

SUPPLEMENTARY MATERIALS

Relatively Preserved Cognition Despite Elevated Adjusted Serum Neurofilament Light in U.S. Adults Aged 60–75 Years: A Survey-Weighted NHANES Analysis

Supplementary Methods, Tables, and Figures

This supplement contains 11 supplementary tables and 7 supplementary figures.

Contents

Supplementary Methods

Supplementary Results

Table S1. Multiple-imputation implementation, missingness, and diagnostic summary

Table S2. Cognitive-completion nonresponse assessment and inverse-probability-of-completion weighting

Table S3. Prespecified sensitivity analyses of the primary candidate-correlate models

Table S4. Alternative cognition definitions and robustness analyses

Table S5. Leave-one-PSU-out influence analyses

Table S6. Prespecified subgroup interaction feasibility

Table S7. Phenotype thresholds and NCHS reliability assessment

Table S8. Complete-case versus multiple-imputation adjusted-NfL associations

Table S9. Nonlinearity and education-interaction analyses

Table S10. Raw log2 NfL sensitivity analysis

Table S11. Post hoc sensitivity analyses of phenotype definition, vascular-metabolic adjustment, and physical-activity specification

Figure S1. Survey-weighted serum NfL distribution and adjusted-NfL diagnostics

Figure S2. Cognition and study-defined NfL-cognition group diagnostics

Figure S3. Survey-weighted adjusted-NfL spline curves across cognitive outcomes

Figure S4. Adjusted-NfL association with global cognition according to educational attainment

Figure S5. Cognitive-completion probabilities and leave-one-PSU-out influence

Figure S6. Alternative-cognition diagnostics

Figure S7. Robustness of the primary education and NfL-cognition findings across sensitivity analyses

Abbreviations

BMI, body mass index; CI, confidence interval; DSST, Digit Symbol Substitution Test; eGFR, estimated glomerular filtration rate; IPCW, inverse-probability-of-completion weighting; LLOQ, lower limit of quantification; MET, metabolic equivalent; MI, multiple imputation; NCHS, National Center for Health Statistics; NfL, neurofilament light chain; PCA, principal-component analysis; PIR, poverty-income ratio; PSU, primary sampling unit; RSE, relative standard error; SE, standard error; SMD, standardized mean difference.

Supplementary Methods

Complex-survey design and analytic domains

All analyses retained the dedicated 2-year serum NfL subsample weight (WTSSNH2Y), masked variance strata (SDMVSTRA), and primary sampling units (SDMVPSU). Taylor-series linearization was used for design-based variance estimation. Participant restrictions were implemented as survey domains rather than by reconstructing the survey design after exclusion. The primary analytic cohort included 502 adults aged 60-75 years with valid serum NfL, survey-design variables, and complete scores for the three prespecified cognitive domains; the final design contained 15 degrees of freedom.

Adjusted NfL, adjusted cognition, and better-than-expected cognition

The adjusted-NfL reference model was fitted among 2,071 positive-weight participants aged 20-75 years and modeled log2 serum NfL as a function of natural cubic splines for age and eGFR, each with 3 degrees of freedom, together with sex. Observed-minus-predicted residuals were standardized using the survey-weighted residual standard deviation among participants aged 60-75 years. Global cognition was constructed from equally weighted memory, Animal Fluency, and DSST z scores and was adjusted for age, sex, and cognitive-test language. The standardized residual from a separate model of global cognition on adjusted NfL, age, sex, eGFR, and test language was termed the better-than-expected cognition score; positive values indicated cognitive performance higher than predicted by the modeled variables.

Multiple imputation and pooled inference

Missing covariate data were handled using 20 imputations with chained equations. Nine variables were actively imputed. Predictive mean matching was used for continuous variables, polytomous logistic regression for unordered categorical variables, and bootstrapped logistic regression for binary variables. Serum NfL, adjusted NfL, cognitive outcomes, study-group indicators, survey weights, strata, and PSUs were not imputed. Survey models were fitted within each completed dataset and pooled using Rubin rules with Barnard-Rubin small-sample degrees of freedom. Table S1 summarizes missingness and the principal imputation diagnostics.

Cognitive-completion nonresponse analysis

Among the 542 age-eligible participants with valid NfL data, the probability of completing the three-domain cognitive composite was modeled using age, sex, binary race/ethnicity, education, eGFR, adjusted NfL, stroke, and diabetes. Stabilized inverse-probability-of-completion weights were truncated at the 1st and 99th percentiles and multiplied by WTSSNH2Y. The primary education, PIR, and physical-activity models were then repeated using the combined weights. Survey-weighted standardized differences were used descriptively to characterize differences between cognitively complete and incomplete participants (Table S2).

Prespecified robustness analyses

Thirteen prespecified sensitivity scenarios evaluated alternative NfL and cognition thresholds, exclusions for stroke or reduced kidney function, assay-limit and extreme-NfL restrictions, and alternative adjustment of the NfL and cognition models (Table S3). Alternative cognition analyses used a standardized summed raw-score composite, the first component from a survey-weighted principal-component analysis, and a failed-practice-pretest-zero definition in which only explicit Animal Fluency or DSST practice-pretest failures were assigned zero (Table S4). Leave-one-PSU-out analyses repeated each primary candidate-correlate model after sequential deletion of each observed PSU (Table S5). Prespecified subgroup interactions were retained for interpretation only when the model was estimable and the pooled denominator degrees of freedom were at least 4 (Table S6).

Post hoc sensitivity analyses

Three post hoc sensitivity analyses were conducted to evaluate the robustness of the principal findings. First, the elevated-adjusted-NfL threshold was retained at the survey-weighted 75th percentile, low cognition remained defined at or below the weighted 25th percentile, and clearly preserved cognition was defined at or above the survey-weighted median; participants between the 25th and 50th percentiles were excluded from this contrast. Second, diabetes and hypertension were added to the primary education model. Third, the continuous physical-activity analysis was repeated using any versus no recreational activity. All post hoc sensitivity models were pooled across the 20 imputed datasets (Table S11).

Reliability and interpretation

Study-defined group proportions were screened using programmed NCHS presentation criteria incorporating unweighted event counts, effective denominators, RSE, confidence-interval width, and design degrees of freedom (Table S7).

Robustness analyses were interpreted according to direction, precision, and stability across analyses rather than by statistical significance alone. Numerically unstable level-specific estimates arising from quasi-complete separation were identified explicitly and were not interpreted according to magnitude.

