

	Overall	Whites	African American
BVRT	$+0.49 \pm 0.04^{\dagger}$	$+0.33 \pm 0.06^{\dagger,****e}$	$+0.59 \pm 0.06^{\dagger}$
BTA	$-0.052 \pm 0.021^{\dagger}$	-0.033 ± 0.030	$-0.059 \pm 0.026^{\dagger}$
AF	$+0.075 \pm 0.038^{\dagger}$	$+0.058 \pm 0.066$	$+0.091 \pm 0.047$
DS-F	$+0.011 \pm 0.015$	$+0.030 \pm 0.026$	$+0.003 \pm 0.019$
DS-B	$+0.024 \pm 0.016$	$+0.055 \pm 0.029$	$+0.009 \pm 0.019$
CDT	-0.015 ± 0.013	-0.028 ± 0.021	-0.004 ± 0.017
Log_e (TRAILS A)	-0.0003 ± 0.0032	$+0.0037 \pm 0.0042$	-0.004 ± 0.005
Log_e(TRAILS B)	$+0.0156 \pm 0.005^{\dagger}$	$+0.0234 \pm 0.008^{\dagger}$	$+0.010 \pm 0.007$

Abbreviations: *AF* Animal Fluency; *BMI* body mass index; *BTA* Brief Test of Attention; *BVRT* Benton Visual Retention Test; *CDT* Clock Drawing Test; *CES-D* Center for Epidemiologic Studies-Depression; *CVLT-DFR* California Verbal Learning Test-Delayed Free Recall; *CVLT-List A* California Verbal Learning Test-List A; *DS-B* Digits Span-Backward; *DS-F* Digits Span-Forward; *HANDLS* Healthy Aging in Neighborhoods of Diversity across the Life Span; *HEI-2010* Healthy Eating Index, 2010 version; *HS* high school; *MMSE* Mini-Mental State Examination; *SE* standard error; *TRAILS A* Trailmaking test, part A; *TRAILS B* Trailmaking test, part B; *WRAT-3* Wide Range Achievement Test, 3rd revision; *X* mean.

*$p < 0.10$; **$p < 0.05$; *** $p < 0.010$; ****$p < 0.001$, *t*-test for the null hypothesis of no between-race differences.

$^{\dagger}p < 0.05$, *t*-test for the null hypothesis of $\gamma_1 = 0$ (fixed effects coefficient for *TIME*) in mixed-effects linear regression models with *TIME* as the only variable.

[a]Values are means $(X) \pm SE$ for continuous variables and % for categorical variables. The sample selected has complete data on MMSE and 10 other cognitive test scores at visits 1 and/or 2 and complete data on ApoE genotypes. Other covariates were multiple imputed (five imputations with ten iterations), using chained equations. All cognitive test scores are in the direction of higher score → better performance with the exception of BVRT (# of errors) and TRAILS A and B (# of sec. to complete).

[b]The co-morbidity index was calculated as the sum of hypertension, diabetes, and dyslipidemia (or statin use), and self-reported history of cardiovascular disease included atrial fibrillation, angina, coronary artery disease, congestive heart failure, or myocardial infarction, ranging from 0 to 5.

[c]Crude baseline cognitive test score. Sample sizes varied between 492 and 624 for the overall sample.

[d]Crude estimated the annual rate of change in cognitive performance based on mixed-effects linear regression model with TIME as the only covariate. Difference by race was determined by interacting TIME with race.

[e]$p < 0.05$ upon further adjustment for age, sex, and poverty status in multiple linear, logistic, multinomial logit, and mixed-effects linear regression models with race entered as the main predictor.

<5% missing on average, we ensured sample sizes were constant between reduced and fully adjusted models by conducting multiple imputations (five imputations, ten iterations), using the chained equations methodology. All covariates were used simultaneously during this estimation process, similar to previous studies [27, 28] and continuous covariates were centered around their means. Thus, for mixed-effects linear regression models, we applied Models 1 and 2 to two exposures (NfL and δNfL), 11 cognitive test scores with up to two repeats (effect of exposures on v_1 cognitive performance (CP_{v1}) and cognitive performance change over time (δCP)), one main stratifying variable (race), and several exploratory stratifying variables (sex, age group, and poverty status). NfL was Log_e transformed in all these analyses, and the annualized changes in the Log_e transformed NfL between v_1 and v_2 were used to operationalize δNfL [i.e., $\delta NfL = (Log_e(NfL_{v2}) - Log_e(NfL_{v1})/(Age_{v2} - Age_{v1})]$, using complete case analysis. *Z*-scoring for exposures was done using the final eligible sample ($N = 625$). These two exposures were constructed in a similar way in other studies (e.g., [21]). Racial differences in the association between NfL exposures and cognitive performance at v_1 was tested using NfL × Race and δNfL × Race interaction terms in separate models, while that of the association between NfL exposures and cognitive change was carried out by testing the NfL/δNfL × TIME × Race term in the same model. Following a similar approach but with a set of OLS linear regression models, race-specific associations of v_1 NfL and δNfL with v_2 cognitive performance (CP_{v2}) as an outcome of interest, were examined, while additionally adjusting models with the time of follow-up (years) between v_1 and v_2. Racial differences were also tested using two-way interaction terms (NfL × Race) in unstratified models, as were differences by age group, sex, and poverty status.

In all models, sample selectivity due to missing exposure and outcome data, relative to the initially recruited sample, was adjusted for using a two-stage Heckman selection strategy. Thus, we first predicted an indicator of selection with socio-demographic factors, namely, v_1 age, race, sex, and poverty status using a probit regression model, which yielded an inverse mills ratio (IMR), a function of the probability of being selected given those socio-demographic factors. At a second stage, we estimated our multiple mixed-effects and OLS linear regression models adjusted for the IMR in addition to the aforementioned covariates [29].

This study set the Type I error rate a priori for main and interactive effects before correction for multiple testing to 0.05 and 0.10, respectively [30]. We accounted for outcome multiplicity (i.e., 11 cognitive test scores) using the approach of familywise Bonferroni correction [31], specifically for Model 1.

Subsequently, the full model (Model 2) was considered a sensitivity model in which potentially confounding and/or mediating factors were included. In addition, a reduced version of Model 2 (Model 3) was tested, whereby only covariates, aside from those included in Model 1, shown to be associated with each of the two exposures were included. This model was only conducted as a sensitivity analysis. Therefore, we adjusted significance levels to $p < 0.00455$ (0.05/11), and for two-way interaction terms to $0.10/11 = 0.00910$, similar to previous work [32]. Moreover, *q*-values (false discovery rates) were also computed as an alternative means to correct for multiple testing in Model 1, accounting for multiplicity in cognitive tests only [33, 34]. *Q*-values < 0.05 were used for statistical significance for main effects (e.g., effect of NfL_{v1}), while $0.05 \leq q$-values < 0.10 were considered as significant for two-way interaction terms (e.g., effect of $NfL_{v1} \times TIME$). In our exploratory stratified analysis, all main hypotheses were tested across sex, age group (≤50 y, >50 y, as 50 y was the approximate median age) and poverty status (above vs. below poverty), separately, using the same modeling approach; and only familywise Bonferroni correction was applied to this part of the analysis (Model 1). Main findings were illustrated using predictive margins (with estimated 95% CI) of outcomes across time, and by exposure, overall or stratified by race and/or the other socio-demographic factors, using a specific mixed-effects or OLS linear regression model. Data analysis code in parts or in full can be made available upon request to the corresponding author.