Supplementary Results

The multiple-imputation procedure showed stable performance: all primary models were fitted across 20 imputed datasets, no primary-procedure logged events occurred, the maximum Monte Carlo error-to-total-SE ratio for primary exposure coefficients was 0.045, and the minimum relative efficiency was 0.998 (Table S1). Cognitive completion was high in weighted analyses. Although noncompletion was socially patterned, IPCW estimates closely reproduced the primary education, PIR, and physical-activity results (Table S2 and Figure S5).

Across the 13 prespecified sensitivity scenarios, the direction of the education association was generally consistent. Education showed evidence of association in 10 of 12 stable scenarios; the weighted 20th-percentile cognition definition and the combined NfL-80th/cognition-20th definition attenuated the omnibus education result, while the fixed minus-1-SD cognition definition produced quasi-complete separation and was not interpreted according to its level-specific magnitude (Table S3 and Figure S7A). Alternative cognition definitions reproduced the primary pattern: education remained associated with relatively preserved-cognition discordance, PIR and physical activity showed no evidence of association in the categorical models, and adjusted NfL remained inversely associated with all three alternative global cognition measures (Table S4 and Figures S6-S7B).

The education omnibus association remained evident after each of 29 leave-one-PSU-out deletions, whereas PIR and physical activity did not show evidence of association across the influence analyses (Table S5 and Figure S5). None of the prespecified subgroup interaction analyses met the predefined pooled-degree-of-freedom criterion for stable interpretation (Table S6). All four study-defined group proportions met the programmed NCHS presentation standards (Table S7). Complete-case and multiple-imputation adjusted-NfL coefficients were closely aligned (Table S8), and the nonlinearity and education-interaction analyses did not provide evidence requiring departure from the primary linear, no-interaction specification (Table S9 and Figures S3-S4). Raw log2 NfL sensitivity estimates remained inverse and were consistent with the primary adjusted-NfL results (Table S10).

In post hoc sensitivity analyses, education remained associated with the stricter preserved-cognition phenotype (omnibus $p=0.009$) and after additional adjustment for diabetes and hypertension (omnibus $p=0.023$). When recreational physical activity was modeled as any versus none, any activity was associated with a 0.467-SD higher better-than-expected cognition score (95% CI, 0.249 to 0.686; $p<0.001$) (Table S11).

Table S1. Multiple-imputation implementation, missingness, and diagnostic summary**Panel A. Overall implementation and performance**

Metric	Value
Number of imputations	20
Chained-equation iterations	20
Variables actively imputed	9
Primary-procedure logged events	0
Primary model failures	0
Maximum Monte Carlo error / total SE	0.045
Minimum relative efficiency	0.998

Panel B. Variables actively imputed

Variable	Missing, n (%)	Imputation method
Educational attainment	1 (0.2%)	Polytomous logistic regression
Poverty-income ratio	38 (7.6%)	Predictive mean matching
PHQ-9 score	7 (1.4%)	Predictive mean matching
Sleep duration	2 (0.4%)	Predictive mean matching
Diabetes	1 (0.2%)	Bootstrapped logistic regression
Hypertension	7 (1.4%)	Bootstrapped logistic regression
Anemia	2 (0.4%)	Bootstrapped logistic regression
Body mass index	5 (1.0%)	Predictive mean matching
Smoking status	1 (0.2%)	Polytomous logistic regression

Serum NfL, adjusted NfL, cognitive outcomes, study-group indicators, survey weights, strata, and PSUs were not imputed. The diagnostic values shown are sufficient to document stability of the primary imputation procedure without reproducing internal model-audit logs.

Table S2. Cognitive-completion nonresponse assessment and inverse-probability-of-completion weighting**Panel A. Completion and IPCW summary**

Metric	Value
Age-eligible NfL participants	542
Cognitively complete participants	502
Cognitively incomplete participants	40
Weighted marginal completion probability	0.954
Truncated IPCW range	0.960 to 1.356

Panel B. Primary omnibus tests before and after IPCW

Exposure	Standard model	IPCW model
Education	F=7.02; p=0.010	F=6.80; p=0.011
PIR spline	F=0.91; p=0.484	F=0.88; p=0.495
Physical activity	F=2.26; p=0.161	F=2.26; p=0.161

Panel C. Selected survey-weighted differences between cognitively complete and incomplete participants

Variable	Comparison	Complete	Incomplete	SMD
Poverty-income ratio	Continuous	3.136	1.348	1.380
Education	Less than high school	0.153	0.535	-0.876
College graduate indicator	Indicator = 1	0.290	0.030	0.759
Race/ethnicity	Non-Hispanic White	0.777	0.463	0.684
Current smoking	Indicator = 1	0.118	0.316	-0.496
Diabetes	Indicator = 1	0.231	0.148	0.214
Hypertension	Indicator = 1	0.751	0.829	-0.194
Self-reported stroke	Indicator = 1	0.063	0.104	-0.149
Female sex	Indicator = 1	0.527	0.588	-0.124
Age	Continuous	66.349	66.403	-0.012

Continuous entries are weighted means; indicator and categorical entries are weighted proportions. Standardized differences are descriptive and were used to characterize noncompletion rather than as hypothesis tests. IPCW, inverse-probability-of-completion weighting; SMD, standardized mean difference.

Table S3. Prespecified sensitivity analyses of the primary candidate-correlate models