RESULTS

Overall, and based on Table 1, participants were ~48 years old at initial testing; African Americans were significantly older than Whites (48.7 vs. 47.3, $p < 0.05$). A significantly higher proportion of Whites than African Americans had <HS education (7.9% vs. 3.7%). Although there were no race differences in poverty status, mean literacy was significantly higher among Whites. Log_e transformed NfL_{v1} plasma concentration was significantly higher among Whites compared with African Americans. However, there were no significant differences between races in the annualized rate of change values of NfL (delta NfL; δNfL). Current drug use was higher among African Americans; CES-D total score was higher among Whites. Although the co-morbidity index did not differ by race, dyslipidemia was more prevalent among Whites and

Table 2. Baseline and annual rates of change in plasma neurofilament light (v^1 NfL, and δNfL) and their association with cognitive performance at v^1 and change over time: overall and race-specific mixed-effects linear regression models: HANDLS 2004–2013[a].

	NfL, pg/mL, (v_1 Log_e transformed, z-scored)		δNfL, pg/mL (annualized change between v^1 and v^2, Log^e transformed, z-score)	
	Model 1	Model 2	Model 1	Model 2
	γ ± SE	γ ± SE	γ ± SE	γ ± SE
Overall	**(N = 625, k = 1.9−2.0)**	**(N = 625, k = 1.9−2.0)**	**(N = 625, k = 1.9−2.0)**	**(N = 625, k = 1.9−2.0)**
Outcome = Cognitive performance test score				
Normalized MMSE				
Exposure, $γ_{0a}$	+1.024 ± 0.667[b]	+0.688 ± 0.607[b]	+0.719 ± 0.574	+1.009 ± 0.508**
Exposure × TIME, $γ_{1a}$	−0.254 ± 0.158[c]	−0.208 ± 0.161[c]	−0.004 ± 0.144	−0.086 ± 0.143
CVLT-List A				
Exposure, $γ_{0a}$	+0.404 ± 0.310	+0.486 ± 0.302	−0.208 ± 0.264	−0.177 ± 0.249
Exposure × TIME, $γ_{1a}$	−0.078 ± 0.067	−0.051 ± 0.070	+0.061 ± 0.060	+0.035 ± 0.060
CVLT-DFR				
Exposure, $γ_{0a}$	+0.164 ± 0.144	+0.226 ± 0.143	−0.127 ± 0.122	−0.139 ± 0.118
Exposure × TIME, $γ_{1a}$	−0.038 ± 0.031	−0.029 ± 0.032	+0.051 ± 0.027*	+0.045 ± 0.028
BVRT				
Exposure, $γ_{0a}$	−0.072 ± 0.223	−0.211 ± 0.216	−0.038 ± 0.192	−0.042 ± 0.181
Exposure × TIME, $γ_{1a}$	+0.059 ± 0.048	+0.058 ± 0.051	+0.005 ± 0.044	+0.015 ± 0.044
BTA				
Exposure, $γ_{0a}$	+0.164 ± 0.105	+0.100 ± 0.105	+0.073 ± 0.088	+0.102 ± 0.086
Exposure × TIME, $γ_{1a}$	−0.036 ± 0.025	-0.037 ± 0.026	−0.013 ± 0.022	−0.016 ± 0.022
AF				
Exposure, $γ_{0a}$	−0.107 ± 0.253	−0.176 ± 0.250	+0.094 ± 0.218	+0.097 ± 0.210
Exposure × TIME, $γ_{1a}$	+0.007 ± 0.046	+0.006 ± 0.048	−0.071 ± 0.041*	−0.077 ± 0.042*
DS-F				
Exposure, $γ_{0a}$	+0.104 ± 0.103	+0.046 ± 0.098	−0.021 ± 0.089	0.0077 ± 0.083
Exposure × TIME, $γ_{1a}$	−0.015 ± 0.018	−0.018 ± 0.019	−0.010 ± 0.016	−0.011 ± 0.017
DS-B				
Exposure, $γ_{0a}$	+0.056 ± 0.097[b]	+0.009 ± 0.091	+0.066 ± 0.084	+0.094 ± 0.076
Exposure × TIME, $γ_{1a}$	−0.023 ± 0.020	−0.025 ± 0.020	+0.005 ± 0.018	+0.000 ± 0.018
CDT				
Exposure, $γ_{0a}$	+0.057 ± 0.056[b]	+0.040 ± 0.057[b]	−0.005 ± 0.048	−0.005 ± 0.048
Exposure × TIME, $γ_{1a}$	−0.011 ± 0.015[c]	−0.010 ± 0.016[c]	−0.003 ± 0.014 [c]	+0.001 ± 0.014 [c]
Log_e(TRAILS A)				
Exposure, $γ_{0a}$	+0.032 ± 0.016*	+0.025 ± 0.017	+0.013 ± 0.014	+0.014 ± 0.014
Exposure × TIME, $γ_{1a}$	−0.001 ± 0.004	−0.001 ± 0.004	+0.002 ± 0.004	+0.0021 ± 0.004
Log_e(TRAILS B)				
Exposure, $γ_{0a}$	+0.026 ± 0.029	+0.023 ± 0.028	+0.014 ± 0.025	+0.010 ± 0.024
Exposure × TIME, $γ_{1a}$	−0.004 ± 0.006	−0.004 ± 0.006	+0.001 ± 0.006	+0.002 ± 0.006
Whites	**(N = 273, k = 1.9−2.0)**	**(N = 273, k = 1.9−2.0)**	**(N = 273, k = 1.9−2.0)**	**(N = 273, k = 1.9−2.0)**
Outcome = Cognitive performance test score				
Normalized MMSE				
Exposure, $γ_{0a}$	+2.334 ± 0.988**	+1.616 ± 0.836*	+0.279 ± 0.880	+0.581 ± 0.731
Exposure × TIME, $γ_{1a}$	**−0.661 ± 0.252*****	−0.565 ± 0.246**	+0.060 ± 0.239	+0.005 ± 0.230
CVLT-List A				
Exposure, $γ_{0a}$	+0.311 ± 0.458	+0.194 ± 0.433	−0.518 ± 0.395	−0.369 ± 0.370
Exposure × TIME, $γ_{1a}$	−0.025 ± 0.116	+0.026 ± 0.120	+0.099 ± 0.106	+0.071 ± 0.108
CVLT-DFR				
Exposure, $γ_{0a}$	+0.238 ± 0.214	+0.190 ± 0.208	−0.299 ± 0.184	−0.290 ± 0.178
Exposure × TIME, $γ_{1a}$	−0.032 ± 0.049	−0.008 ± 0.050	+0.092 ± 0.045**	+0.086 ± 0.045*

Table 2 continued

	NfL, pg/mL, (v$_1$ Log$_e$ transformed, z-scored)		δNfL, pg/mL (annualized change between v^1 and v^2, Log^e transformed, z-score)	
BVRT				
Exposure, γ_{0a}	-0.217 ± 0.319	-0.206 ± 0.294	+0.147 ± 0.283	+0.008 ± 0.257
Exposure × TIME, γ_{1a}	+0.076 ± 0.067	+0.066 ± 0.069	−0.058 ± 0.063	−0.027 ± 0.063
BTA				
Exposure, γ_{0a}	+0.375 ± 0.160**	+0.273 ± 0.158*	−0.010 ± 0.135	+0.013 ± 0.130
Exposure × TIME, γ_{1a}	−0.072 ± 0.040*	−0.055 ± 0.040	−0.002 ± 0.036	−0.003 ± 0.035
AF				
Exposure, γ_{0a}	−0.282 ± 0.379	−0.442 ± 0.363	+0.048 ± 0.335	−0.004 ± 0.321
Exposure × TIME, γ_{1a}	+0.004 ± 0.075	−0.005 ± 0.077	−0.075 ± 0.070	−0.063 ± 0.071
DS-F				
Exposure, γ_{0a}	0.122 ± 0.154	+0.030 ± 0.136	+0.118 ± 0.136	+0.108 ± 0.118
Exposure × TIME, γ_{1a}	−0.021 ± 0.030	−0.025 ± 0.030	+0.012 ± 0.028	+0.016 ± 0.028
DS-B				
Exposure, γ_{0a}	+0.217 ± 0.153	+0.103 ± 0.137	+0.167 ± 0.136	+0.207 ± 0.119*
Exposure × TIME, γ_{1a}	−0.048 ± 0.033	−0.044 ± 0.034	+0.031 ± 0.031	+0.022 ± 0.031
CDT				
Exposure, γ_{0a}	−0.075 ± 0.079	−0.066 ± 0.078	+0.085 ± 0.070	+0.091 ± 0.068
Exposure × TIME, γ_{1a}	+0.036 ± 0.023	+0.035 ± 0.024	−0.041 ± 0.022*	−0.043 ± 0.022*
Log$_e$(TRAILS A)				
Exposure, γ_{0a}	+0.034 ± 0.021	+0.038 ± 0.020*	+0.004 ± 0.019	+0.004 ± 0.018
Exposure × TIME, γ_{1a}	−0.002 ± 0.005	−0.001 ± 0.005	+0.002 ± 0.004	+0.001 ± 0.005
Log$_e$(TRAILS B)				
Exposure, γ_{0a}	+0.000 ± 0.041	+0.018 ± 0.037	+0.029 ± 0.036	+0.019 ± 0.032
Exposure × TIME, γ_{1a}	−0.002 ± 0.009	−0.004 ± 0.009	+0.002 ± 0.008	+0.002 ± 0.008
African American	**(N = 352, k = 1.9)**	**(N = 352, k = 1.9)**	**(N = 352, k = 1.9)**	**(N = 352, k = 1.9)**
Outcome = Cognitive performance test score				
Normalized MMSE				
Exposure, γ_{0a}	−0.183 ± 0.910	−0.075 ± 0.870	+1.153 ± 0.756	+1.305 ± 0.705*
Exposure × TIME, γ_{1a}	+0.047 ± 0.203	+0.025 ± 0.210	−0.053 ± 0.180	−0.102 ± 0.182
CVLT-List A				
Exposure, γ_{0a}	+0.336 ± 0.413	+0.562 ± 0.412	+0.097 ± 0.343	+0.053 ± 0.332
Exposure × TIME, γ_{1a}	−0.099 ± 0.080	−0.108 ± 0.084	+0.040 ± 0.068	+0.023 ± 0.070
CVLT-DFR				
Exposure, γ_{0a}	+0.066 ± 0.193	+0.216 ± 0.197	+0.039 ± 0.160	−0.009 ± 0.158
Exposure × TIME, γ_{1a}	−0.046 ± 0.040	−0.052 ± 0.042	+0.022 ± 0.034	+0.017 ± 0.035
BVRT				
Exposure, γ_{0a}	+0.158 ± 0.317	−0.117 ± 0.314	−0.228 ± 0.264	−0.139 ± 0.255
Exposure × TIME, γ_{1a}	+0.025 ± 0.069	+0.034 ± 0.072	+0.046 ± 0.060	+0.043 ± 0.061
BTA				
Exposure, γ_{0a}	+0.020 ± 0.140	+0.005 ± 0.141	+0.157 ± 0.116	+0.173 ± 0.115
Exposure × TIME, γ_{1a}	−0.023 ± 0.032	−0.036 ± 0.034	−0.025 ± 0.028	−0.022 ± 0.028
AF				
Exposure, γ_{0a}	−0.110 ± 0.336	−0.062 ± 0.337	+0.193 ± 0.280	+0.182 ± 0.275
Exposure × TIME, γ_{1a}	+0.007 ± 0.059	+0.016 ± 0.061	−0.075 ± 0.051	−0.098 ± 0.052*
DS-F				
Exposure, γ_{0a}	+0.081 ± 0.139	+0.043 ± 0.137	−0.130 ± 0.116	−0.085 ± 0.112
Exposure × TIME, γ_{1a}	−0.011 ± 0.024	−0.013 ± 0.025	−0.023 ± 0.021	−0.025 ± 0.021
DS-B				
Exposure, γ_{0a}	−0.086 ± 0.126	−0.082 ± 0.118	−0.020 ± 0.103	−0.021 ± 0.096
Exposure × TIME, γ_{1a}	−0.004 ± 0.024	−0.007 ± 0.025	−0.008 ± 0.021	−0.007 ± 0.022

Table 2 continued

	NfL, pg/mL, (v$_1$ Log$_e$ transformed, z-scored)		δNfL, pg/mL (annualized change between v^1 and v^2, Log^e transformed, z-score)	
CDT				
Exposure, γ$_{0a}$	+0.156 ± 0.079*	+0.134 ± 0.081	−0.072 ± 0.066	−0.070 ± 0.066
Exposure × TIME, γ$_{1a}$	−0.046 ± 0.021**	−0.043 ± 0.021**	+0.022 ± 0.0180	+0.026 ± 0.018
Log$_e$(TRAILS A)				
Exposure, γ$_{0a}$	+0.025 ± 0.025	+0.008 ± 0.025	+0.020 ± 0.021	+0.027 ± 0.020
Exposure × TIME, γ$_{1a}$	−0.001 ± 0.006	−0.001 ± 0.006	+0.003 ± 0.005	+0.003 ± 0.005
Log$_e$(TRAILS B)				
Exposure, γ$_{0a}$	+0.039 ± 0.042	+0.019 ± 0.042	−0.001 ± 0.035	−0.001 ± 0.034
Exposure × TIME, γ$_{1a}$	−0.004 ± 0.009	−0.003 ± 0.009	+0.001 ± 0.007	+0.003 ± 0.007

Abbreviations: AF Animal Fluency; *BTA* Brief Test of Attention; *BVRT* Benton Visual Retention Test; *CDT* Clock Drawing Test; *CES-D* Center for Epidemiologic Studies-Depression; *CVLT-DFR* California Verbal Learning Test-Delayed Free Recall; *CVLT-List A* California Verbal Learning Test-List A; *DS-B* Digits Span-Backward; *DS-F* Digits Span-Forward; *HANDLS* Healthy Aging in Neighborhoods of Diversity across the Life Span; *HEI-2010* Healthy Eating Index, 2010 version; *k* number of observations/participant; *MMSE* Mini-Mental State Examination; *SD* standard deviation; *SE* standard error; *NfL* neurofilament light; *TRAILS A* Trailmaking test, part A; *TRAILS B* Trailmaking test, part B; *WRAT-3* Wide Range Achievement Test, 3rd revision; *X* mean.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.010$; ****$p < 0.001$, test for null hypothesis of γ = 0. Bolded values passed correction for multiple testing; underlined values passed $q < 0.05$ correction for multiple testing in Model 1.

[a]Models 1A.1−1K.2 included each of NfL (Log$_e$ transformed, z-scored) or δNfL (annualized change in Log$_e$ transformed NfL, z-scored), separately as the main predictor for v1 cognitive performance and cognitive change over time (11 test scores), using a series of mixed-effects linear regression models, carried out in the overall population, and stratified by race. These models adjusted only for age, sex, race, poverty status, and the inverse mills ratio. Models 2A.1−2K.2 followed a similar approach but adjusted further for selected socio-demographic, lifestyle, and health-related factors, namely educational attainment, the WRAT-3 score, current drug use, current tobacco use, body mass index, self-rated health, co-morbidity index, HEI-2010, total energy intake, and the CES-D total score. 1 SD of baseline Log$_e$(NfL) is estimated at 0.51; mean = 1.98. dNfL values are annualized changes in Log$_e$ transformed NfL between v$_1$ and v$_2$, z-scored. 1 SD of annualized change in Log$_e$(NfL) is estimated at 0.101; mean = 0.044.