Scenario	n	Concordant impairment n	Relatively preserved-cognition discordance n	Education p	PIR p	Activity p	College graduate OR (95% CI)	Interpretation
Primary Q75/Q25 phenotype reproduced	127	57	70	0.010	0.519	0.161	28.04 (2.62 to 300.27)	Evidence of association
Elevated NfL defined by weighted top tertile	169	67	102	0.003	0.209	0.204	13.81 (2.62 to 72.74)	Evidence of association
Elevated NfL defined by weighted 80th percentile	98	45	53	0.044	0.528	0.161	21.05 (1.75 to 253.62)	Evidence of association
Low cognition defined by weighted 20th percentile	127	47	80	0.065	0.485	0.133	26.61 (1.44 to 491.05)	Direction similar; omnibus p ≥ 0.05
Low cognition defined as adjusted cognition z ≤ -1	127	36	91	<0.001	0.470	0.015	Unstable	Quasi-complete separation; level-specific magnitude not interpreted
Combined NfL 80th and cognition 20th percentiles	98	37	61	0.104	0.456	0.161	20.98 (1.04 to 422.91)	Direction similar; omnibus p ≥ 0.05
Primary phenotype after excluding self-reported stroke	121	53	68	0.018	0.468	0.166	22.16 (2.24 to 219.04)	Evidence of association
Primary phenotype restricted to eGFR ≥ 30	124	54	70	0.025	0.700	0.179	20.91 (1.82 to 240.72)	Evidence of association
Primary phenotype restricted to eGFR ≥ 45	123	54	69	0.024	0.702	0.176	21.17 (1.83 to 244.61)	Evidence of association
Primary phenotype excluding NfL below the LLOQ	127	57	70	0.010	0.519	0.161	28.04 (2.62 to 300.27)	Evidence of association
Primary phenotype excluding raw NfL above weighted P99	123	55	68	0.036	0.601	0.200	22.01 (1.97 to 246.42)	Evidence of association
Phenotype using cognition residual additionally adjusted for race	127	48	79	0.004	0.207	0.224	68.31 (3.95 to 1181.78)	Evidence of association
Phenotype using NfL reference model additionally adjusted for BMI	127	57	70	0.010	0.699	0.145	23.16 (2.22 to 241.53)	Evidence of association

Outcome counts refer to the elevated-NfL comparison. Overall p values are Rubin-pooled covariance Wald tests from separate education, PIR-spline, and physical-activity models adjusted for binary race/ethnicity and self-reported stroke, except that stroke adjustment was omitted in the stroke-exclusion scenario. The fixed adjusted-cognition ≤ -1 -SD scenario produced quasi-complete separation and its level-specific magnitude was not interpreted.

Table S4. Alternative cognition definitions and robustness analyses**Panel A. Definition, cohort size, and concordance with the primary cognition construction**

Definition	Analytic n	Global-score correlation	Adjusted-residual correlation	Additional information
Raw-score cognitive composite	502	0.967	0.964	Summed raw-score composite; survey-weight standardized before adjustment
Survey-weighted first principal-component cognition	502	0.999	0.999	PC1 explained 64.8% of weighted variance
Failed Animal Fluency/DSST practice pretests assigned zero	516	1.000	0.999	14 participants added: 1 Animal Fluency and 13 DSST practice failures

Panel B. Omnibus p values for categorical relatively preserved-cognition discordance models

Cognition definition	Education p	PIR p	Physical activity p
Raw-score cognitive composite	0.009	0.234	0.320
Survey-weighted PCA cognition	0.010	0.502	0.154
Failed-pretest-zero	0.014	0.388	0.134

Panel C. Adjusted-NfL association with alternative global cognition definitions

Cognition definition	Beta (95% CI)	p value	Pooled df
Raw-score cognitive composite	-0.108 (-0.194 to -0.023)	0.023	4.667
Survey-weighted PCA cognition	-0.117 (-0.205 to -0.029)	0.020	4.667
Failed-pretest-zero	-0.112 (-0.202 to -0.021)	0.025	4.667

Categorical models were adjusted for binary race/ethnicity and self-reported stroke and pooled across 20 imputations. Adjusted-NfL models used the same covariate structure as the primary biomarker-cognition analysis. PCA, principal-component analysis; PIR, poverty-income ratio.

Table S5. Leave-one-PSU-out influence analyses

Exposure	Baseline	Leave-one-PSU-out F range	Leave-one-PSU-out p range	Successful deletions
Education	F=7.02; p=0.010	4.87 to 9.94	0.004 to 0.033	29
PIR spline	F=0.87; p=0.500	0.61 to 2.09	0.203 to 0.632	29
Physical activity	F=2.26; p=0.161	0.69 to 2.91	0.119 to 0.426	29

All 29 leave-one-PSU-out deletions were fitted successfully. Education retained evidence of association after every deletion; PIR and physical activity did not show evidence of association across the influence analysis. PIR, poverty-income ratio; PSU, primary sampling unit.

Table S6. Prespecified subgroup interaction feasibility

Subgroup	Comparison	Interaction beta (95% CI)	p value	Pooled df	Reporting status
Sex	Male vs Female	0.107 (-0.131 to 0.345)	0.274	3.750	Not interpreted: pooled df <4
Age group	Age 60-67 vs Age 68-75	-0.070 (-0.318 to 0.178)	0.468	3.750	Not interpreted: pooled df <4
Diabetes status	No vs Yes	0.087 (-0.160 to 0.334)	0.337	2.857	Not interpreted: pooled df <4
Obesity status	BMI <30 kg/m ² vs BMI ≥30 kg/m ²	0.118 (-0.189 to 0.426)	0.301	2.852	Not interpreted: pooled df <4
Kidney-function group	eGFR ≥60 vs eGFR <60	-0.247 (-0.665 to 0.170)	0.171	3.750	Not interpreted: pooled df <4
Self-reported stroke history	No vs Yes	—	—	—	Not modelable under prespecified rule

The interaction beta represents the difference in the adjusted-NfL slope for the comparison level relative to the reference level. Stable detailed reporting required finite pooled quantities and at least 4 pooled denominator degrees of freedom; none of the prespecified subgroup analyses met all criteria.

Table S7. Phenotype thresholds and NCHS reliability assessment**Panel A. Primary and sensitivity thresholds**

Construct	Threshold	Rule	Value	Role
Adjusted NfL	Top tertile sensitivity	Survey-weighted 66.67th percentile	0.189	Sensitivity
Adjusted NfL	Primary elevated-NfL threshold	Survey-weighted 75th percentile	0.467	Primary
Adjusted NfL	80th-percentile sensitivity	Survey-weighted 80th percentile	0.632	Sensitivity
Adjusted cognition	20th-percentile sensitivity	Survey-weighted 20th percentile	-0.757	Sensitivity
Adjusted cognition	Primary low-cognition threshold	Survey-weighted 25th percentile	-0.621	Primary
Adjusted cognition	Minus-1-SD sensitivity	Fixed standardized value	-1.000	Sensitivity

Panel B. NCHS reliability assessment for study-defined group proportions

Study-defined group	Unweighted n	Weighted proportion	Korn-Graubard 95% CI	RSE	Presentation status
Concordant reference	257	58.4%	49.2% to 67.2%	7.0%	Meets criteria
Low-cognition discordance	118	16.4%	12.3% to 21.2%	12.3%	Meets criteria
Relatively preserved-cognition discordance	70	16.3%	12.0% to 21.4%	13.1%	Meets criteria
Concordant impairment	57	8.9%	5.5% to 13.6%	20.2%	Meets criteria

The programmed NCHS screen incorporated unweighted event counts, effective denominators, relative standard errors, confidence-interval widths, and design degrees of freedom. All four study-defined group proportions met the presentation criteria. NCHS, National Center for Health Statistics; RSE, relative standard error.

Table S8. Complete-case versus multiple-imputation adjusted-NfL associations

Analysis	Outcome	Beta (95% CI)	p value	Degrees of freedom
Complete case	Global cognition	-0.117 (-0.199 to -0.034)	0.014	6.000
Multiple imputation	Global cognition	-0.117 (-0.206 to -0.028)	0.020	4.667
Complete case	Memory	-0.082 (-0.170 to 0.006)	0.062	6.000
Multiple imputation	Memory	-0.082 (-0.177 to 0.012)	0.074	4.667
Complete case	Animal Fluency	-0.107 (-0.214 to -0.000)	0.050	6.000
Multiple imputation	Animal Fluency	-0.107 (-0.222 to 0.007)	0.061	4.667
Complete case	DSST	-0.086 (-0.166 to -0.007)	0.038	6.000
Multiple imputation	DSST	-0.086 (-0.172 to -0.001)	0.049	4.667

Beta coefficients represent the change in cognitive outcome z score per 1-SD increase in adjusted NfL. Models adjusted for age, sex, binary race/ethnicity, college-graduate status, cognitive-test language, eGFR, and self-reported stroke. Multiple-imputation estimates were pooled across 20 completed datasets.

Table S9. Nonlinearity and education-interaction analyses**Panel A. Design-based nonlinear-component tests comparing spline and linear adjusted-NfL specifications**

Outcome	F statistic	df	p value
Global cognition z-score	5.74	2, 4	0.067
Memory-domain z-score	2.35	2, 4	0.211
Animal Fluency z-score	6.28	2, 4	0.058
DSST z-score	1.85	2, 4	0.270

Panel B. Adjusted-NfL-by-college interaction for global cognition

Analysis	Interaction beta (95% CI)	p value	Degrees of freedom
Complete case	0.031 (-0.173 to 0.235)	0.709	5.000
Multiple imputation	0.031 (-0.195 to 0.258)	0.714	3.750

The nonlinear-component tests evaluate whether the spline specification improved on the corresponding linear model. The interaction beta is the difference in the adjusted-NfL slope for college graduates versus participants with lower educational attainment. The results did not provide evidence requiring departure from the primary linear, no-interaction specification.

Table S10. Raw log2 NfL sensitivity analysis

Outcome	Beta (95% CI)	p value	Pooled df
Global cognition	-0.184 (-0.309 to -0.058)	0.014	4.667
Memory	-0.132 (-0.261 to -0.003)	0.047	4.667
Animal Fluency	-0.164 (-0.331 to 0.004)	0.054	4.667
DSST	-0.139 (-0.254 to -0.024)	0.027	4.667

Beta coefficients represent the change in cognitive outcome z score per doubling of measured serum NfL. Models adjusted directly for age, sex, eGFR, binary race/ethnicity, college-graduate status, cognitive-test language, and self-reported stroke and were pooled across 20 imputations.

Table S11. Post hoc sensitivity analyses of phenotype definition, vascular-metabolic adjustment, and physical-activity specification

Panel A. Stricter preserved-cognition definition

Exposure	Contrast	OR (95% CI)	Coefficient p	Omnibus F (df1, df2)	Overall p
Education	Overall 4-level test			7.17 (3, 9.00)	0.009
Education	Less than high school	1.00 (reference)	—	—	—
Education	High school or GED vs less than high school	0.60 (0.12 to 3.08)	0.499	—	—
Education	Some college or associate degree vs less than high school	5.65 (0.93 to 34.33)	0.058	—	—
Education	College graduate or above vs less than high school	33.00 (2.63 to 414.03)	0.012	—	—
Poverty-income ratio	Natural-spline overall test			0.94 (3, 7.21)	0.470
Recreational physical activity	Per 500 MET-min/week	1.25 (0.84 to 1.88)	0.244	1.51 (1, 11.00)	0.244