[b]$p < 0.05$ for Race × NfL/δNfL in models that are unstratified by race to which this three-way interaction was included.

[c]$p < 0.05$ for Race × NfL/δNfL × TIME in models that are unstratified by race to which this two-way interaction was included.

hypertension was more prevalent among African Americans. In this select sample, Whites performed better than African Americans on most cognitive tests at v$_1$. Whites had a greater rate of decline on CVLT-List A and a smaller rate of decline on the BVRT than African Americans.

Our main hypotheses of associations between plasma NfL exposures and time-dependent cognitive outcomes were examined by mixed-effects and OLS regression models (Tables 2, 3) and are summarized in Fig. S1. Our exploratory analyses by age group, sex, and poverty status are presented in Tables S1−S3. Over a mean follow-up of 4.3 years, no association retained statistical significance upon correction for multiple testing in the total sample. However, we found initial NfL (i.e., NfL$_{v1}$) was associated with faster decline on normalized mental status scores in Whites only (δMMSE$_{norm.}$: γ$_{11}$ = −0.661 ± 0.252, P = 0.0085, q = 0.094, reduced model), an association that retained significance in the fully adjusted model 2. This association (NfL$_{v1}$ vs. decline in performance) was also found in those >50 years of age (δMMSE$_{norm}$: γ$_{11}$ = −0.705 ± 0.242, P = 0.004, reduced model); (Tables 2 and S2). Annualized increase in NfL was associated with greater decline in verbal fluency in men (δAF: γ$_{11}$ = −0.181 ± 0.058, P = 0.002, full model); (Table S1). In other exploratory analyses (Tables S1−S3), annualized increase in NfL was associated with slower decline in verbal memory among individuals living above poverty (δCVLT-DFR: +0.104 ± 0.036, P = 0.004, reduced model), while, in the older group (>50 years), first-visit NfL was linked with better performance at baseline in global mental status and verbal memory (P < 0.004). Finally, and upon correction for multiple testing, no stratum-specific associations were found between NfL$_{v1}$ (or δNfL) and follow-up cognitive performance. Reduction of Model 2 to Model 3, leaving in only additional covariates (in addition to socio-demographics) that were associated with NfL exposures, did not alter our main findings.

The main finding among Whites, for NfL$_{v1}$ vs. normalized MMSE scores across time is presented in terms of predictive margins of

outcome per SD of exposure in Fig. 2A. The Figure indicates that among those with higher NfL$_{v1}$ (i.e., v$_1$ Log$_e$ transformed plasma NfL, z-scored: mean + 1 SD), normalized MMSE score was on a decline over a period of 5 years as opposed to participants with NfL$_{v1}$ at the mean or at mean − 1 SD, whose performance was improving over time, from an initial low level. This was not the case among African Americans. Figure S1 summarizes findings from Model 1, across race, for all regression analyses with 11 cognitive test scores, three types of outcomes, and two exposures. Figure 2B−E shows predictive margins of cognitive performance tests across exposure levels (NfL$_{v1}$ and δNfL: z-score for annualized change in Log$_e$ transformed plasma NfL between v$_1$ and v$_2$) and by sex, age group, and poverty status, highlighting the key exploratory findings.

DISCUSSION
Main findings
This study is one of the few to examine plasma NfL baseline level (NfL$_{v1}$) and its annualized rate of change over a 5 y follow-up (δNfL) and the longitudinal associations with cognitive performance in middle-aged adults over the same period of time. The study was specifically conducted among a bi-racial urban cohort of middle-aged men and women who were free from dementia at baseline. The sampling strategy allowed us to examine key tested associations across racial groups, and secondarily across sex, age, and poverty status groups. Cognitive performance was measured twice for most selected participants, reflecting global mental status and domains of verbal memory and fluency, visual memory and visuo-spatial abilities, attention, and executive functions. Over a mean follow-up of 4.3 years, we found initial NfL was associated with a faster decline on normalized mental status scores in Whites only and in those >50 years old. Annualized increase in NfL was associated with a greater decline in verbal fluency in men. In other exploratory analyses, annualized increase in NfL was associated

Table 3. Baseline plasma neurofilament light (v[1] NfL and δNfL) and their association with cognitive performance at v[2]: overall and race-specific multiple ordinary least square linear regression models: HANDLS 2004−2013[a].

	Whites		African Americans	
	Model 1	Model 2	Model 1	Model 2
	$\beta \pm SE$	$\beta \pm SE$	$\beta \pm SE$	$\beta \pm SE$
NfL, pg/mL, (v[1] Log$_e$ transformed, z-scored)				
Outcome = cognitive performance test score				
Normalized MMSE	$N = 258$	$N = 258$	$N = 330$	$N = 330$
	-0.622 ± 0.915	-0.852 ± 0.869	$+0.134 \pm 0.923$	$+0.048 \pm 0.935$
CVLT-List A	$N = 272$	$N = 272$	$N = 349$	$N = 349$
	$+0.208 \pm 0.515$	$+0.243 \pm 0.521$	-0.180 ± 0.442	-0.062 ± 0.447
CVLT-DFR	$N = 272$	$N = 272$	$N = 349$	$N = 349$
	$+0.117 \pm 0.226$	$+0.148 \pm 0.230$	-0.139 ± 0.189	-0.033 ± 0.193
BVRT	$N = 273$	$N = 273$	$N = 350$	$N = 350$
	$+0.127 \pm 0.330$	$+0.146 \pm 0.318$	$+0.311 \pm 0.316$	$+0.089 \pm 0.310$
BTA	$N = 266$	$N = 266$	$N = 346$	$N = 346$
	$+0.072 \pm 0.153$	$+0.015 \pm 0.149$	-0.091 ± 0.146	-0.163 ± 0.146
AF	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	-0.289 ± 0.386	-0.488 ± 0.387	-0.111 ± 0.332	-0.032 ± 0.338
DS-F	$N = 258$	$N = 258$	$N = 345$	$N = 345$
	$+0.041 \pm 0.169$	-0.073 ± 0.155	$+0.012 \pm 0.145$	-0.039 ± 0.145
DS-B	$N = 257$	$N = 257$	$N = 343$	$N = 343$
	-0.037 ± 0.170	-0.160 ± 0.153	-0.103 ± 0.136	-0.123 ± 0.133
CDT	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	$+0.065 \pm 0.088$	$+0.067 \pm 0.090$	-0.055 ± 0.080	-0.074 ± 0.083
Log$_e$(TRAILS A)	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	$+0.024 \pm 0.023$	$+0.031 \pm 0.022$	$+0.021 \pm 0.025$	$+0.007 \pm 0.026$
Log$_e$(TRAILS B)	$N = 272$	$N = 272$	$N = 351$	$N = 351$
	-0.001 ± 0.043	$+0.008 \pm 0.04$	$+0.023 \pm 0.046$	$+0.01 \pm 0.045$
δNfL, pg/mL (annualized change between v[1] and v[2], Log$_e$ transformed, z-scored)				
Outcome = Cognitive performance test score				
Normalized MMSE	$N = 258$	$N = 258$	$N = 330$	$N = 330$
	$+0.484 \pm 0.812$	$+0.692 \pm 0.756$	$+1.371 \pm 0.780$*	$+1.144 \pm 0.783$
CVLT-List A	$N = 272$	$N = 272$	$N = 349$	$N = 349$
	-0.193 ± 0.457	-0.135 ± 0.456	$+0.270 \pm 0.368$	$+0.186 \pm 0.365$
CVLT-DFR	$N = 272$	$N = 272$	$N = 349$	$N = 349$
	$+0.032 \pm 0.201$	$+0.043 \pm 0.202$	$+0.154 \pm 0.157$	$+0.097 \pm 0.157$
BVRT	$N = 273$	$N = 273$	$N = 350$	$N = 350$
	-0.048 ± 0.293	-0.062 ± 0.272	-0.049 ± 0.265	$+0.034 \pm 0.253$
BTA	$N = 266$	$N = 266$	$N = 346$	$N = 346$
	$+0.007 \pm 0.135$	$+0.016 \pm 0.129$	$+0.054 \pm 0.122$	$+0.073 \pm 0.118$
AF	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	-0.217 ± 0.343	-0.234 ± 0.334	-0.141 ± 0.277	-0.211 ± 0.275
DS-F	$N = 258$	$N = 258$	$N = 345$	$N = 345$
	$+0.168 \pm 0.147$[b]	$+0.165 \pm 0.132$	-0.219 ± 0.120 *[b]	-0.186 ± 0.119
DS-B	$N = 257$	$N = 257$	$N = 343$	$N = 343$
	$+0.296 \pm 0.148$**[b]	$+0.315 \pm 0.13$**	-0.060 ± 0.115[b]	-0.067 ± 0.110
CDT	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	-0.064 ± 0.078	-0.069 ± 0.078	$+0.023 \pm 0.067$	$+0.042 \pm 0.068$
Log$_e$(TRAILS A)	$N = 273$	$N = 273$	$N = 351$	$N = 351$
	$+0.012 \pm 0.020$	$+0.013 \pm 0.019$	$+0.033 \pm 0.021$	$+0.037 \pm 0.021$*
Log$_e$(TRAILS B)	$N = 272$	$N = 272$	$N = 351$	$N = 351$
	$+0.033 \pm 0.038$	$+0.031 \pm 0.035$	0.000 ± 0.038	$+0.009 \pm 0.037$