Panel B. Limited vascular-metabolic adjustment of the education model

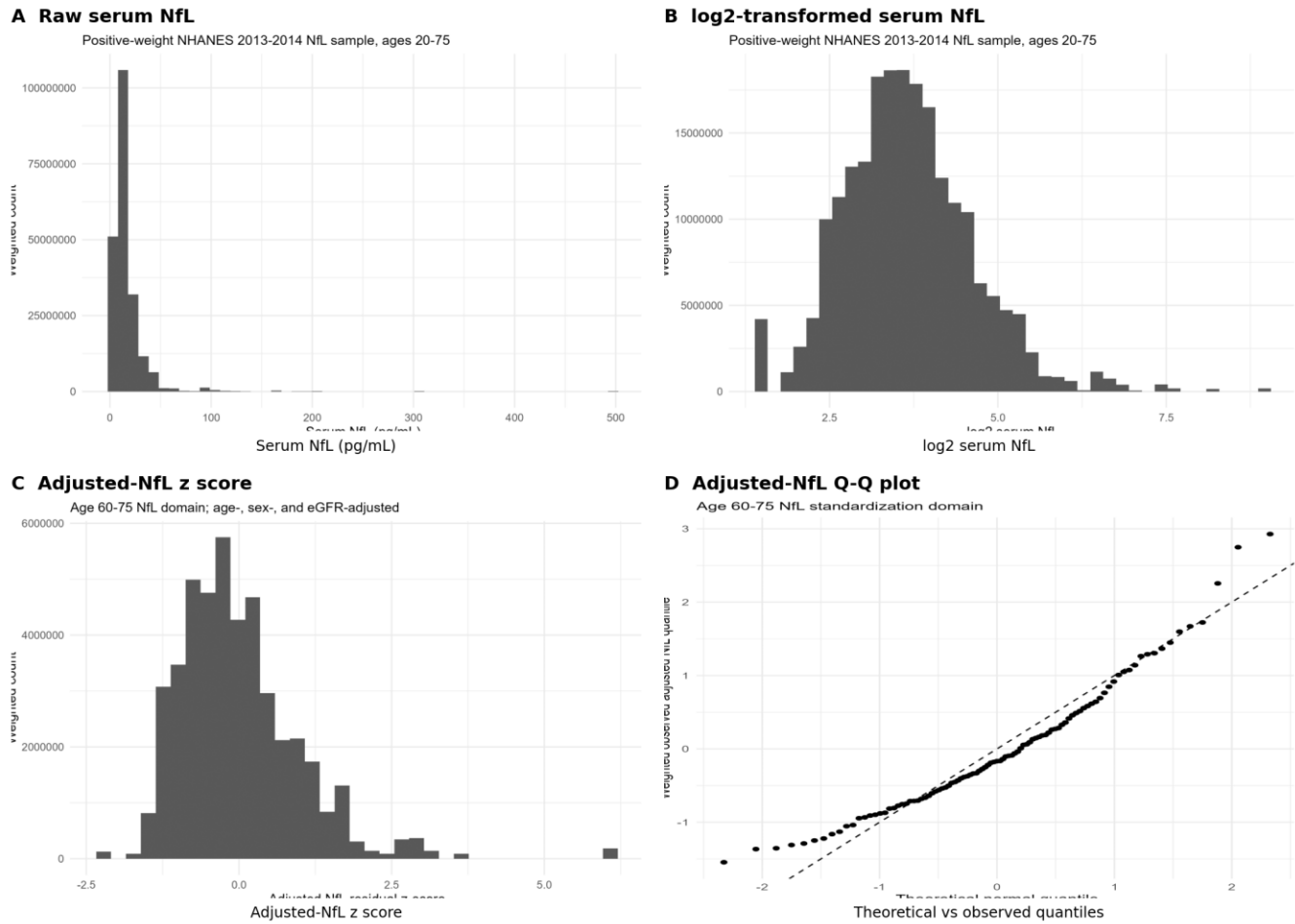
Exposure	Contrast	OR (95% CI)	Coefficient p	Omnibus F (df1, df2)	Overall p
Education	Overall 4-level test			7.32 (3, 5.60)	0.023
Education	Less than high school	1.00 (reference)	—	—	—
Education	High school or GED vs less than high school	1.61 (0.37 to 6.92)	0.449	—	—
Education	Some college or associate degree vs less than high school	7.65 (1.70 to 34.47)	0.017	—	—
Education	College graduate or above vs less than high school	31.40 (2.02 to 488.21)	0.022	—	—

Panel C. Alternative physical-activity representation

Exposure	Contrast	Beta (95% CI)	Coefficient p	Omnibus F (df1, df2)	Overall p
Recreational physical activity	Any recreational activity vs none	0.467 (0.249 to 0.686)	<0.001	21.69 (1, 12.00)	<0.001

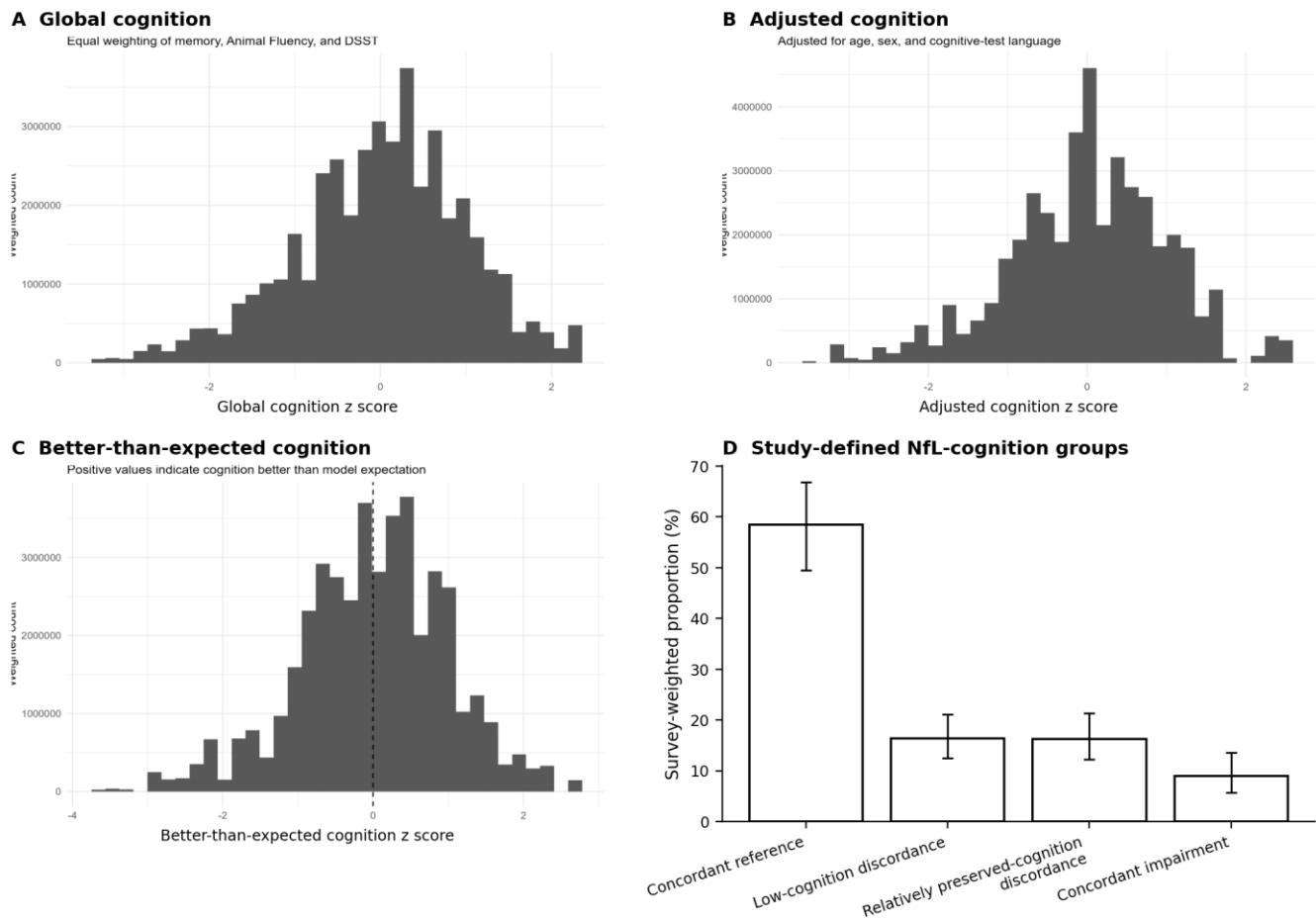
The stricter phenotype retained elevated adjusted NfL at the survey-weighted 75th percentile and low cognition at or below the weighted 25th percentile; clearly preserved cognition was defined as adjusted cognition at or above the weighted median, and participants between the 25th and 50th percentiles were excluded. The vascular-metabolic model additionally adjusted for diabetes and hypertension. All analyses were pooled across 20 imputations. These analyses were post hoc and were not part of the prespecified primary analysis. CI, confidence interval; MET, metabolic equivalent; OR, odds ratio; PIR, poverty-income ratio.

Figure S1. Survey-weighted serum NfL distribution and adjusted-NfL diagnostics



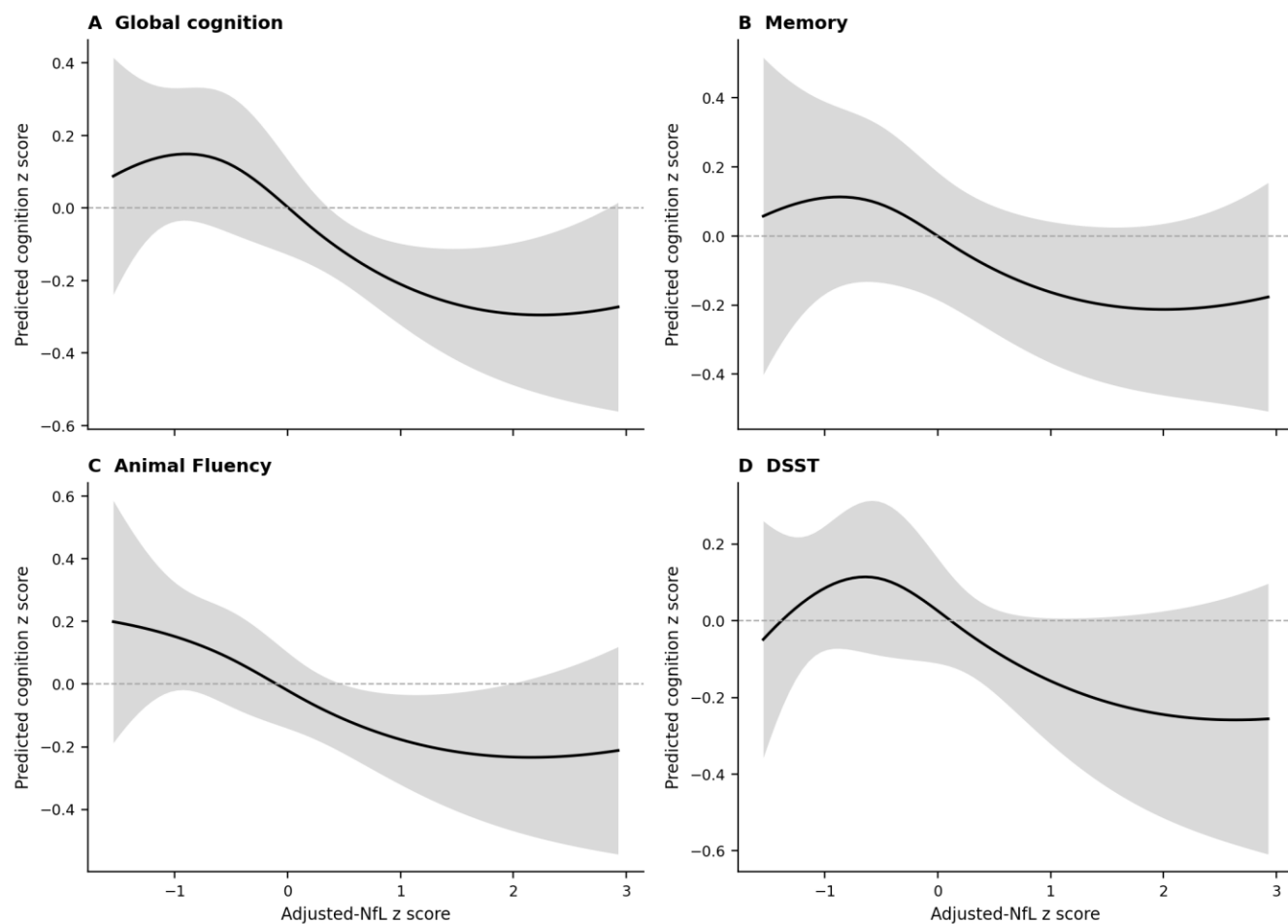
Panel A shows the survey-weighted raw serum NfL distribution; Panel B shows the log2-transformed distribution; Panel C shows the standardized adjusted-NfL distribution; and Panel D shows the weighted quantile-quantile assessment. The adjusted-NfL model accounted for age, sex, and eGFR. eGFR, estimated glomerular filtration rate; NfL, neurofilament light chain.