Abbreviations: *AF* Animal Fluency; *BTA* Brief Test of Attention; *BVRT* Benton Visual Retention Test; *CDT* Clock Drawing Test; *CES-D* Center for Epidemiologic Studies-Depression; *CVLT-DFR* California Verbal Learning Test-Delayed Free Recall; *CVLT-List A* California Verbal Learning Test-List A; *DS-B* Digits Span-Backward; *DS-F* Digits Span-Forward; *HANDLS* Healthy Aging in Neighborhoods of Diversity across the Life Span; *HEI-2010* Healthy Eating Index, 2010 version; *MMSE* Mini-Mental State Examination; *k* number of observations/participant; *SD* standard deviation; *NfL* neurofilament light; *TRAILS A* Trailmaking test, part A; *TRAILS B* Trailmaking test, part B; *WRAT-3* Wide Range Achievement Test, 3rd revision.

with slower decline in verbal memory among individuals living above poverty, while, in the older group (>50 years), first-visit NfL was linked with better performance at baseline in global mental status and verbal memory.

Previous studies and biological mechanisms

Currently, methods to diagnose and monitor neuropathology are based on various imaging modalities, which are expensive with limited availability. CSF biomarkers, including NfL, have also been utilized, but require invasive procedures. Therefore, non-invasive biomarkers of neurocognitive decline are needed to identify those individuals at risk for AD and other neurodegenerative diseases. Plasma NfL may be one such non-invasive biomarker. Recent technological advances indicate that NfL levels measured in the blood, i.e., plasma NfL, are associated with AD diagnosis and with various cognitive, imaging, and biochemical disease measures [1, 15, 35]. CSF NfL also was inversely associated with the clinical dementia rating scale, the Recognition Memory Test [9], and the cognitive sub-scale of an AD assessment battery [10]. Several studies have indicated that CSF NfL is elevated in the early stages of dementia and is a strong predictor for cognitive decline in Aβ positive individuals [36, 37], and in the general non-demented older adult population [22]. Given that Aβ positivity alone was not sufficient to predict symptoms of cognitive decline in AD, identifying additional markers of neurodegeneration that are downstream from Aβ accumulation has high utility for screening individuals in pre-symptomatic trials [9].

Given the high correlation between plasma and CSF NfL levels, and the invasiveness of acquiring CSF, plasma NfL may have greater overall utility as a screening tool. Several recent studies have shown that plasma NfL may accurately predict the estimated year of onset for dementia [38, 39]. In fact, several recent studies have shown that serum or plasma NfL are direct indicators of axonal degeneration based on neuroimaging markers, including gray and white matter pathology [21, 40, 41], and can act as a proxy for hypometabolism in AD-vulnerable brain regions, particularly in Aβ-positive individuals [42]. Generally, the demyelination of axons triggers inefficiency in energy utilization, dysfunction of the mitochondria, and oxidative stress accumulation, alterations that increase axonal fragmentation and result in neurodegeneration [43]. The spread of such pathology can occur at independent tract locations and their associated gray matter structures [44]. Since such axonal retraction does not often occur simultaneously, it is more likely that baseline plasma NfL rather than follow-up or change in NfL, is associated with change or follow-up outcome of neurodegeneration, as well as adverse cognitive performance outcomes [40, 45]. This is in line with our main findings.

Among older adults, several studies have indicated that plasma NfL is a good predictor for cognitive decline or impairment, independently of neuroimaging markers. One recent study found that individuals with AD or fronto-temporal dementia cases had higher plasma NfL compared to cognitively normal controls, with no differences detected for other neuropsychiatric disorders [46]. Upon adjustment for baseline hippocampal atrophy and memory scores, plasma NfL predicted greater cognitive decline among the cognitively impaired [46]. Another study among older adults suggested that a combination of markers (low plasma Aβ42/Aβ40

ratio and high plasma NfL level) was associated with a greater decline in cognitive performance over time [20]. These findings were recently corroborated by Mielke and colleagues who examined both plasma and CSF NfL in relation to cognitive and neuroimaging outcomes in a small sample of older adults ($N = 79$, median age: 76 y) participating in the Alzheimer's Disease Neuroimaging (ADNI) study. Their findings indicate that elevated baseline plasma NfL may adequately predict cognitive decline and brain imaging neurodegenerative measures, with comparable effect sizes to baseline CSF NfL [21]. Furthermore, Rajan and colleagues found that 1,327 older participants, plasma NfL > 25.5 pg/ml (determined 4−8 y prior to AD onset) was associated with 110% faster cognitive decline over 16 y of follow-up, as well as a faster decline in cortical thickness [18]. Similarly, He and colleagues found that among 452 older adults, a combination of elevated Aβ and plasma NfL was associated with faster decline on the MMSE compared with lower levels, even upon adjustment for APOE4 status [20]. Moreover, Nyberg and colleagues found that plasma NfL, while reflecting white matter alteration, may not be a good predictor for cognitive impairment or impending AD [19]. Most recently, Rübsamen et. al. (2021) evaluated associations between NfL and tau serum levels, neuropsychological functioning, and brain structure among a sample of 385 adults aged 65+ years enrolled in the Memory and Morbidity in Augsburg Elderly study [16]. The authors used linear regression models adjusted for age, sex, years of education, and comorbidities and reported a cross-sectional association between NfL serum levels and neuropsychological functioning which included standardized cognitive tests spanning the domains of short-term memory, cognitive speed, attention, and motor speed [16]. Furthermore, in a study by Khalil and colleagues (2020), the authors examined age-related changes in NfL serum levels and their associations with brain structure and functioning [17]. In a sample of 335 men and women drawn from the prospective and ongoing Austrian Stroke Prevention Family Study, the authors used backwards stepwise regression while considering comorbidities and observed that individuals with elevated and more variable NfL serum levels tended to show accelerated rates of neuronal injury which may be attributed to subclinical comorbid pathologies [17]. Moreover, the authors reported that baseline NfL serum levels were negatively associated with annualized changes in scores obtained from the Mini-Mental State Examination [17]. Taken together, these studies may suggest associations between NfL levels and changes in brain volume which may, in turn, influence neuropsychological functioning.