Figure S2. Cognition and study-defined NfL-cognition group diagnostics



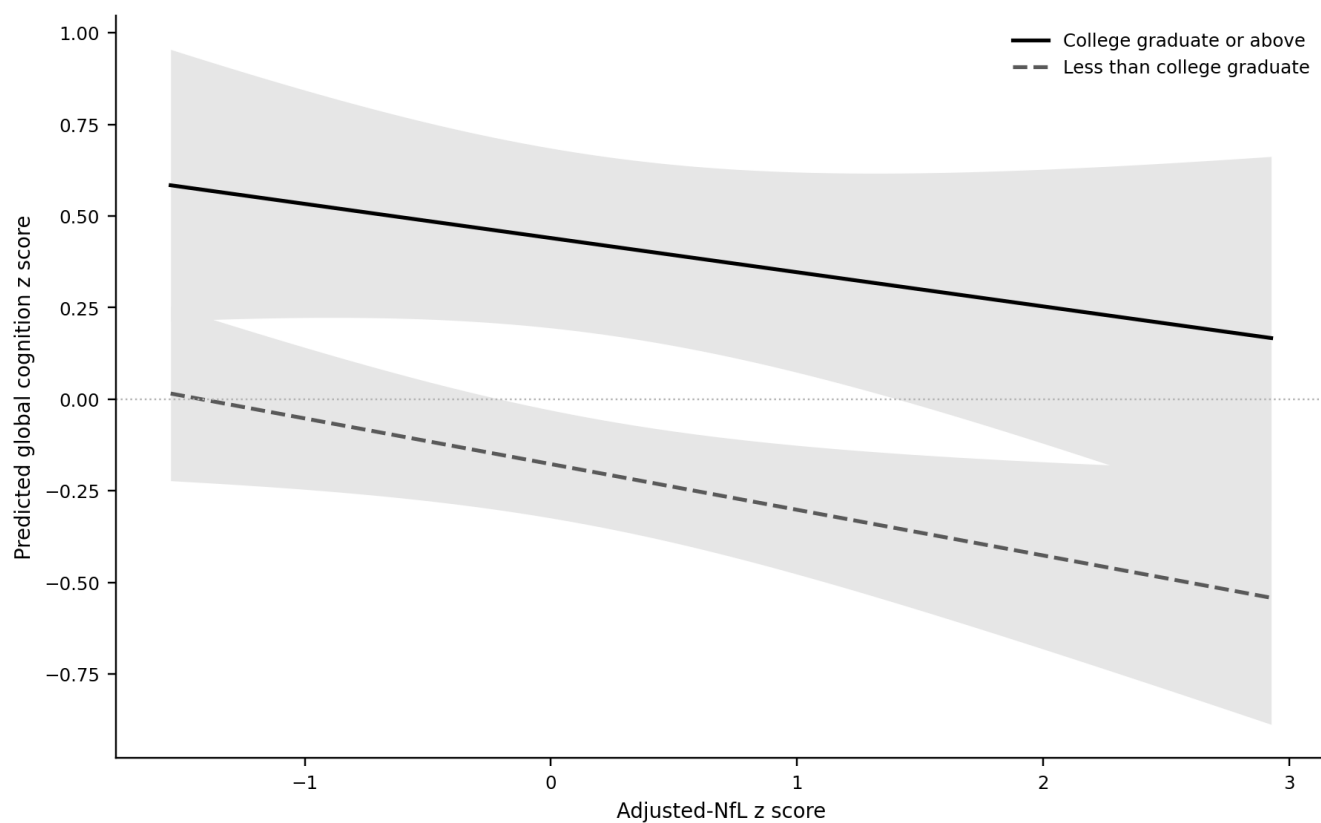
Panels A-C show the survey-weighted distributions of global cognition, adjusted cognition, and the better-than-expected cognition score. Panel D shows the survey-weighted proportions of the four study-defined NfL-cognition groups with 95% confidence intervals. The study-defined thresholds were the weighted 75th percentile of adjusted NfL and the weighted 25th percentile of adjusted cognition.

Figure S3. Survey-weighted adjusted-NfL spline curves across cognitive outcomes



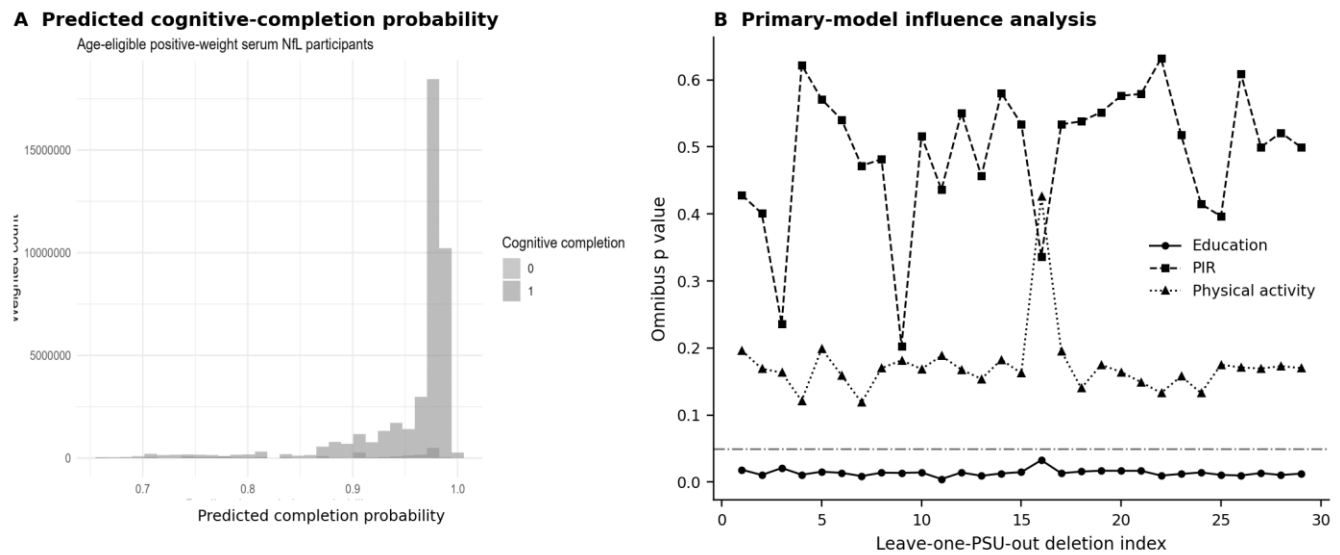
Survey-weighted natural-spline associations are shown for global cognition (Panel A), memory (Panel B), Animal Fluency (Panel C), and DSST (Panel D) over the central range of adjusted NfL. Curves are adjusted predictions with 95% confidence bands. These analyses were used to assess potential departure from linearity. DSST, Digit Symbol Substitution Test; NfL, neurofilament light chain.