Our data in middle-aged adults is in agreement with other studies among older adults, indicating the utility of blood-based NfL as a non-invasive biomarker of cognitive decline, which may allow for disease monitoring. Few studies have examined longitudinal change in blood levels of NfL. In one study of AD, longitudinal plasma NfL levels increased in individuals with several baseline AD-disease measures [10]. Here, we examined longitudinal changes in plasma NfL in non-demented middle-aged adults. Therefore, we were able to assess baseline and rates of change of NfL in relation to longitudinal cognitive test performance across race and other socio-demographic variables (sex,
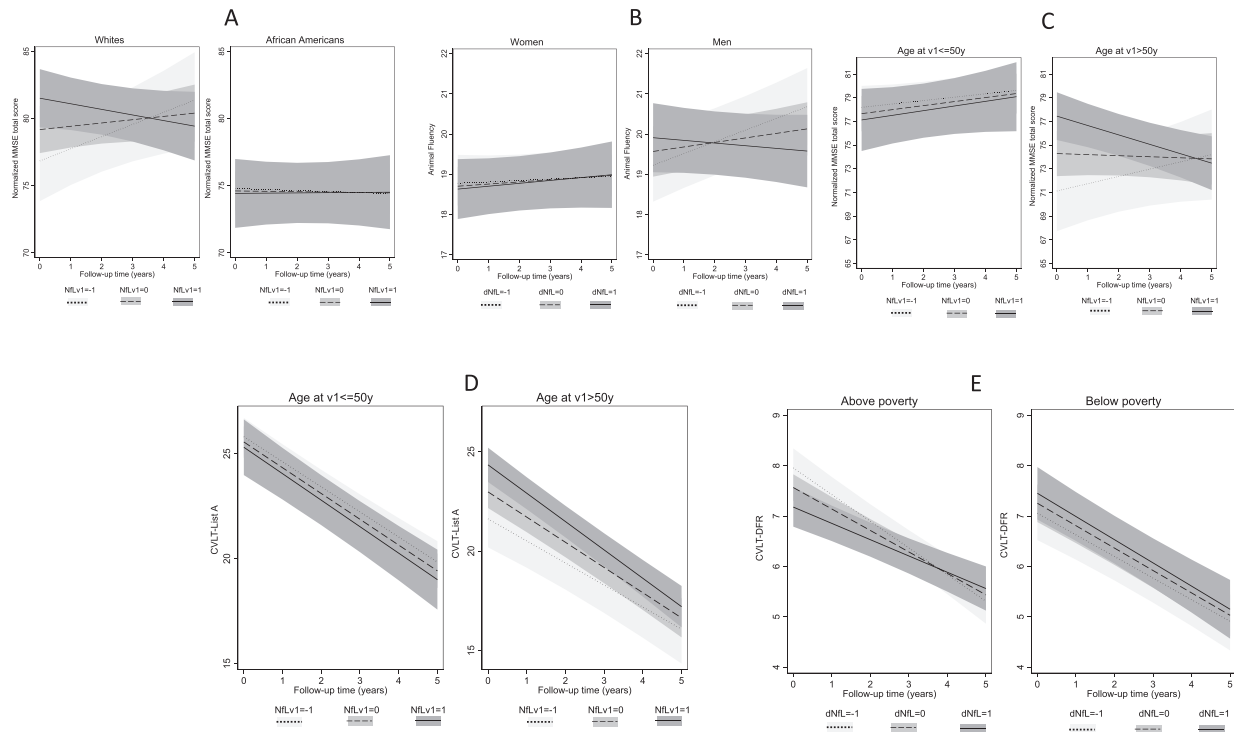
Fig. 2 Summary of key findings by race, sex, age group, and poverty status across NfL exposures[a,b]. [a]NfLv$_1$ values are Log$_e$ transformed and z-scored. Levels of exposure are −1: mean − 1 SD; 0: at mean; +1: mean + 1 SD. 1 SD of baseline Log$_e$(NfL) is estimated at 0.51; mean = 1.98. dNfL values are annualized changes in Log$_e$ transformed NfL between v$_1$ and v$_2$, z-scored. 1 SD of annualized change in Log$_e$(NfL) is estimated at 0.101; mean = 0.044. All test scores presented in these figures are coded in the direction of higher score → better performance. [b]**A** Predicted margins for normalized MMSE total score across NfLv1 are based on Model 1 among Whites and African Americans in Table 2; **B** predicted margins for animal fluency scores across dNfL are based on Model 2 among women and men in Table S1; **C** predicted margins for normalized MMSE total score across NfLv1 are based on Model 1 among ≤50 y vs. >50 y age groups in Table S2; **D** predicted margins for CVLT-List A across NfLv1 are based on Model 1 among ≤50 y vs. >50 y age groups in Table S2; **E** predicted margins for CVLT-DFR across dNfL are based on Model 2 among "above poverty" vs. "below poverty" groups in Table S3. Abbreviations: AF Animal Fluency; BC baseline cognitive performance; CVLT-DFR California Verbal Learning Test-Delayed Free Recall; CVLT-List A California Verbal Learning Test-List A; dNfL z-scores of annualized rates of change NfL, Loge transformed; NfL$_{v1}$ plasma NfL levels, Log$_e$ transformed, z-scored at v$_1$.

age group, and poverty status). This is important given the limited information about the longitudinal changes in plasma NfL, especially in non-diseased cohorts. These associations we found, highlight the underlying neurodegeneration that occurs over time and suggests that baseline plasma NfL levels in Whites and in individuals >50 y may be valuable to predict those individuals who will cognitively decline faster than others. The lack of association between NfL and cognitive decline among African Americans may be due to less variability in NfL and limited change in cognitive performance over time within this racial group, especially among middle-aged adults, as compared with Whites and therefore a reduced statistical power to detect such an association. Among Whites, the only other cognitive performance test that was suggestive of an association between first-visit NfL and cognitive decline over time was BTA, reflecting attention, though this relationship did not survive correction for multiple testing ($\gamma_{11} = -0.072 \pm 0.040$, $p < 0.10$, Model 1).

More generally, our study detected few associations between plasma NfL and cognitive decline compared with other studies, due to several possible reasons. First, our sample consisted of middle-aged adults, while most other studies were conducted among older adults aged over 60 y at baseline. This would result in a less steep decline in cognition over time in our sample compared to others of older mean age at baseline, which in turn would reduce the statistical power to detect an association between exposure and change in cognition over time, keeping exposure variability the same across samples. However, younger age also results in less variability in the plasma NfL exposures,

further reducing statistical power. Second, our sample consisted of a diverse group of middle-aged adults, whereas most other studies recruited middle to upper-middle-class White older adults. This difference in age group, racial, and SES composition is expected to yield diverging findings between our study and those of others, mainly due to differing baseline exposure and outcome levels. Finally, we have adjusted for a large number of potential confounders, including body mass index, and cardio-metabolic risk factors, some of which were shown to be associated with plasma NfL in previous studies [47, 48]. We also accounted for literacy, depressive symptoms, and other important factors that most other studies have not controlled for.

Strengths and limitations
Our study has several notable strengths. First, it is one of the largest longitudinal studies to examine plasma NfL levels in relation to cognition, using data from a community-based population, and the first to do so among middle-aged adults. In addition, plasma NfL was detected and quantified in non-demented individuals, which adds value to utilizing this biomarker as an early marker to monitor cognitive decline over time. Second, we had access to an extensive battery of cognitive tests that spanned the main domains of cognition, as well as measuring global mental status. Test scores had mostly two repeats, as did the main exposure of interest, plasma NfL. Third, the well-balanced sampling of HANDLS allowed for stratification of our analyses by race, sex, age group, and poverty status. Fourth, we used advanced statistical techniques, including mixed-effects

linear regression models, multiple imputations, and 2-stage Heckman selection to test our key hypotheses, while reducing confounding and selection biases. The availability of two concurrent repeats of exposures and outcomes, allowed us to examine relationships in a detailed and bi-directional manner, though mainly focusing on the potential impact of NfL on cognition, rather than the reverse direction. Nevertheless, our study also has some limitations. First, our study sample was relatively young with a low mean NfL at baseline, when compared to previous studies that examined these questions in older adults. In addition, cognitive decline was limited in that age group, and was only evident above the age of 50 y. This may have reduced our ability to detect an association between NfL at v_1 and change in cognitive function in the overall population. However, our results among Whites and the older group, suggest that NfL at v_1 may be a predictor of decline in global mental status in middle-age in those groups who have a high performance on the MMSE at baseline and are prone to decline over a period of ~5 y.

CONCLUSIONS
In summary, first-visit NfL was primarily associated with the global mental status decline among Whites, while exhibiting inconsistent relationships in some exploratory analyses. More comparable longitudinal studies are needed among middle-aged adults to determine the utility of plasma NfL both at baseline and as a marker of change over time in relationship to cognitive performance and decline.

DISCLAIMER
The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of Fort Belvoir Community Hospital, the Defense Health Agency, Department of Defense, or U.S. Government. Reference to any commercial products within this publication does not create or imply any endorsement by Fort Belvoir Community Hospital, theDefense Health Agency, Department of Defense, or U.S. Government.

REFERENCES
1. Zhao Y, Xin Y, Meng S, He Z, Hu W. Neurofilament light chain protein in neurodegenerative dementia: a systematic review and network meta-analysis. Neurosci Biobehav Rev. 2019;102:123–38. https://doi.org/10.1016/j.neubiorev.2019.04.014.
2. Kuhle J, Barro C, Andreasson U, Derfuss T, Lindberg R, Sandelius Å, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay, and Simoa. Clin Chem Lab Med. 2016;54:1655–61. https://doi.org/10.1515/cclm-2015-1195.
3. Raket LL, Kühnel L, Schmidt E, Blennow K, Zetterberg H, Mattsson-Carlgren N. Utility of plasma neurofilament light and total tau for clinical trials in Alzheimer's disease. Alzheimers Dement. 2020;12:e12099. https://doi.org/10.1002/dad2.12099.
4. Hoffman PN, Cleveland DW, Griffin JW, Landes PW, Cowan NJ, Price DL. Neurofilament gene expression: a major determinant of axonal caliber. Proc Natl Acad Sci USA. 1987;84:3472–6. https://doi.org/10.1073/pnas.84.10.3472.
5. Norgren N, Rosengren L, Stigbrand T. Elevated neurofilament levels in neurological diseases. Brain Res. 2003;987:25–31. https://doi.org/10.1016/s0006-8993(03)03219-0
6. Hansson O, Janelidze S, Hall S, Magdalinou N, Lees AJ, Andreasson U, et al. Blood-based NfL: a biomarker for differential diagnosis of parkinsonian disorder. Neurology. 2017;88:930–7. https://doi.org/10.1212/WNL.0000000000003680.
7. Preische O, Schultz SA, Apel A, Kuhle J, Kaeser SA, Barro C, et al. Serum neurofilament dynamics predicts neurodegeneration and clinical progression in pre-symptomatic Alzheimer's disease. Nat Med. 2019;25:277–83. https://doi.org/10.1038/s41591-018-0304-3.
8. de Wolf F, Ghanbari M, Licher S, McRae-McKee K, Gras L, Weverling GJ, et al. Plasma tau, neurofilament light chain, and amyloid-beta levels and risk of dementia; a population-based cohort study. Brain. 2020;143:1220–32. https://doi.org/10.1093/brain/awaa054.
9. Weston P, Poole T, O'Connor A, Heslegrave A, Ryan NS, Liang Y, et al. Longitudinal measurement of serum neurofilament light in presymptomatic familial Alzheimer's disease. Alzheimers Res Ther. 2019;11:19. https://doi.org/10.1186/s13195-019-0472-5.
10. Mattsson N, Cullen NC, Andreasson U, Zetterberg H, Blennow K. Association between longitudinal plasma neurofilament light and neurodegeneration in patients with alzheimer disease. JAMA Neurol. 2019;76:791–9. https://doi.org/10.1001/jamaneurol.2019.0765.
11. Scherling CS, Hall T, Berisha F, Klepac K, Karydas A, Coppola G, et al. Cerebrospinal fluid neurofilament concentration reflects disease severity in frontotemporal degeneration. Ann Neurol. 2014;75:116–26. https://doi.org/10.1002/ana.24052.
12. Teunissen CE, Dijkstra C, Polman C. Biological markers in CSF and blood for axonal degeneration in multiple sclerosis. Lancet Neurol. 2005;4:32–41. https://doi.org/10.1016/S1474-4422(04)00964-0.
13. Shahim P, Gren M, Liman V, Andreasson U, Norgren N, Tegner Y, et al. Serum neurofilament light protein predicts clinical outcome in traumatic brain injury. Sci Rep. 2016;6:36791. https://doi.org/10.1038/srep36791.
14. Khalil M, Teunissen CE, Otto M, Piehl F, Sormani MP, Gattringer T, et al. Neurofilaments as biomarkers in neurological disorders. Nat Rev Neurol. 2018;14:577–89. https://doi.org/10.1038/s41582-018-0058-z.
15. Mattsson N, Andreasson U, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging I. Association of plasma neurofilament light with neurodegeneration in patients with Alzheimer disease. JAMA Neurol. 2017;74:557–66. https://doi.org/10.1001/jamaneurol.2016.6117.
16. Rübsamen N, Maceski A, Leppert D, Benkert P, Kuhle J, Wiendl H, et al. Serum neurofilament light and tau as prognostic markers for all-cause mortality in the elderly general population-an analysis from the MEMO study. BMC Med. 2021;19:38. https://doi.org/10.1186/s12916-021-01915-8.
17. Khalil M, Pirpamer L, Hofer E, Voortman MM, Barro C, Leppert D, et al. Serum neurofilament light levels in normal aging and their association with morphologic brain changes. Nat Commun. 2020;11:812. https://doi.org/10.1038/s41467-020-14612-6.
18. Rajan KB, Aggarwal NT, McAninch EA, Weuve J, Barnes LL, Wilson RS, et al. Remote blood biomarkers of longitudinal cognitive outcomes in a population study. Ann Neurol. 2020;88:1065–76. https://doi.org/10.1002/ana.25874.
19. Nyberg L, Lundquist A, Nordin Adolfsson A, Andersson M, Zetterberg H, Blennow K, et al. Elevated plasma neurofilament light in aging reflects brain white-matter alterations but does not predict cognitive decline or Alzheimer's disease. Alzheimers Dement. 2020;12:e12050. https://doi.org/10.1002/dad2.12050.
20. He L, de Souto Barreto P, Aggarwal G, Nguyen AD, Morley JE, Li Y, et al. Plasma Abeta and neurofilament light chain are associated with cognitive and physical function decline in non-dementia older adults. Alzheimers Res Ther. 2020;12:128. https://doi.org/10.1186/s13195-020-00697-0.
21. Mielke MM, Syrjanen JA, Blennow K, Zetterberg H, Vemuri P, Skoog I, et al. Plasma and CSF neurofilament light: Relation to longitudinal neuroimaging and cognitive measures. Neurology. 2019;93:e252–e260. https://doi.org/10.1212/WNL.0000000000007767.
22. Merluzzi AP, Vogt NM, Norton D, Jonaitis E, Clark LR, Carlsson CM, et al. Differential effects of neurodegeneration biomarkers on subclinical cognitive decline. Alzheimers Dement . 2019;5:129–38. https://doi.org/10.1016/j.trci.2019.02.004.
23. Philipps V, Amieva H, Andrieu S, Dufouil C, Berr C, Dartigues JF, et al. Normalized mini-mental state examination for assessing cognitive change in population-based brain aging studies. Neuroepidemiology. 2014;43:15–25. https://doi.org/10.1159/000365637.
24. Bureau, UC US Census Bureau. Social, Economic, and Housing Statistics Division. Poverty Thresholds, 2004, 2014. (2014). http://www.census.gov/hhes/www/poverty/data/threshld/ thresh04.html. Accessed January 2016.
25. Beydoun MA, Hossain S, Beydoun HA, Shaked D, Weiss J, Evans MK, et al. Red cell distribution width is directly associated with poor cognitive performance among nonanemic, middle-Aged, Urban Adults. J Nutr. 2020;150:128–39. https://doi.org/10.1093/jn/nxz182
26. STATA. Statistics/Data Analysis: Release 16.0. Stata Corporation: Texas; 2019.
27. Beydoun, HA, Huang, S, Beydoun, MA, Hossain, S, Zonderman, AB. Mediating-moderating effect of allostatic load on the association between dietary approaches to stop hypertension diet and all-cause and cause-specific mortality: 2001−2010 National Health and Nutrition Examination Surveys. Nutrients. 2019. https://doi.org/10.3390/nu11102311.
28. Beydoun MA, Beydoun HA, Mode N, Dore GA, Canas JA, Eid SM, et al. Racial disparities in adult all-cause and cause-specific mortality among us adults: mediating and moderating factors. BMC Public Health. 2016;16:1113. https://doi.org/10.1186/s12889-016-3744-z.
29. Beydoun MA, Beydoun HA, Kitner-Triolo MH, Kaufman JS, Evans MK, Zonderman AB. Thyroid hormones are associated with cognitive function: moderation by sex, race, and depressive symptoms. J Clin Endocrinol Metab. 2013;98:3470–81. https://doi.org/10.1210/jc.2013-1813.