Figure S4. Adjusted-NfL association with global cognition according to educational attainment



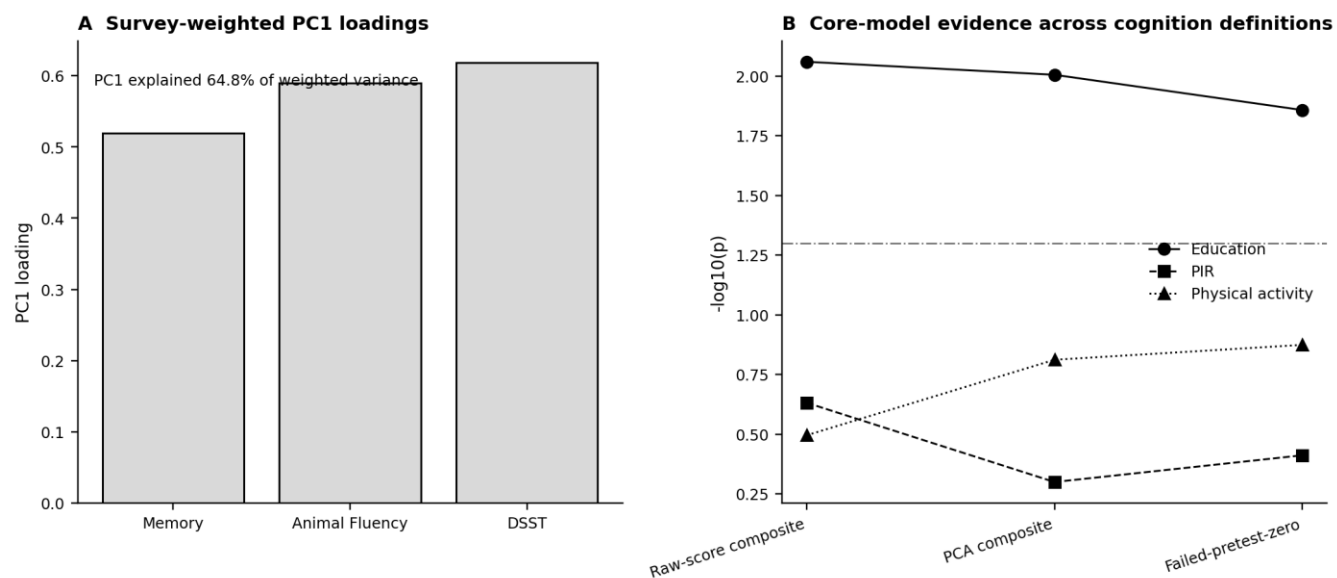
Predicted global-cognition z scores are shown across adjusted NfL for college graduates and participants with lower educational attainment. The complete-case adjusted-NfL-by-college interaction was $\beta=0.031$ ($p=0.709$), and the multiple-imputation pooled interaction was $\beta=0.031$ (95% CI, -0.195 to 0.258; $p=0.714$), providing no evidence that college education modified the adjusted-NfL-cognition association.

Figure S5. Cognitive-completion probabilities and leave-one-PSU-out influence



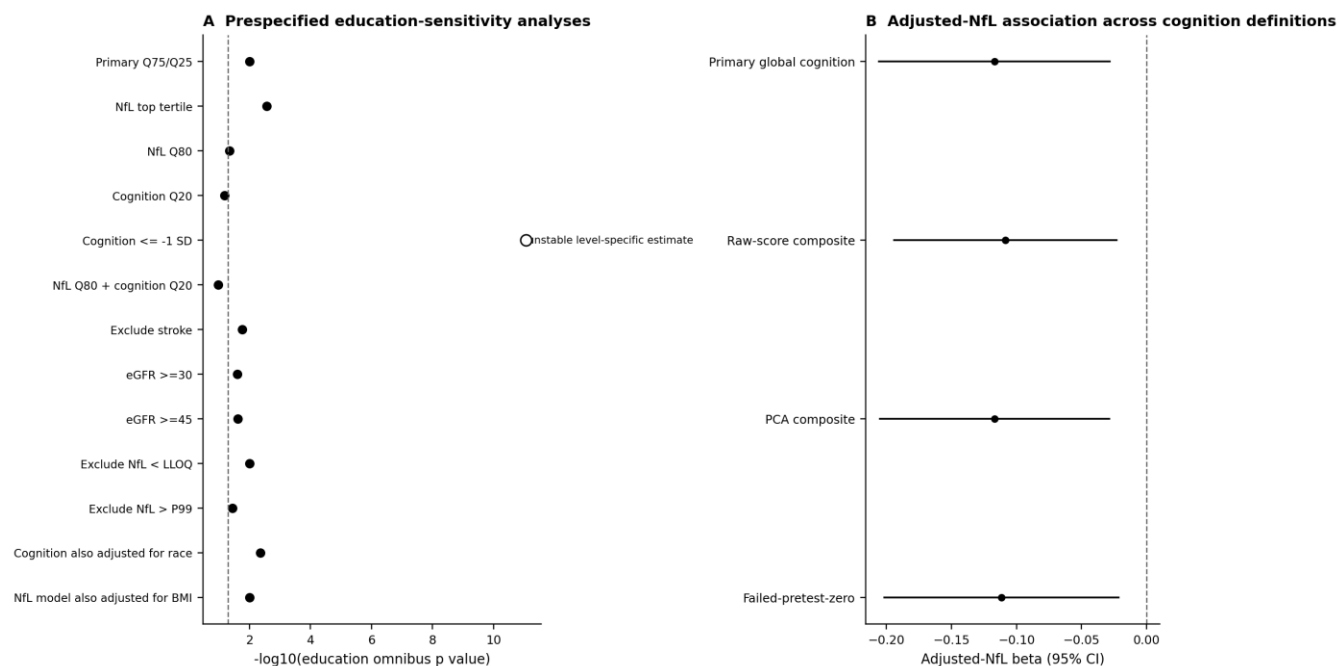
Panel A displays predicted probabilities from the cognitive-completion model among age-eligible participants with valid NFL. Panel B shows omnibus p values for the education, PIR, and physical-activity models after sequential deletion of each of 29 PSUs; the reference line marks $p=0.05$. Education retained evidence of association after every deletion, whereas PIR and physical activity did not show evidence of association. PIR, poverty-income ratio; PSU, primary sampling unit.

Figure S6. Alternative-cognition diagnostics



Panel A shows the survey-weighted first-principal-component loadings for memory, Animal Fluency, and DSST; PC1 explained 64.8% of weighted variance. Panel B summarizes omnibus evidence for education, PIR, and physical activity across the raw-score, PCA, and failed-pretest-zero cognition definitions; the horizontal reference corresponds to $p=0.05$. DSST, Digit Symbol Substitution Test; PCA, principal-component analysis; PIR, poverty-income ratio.

Figure S7. Robustness of the primary education and NfL-cognition findings across sensitivity analyses



Panel A summarizes the education omnibus p values across the 13 prespecified sensitivity scenarios. The open marker denotes the fixed minus-1-SD cognition scenario, in which the level-specific college estimate was unstable because of quasi-complete separation. Panel B shows adjusted-NfL associations with the primary and alternative global cognition definitions. NfL, neurofilament light chain.