30. Selvin, S. Statistical analysis of epidemiologic data. 3rd edn. New York: Oxford University Press; 2004.

31. Hochberg, Y, Tamhane, AC. Multiple comparison procedures. New York: Wiley; 1987.

32. Beydoun MA, Canas JA, Dore GA, Beydoun HA, Rostant OS, Fanelli-Kuczmarski MT, et al. Serum uric acid and its association with longitudinal cognitive change among urban adults. J Alzheimers Dis. 2016;52:1415–30. https://doi.org/10.3233/JAD-160028.

33. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med. 1990;9:811–8. https://doi.org/10.1002/sim.4780090710.

34. Liang K. False discovery rate estimation for large-scale homogeneous discrete p-values. Biometrics. 2016;72:639–48. https://doi.org/10.1111/biom.12429.

35. Jin M, Cao L, Dai YP. Role of neurofilament light chain as a potential biomarker for Alzheimer's disease: a correlative meta-analysis. Front Aging Neurosci. 2019;11:254. https://doi.org/10.3389/fnagi.2019.00254.

36. Bos I, Vos S, Verhey F, Scheltens P, Teunissen C, Engelborghs S, et al. Cerebrospinal fluid biomarkers of neurodegeneration, synaptic integrity, and astroglial activation across the clinical Alzheimer's disease spectrum. Alzheimers Dement. 2019;15:644–54. https://doi.org/10.1016/j.jalz.2019.01.004.

37. Lewczuk P, Ermann N, Andreasson U, Schultheis C, Podhorna J, Spitzer P, et al. Plasma neurofilament light as a potential biomarker of neurodegeneration in Alzheimer's disease. Alzheimers Res Ther. 2018;10:71. https://doi.org/10.1186/s13195-018-0404-9.

38. Weston P, Poole T, Ryan NS, Nair A, Liang Y, Macpherson K, et al. Serum neurofilament light in familial Alzheimer disease: a marker of early neurodegeneration. Neurology. 2017;89:2167–75. https://doi.org/10.1212/WNL.0000000000004667.

39. Sánchez-Valle R, Heslegrave A, Foiani MS, Bosch B, Antonell A, Balasa M, et al. Serum neurofilament light levels correlate with severity measures and neurodegeneration markers in autosomal dominant Alzheimer's disease. Alzheimers Res Ther. 2018;10:113. https://doi.org/10.1186/s13195-018-0439-y.

40. Jakimovski D, Kuhle J, Ramanathan M, Barro C, Tomic D, Hagemeier J, et al. Serum neurofilament light chain levels associations with gray matter pathology: a 5-year longitudinal study. Ann Clin Transl Neurol. 2019;6:1757–70. https://doi.org/10.1002/acn3.50872.

41. Sun Y, Tan L, Xu W, Wang ZT, Hu H, Li JQ, et al. Plasma neurofilament light and longitudinal progression of white matter hyperintensity in elderly persons without dementia. J Alzheimers Dis. 2020;75:729–37. https://doi.org/10.3233/JAD-200022.

42. Benedet AL, Ashton NJ, Pascoal TA, Leuzy A, Mathotaarachchi S, Kang MS, et al. Plasma neurofilament light associates with Alzheimer's disease metabolic decline in amyloid-positive individuals. Alzheimers Dement. 2019;11:679–89. https://doi.org/10.1016/j.dadm.2019.08.002.

43. Fischer MT, Wimmer I, Höftberger R, Gerlach S, Haider L, Zrzavy T, et al. Disease-specific molecular events in cortical multiple sclerosis lesions. Brain. 2013;136:1799–815. https://doi.org/10.1093/brain/awt110.

44. Bergsland N, Tavazzi E, Laganà MM, Baglio F, Cecconi P, Viotti S, et al. White matter tract injury is associated with deep gray matter iron deposition in multiple sclerosis. J Neuroimaging. 2017;27:107–13. https://doi.org/10.1111/jon.12364.

45. Trapp BD, Peterson J, Ransohoff RM, Rudick R, Mörk S, Bö L. Axonal transection in the lesions of multiple sclerosis. N Engl J Med. 1998;338:278–85. https://doi.org/10.1056/NEJM199801293380502.

46. Barker W, Quinonez C, Greig MT, Behar R, Chirinos C, Rodriguez RA, et al. Utility of plasma neurofilament light in the 1Florida Alzheimer's Disease Research Center (ADRC). J Alzheimers Dis. 2021;79:59–70. https://doi.org/10.3233/JAD-200901.

47. Manouchehrinia A, Piehl F, Hillert J, Kuhle J, Alfredsson L, Olsson T, et al. Confounding effect of blood volume and body mass index on blood neurofilament light chain levels. Ann Clin Transl Neurol. 2020;7:139–43. https://doi.org/10.1002/acn3.50972.

48. Akamine S, Marutani N, Kanayama D, Gotoh S, Maruyama R, Yanagida K, et al. Renal function is associated with blood neurofilament light chain level in older adults. Sci Rep. 2020;10:20350. https://doi.org/10.1038/s41598-020-76990-7.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01563-9.

Correspondence and requests for materials should be addressed to M.A.B.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.