Investigating the Relationship Between Neurofilament Light (NfL) Concentrations in Blood Plasma and Cognitive Task Performance in Healthy Older Adults

Cassandra Blew

Department of Psychology and Neuroscience

University of Colorado – Boulder

Senior Honors Thesis

Defense Date: April 3, 2023

Honors Committee

Thesis Advisor: Dr. Angela Bryan, Department of Psychology and Neuroscience

Honors Council Representative: Dr. Lauren Chun, Department of Psychology and Neuroscience

Outside Reader: Dr. Robert Buchwald, Honors Residential Academic Program

Additional Committee Member: Dr. Kelly Gildersleeve, Department of Psychology and Neuroscience

Abstract

When someone is suffering from neurodegeneration, levels of neurofilament light (NfL), a neuronal cytoskeleton protein important for axonal integrity, in blood plasma increase. A key hallmark of neurodegenerative disorders is cognitive impairment, and previous literature suggests that NfL has potential to gauge the extent of cognitive change in those actively suffering from neurodegenerative diseases. There have been fewer studies, however, on if NfL can be used to assess healthy, non-diseased older individuals' cognitive ability. If such a link could be identified, NfL could be a useful biomarker for pre-onset risk assessments in degenerative disorders of the brain. A sample of 40 participants aged 65 and older were utilized to investigate if there was an association between cognitive performance and NfL in healthy older adults. It was hypothesized that as NfL concentrations in plasma increased, scores on three cognitive tasks (the Rey Auditory Verbal Learning Task, Flanker task, and Picture Sequence Memory Task) would decrease. Three separate multiple regression analyses conducted for each cognitive task compared to concentrations of NfL (controlled for age and gender) showed that only performance on the Picture Sequence Memory Task could be predicted by NfL levels in the blood. This suggests that NfL concentrations may be useful in predicting the capacity of someone's visual working memory, but more research on this potential cognitive blood biomarker must be conducted to imply efficacy.

Keywords: neurofilament light, cognition, older adults, biomarker, neurodegeneration

Investigating the Relationship Between Neurofilament Light (NfL) Concentrations in Blood Plasma and Cognitive Task Performance in Healthy Older Adults

Changes in cognition have been well established for decades as a normal course of aging. Cognition describes the metal processing of thoughts, experiences, and sensations from the environment; it is how we think, learn, and remember. Most older adults have a harder time with specific cognitive tasks compared to younger adults, including but not limited to, delayed free recall, source memory, and prospective memory tasks (Harada et al., 2014). The majority of research around aging and cognition at present is related to dementias (e.g., Alzheimer's disease (AD), Lewy body dementia, frontotemporal dementia, etc.) and mild cognitive impairment (MCI), so there has been less focus on how physiologically normal processes of aging can impede cognitive abilities. Currently, the most common way to assess cognitive ability is through a series of time-intensive mental tasks. However, if a biological measure of cognitive aging could be identified, it could help determine the risk of cognitive decline and potentially serve as an indicator of whether particular interventions are successful at decreasing the risk.

Within current studies of aging and cognition, there has been evidence supporting a neuronal cytoplasmic protein, neurofilament light (NfL), as having potential diagnostic value in neurodegenerative disorders. Research has found that concentrations of NfL in the blood are proportional to the level of axonal damage in the brain, with concentrations increasing as axonal integrity decreases (Gaetani et al., 2019). Benedet et al. (2019) showed that NfL blood concentration has potential to predict a temporal risk for developing Alzheimer's disease (AD) in individuals positive for β-amyloid plaque. NfL has also gained support as a potential proxy for decline in global cognition and attention in individuals actively suffering from AD (Aschenbrenner et al., 2020). Yet, there remains a gap in knowledge about the extent to which NfL correlates with

cognition in non-diseased older adults. One of the few studies to date that focused solely on a cognitively healthy population aged 65 and up found that NfL concentration levels in serum do not depend on comorbidities or depression, except for renal failure, with which it is strongly correlated (Ladang et al., 2022), but it did not compare protein concentration levels to cognitive performance. Another study found that NfL did not correlate with performance on the Keep Track Task, a cognitive test that assesses the flexibility of working memory, in healthy older adults (Karoly et al., 2021).

Because high concentrations of NfL are indicative of axonal damage (Gaetani et al., 2019), it could feasibly be used as a measure of cognitive ability. In traumatic brain injury and multiple sclerosis, axons and myelin sheaths are compromised, often leading to cognitive impairment (Fields, 2008). It is believed that as myelin that surrounds axons is damaged, synaptic signaling loses efficiency, leading to decreases in cognitive ability (Peters, 2002). It has also been postulated that cognitive decline associated with normal aging is a result of the loss of myelin as brains age (Madden & Bennett, 2009). Because NfL is a neuronal cytoskeleton protein highly involved in axonal integrity, if someone is experiencing cognitive decline because of neurodegeneration, other forms of axonal damage, or the age-related loss of myelin, measures of NfL in blood serum might be able to detect this. A meta-analysis conducted in 2022 determined that NfL has strong potential to assess cognition but recommends that more studies be conducted on the topic (Travica et al.).

The purpose of this study, then, was to investigate if NfL is associated with cognitive decline in healthy older individuals. If such a link can be established, then NfL could be a useful pre-diagnostic and risk assessment tool for MCI and dementias.

Cognitive ability was operationalized with three cognitive tasks: the Rey Auditory-Verbal Learning Task (RAVLT) that measures verbal learning and memory (de Sousa Magalhães, 2012); the Flanker task that measures spatial selective attention, reactivity, and executive inhibitory control (Kopp et al., 1996); and the Picture Sequence Memory Task (PSMT) which measures episodic and spatial memory (Dikmen et al., 2014). These cognitive tasks were chosen for analysis because each measures a different aspect of cognition. Having a wide range of cognitive and memory ability measures to compare to NfL levels allowed a full view of the extent to which NfL is related to cognitive ability. Based on the current literature, it was hypothesized that increases in NfL concentrations in blood plasma would be associated with poorer cognitive performance on all three tasks.

Methods

Design

This project utilized a subset of data from a larger study (Aging and Marijuana: Benefits, Effects, and Risks – AMBER; 1R01AG066698, PI: A. Bryan) still currently being conducted at the Center for Health and Addiction: Neuroscience, Genes, and Environment (CUChange) at the University of Colorado at Boulder. Project AMBER is investigating the benefits and/or harms of using cannabis for depression, anxiety, chronic pain, and/or trouble sleeping among older adults (i.e., 60+ years old), and both individuals who are or are not interested in using cannabis to help with these symptoms are currently being recruited. To be eligible to participate in AMBER, the participants must satisfy the inclusion/exclusion criteria listed in Table 1. All participants included in AMBER complete a baseline appointment (duration: 3 hours) and a mobile laboratory session (4 hours) spaced one month apart that include cognitive tasks and blood draws for biological data, followed by three monthly follow-up online surveys. For the purposes of this

project, only baseline task and blood data from AMBER were used so that cannabis did not influence the findings.

Participants

The sub-sample for this study included 40 participants over the age of 65. The sample was 50% male and 50% female, and participants ranged from 65 to 84 years of age (M = 72.38, SD = 5.19). By self-report, 95% of participants identified as white, 2.5% as Asian, and 2.5% preferred not to answer. For self-reported educational attainment, 40% of participants had a master's degree, 30% had a bachelor's degree, 17.5% had a doctoral degree, 7.5% reported some college experience, and 5% had an associate degree or a certificate. By self-report, 25% of participants had a household income <$50,000, 35% had an income between $50,000 and $99,999, 22.5% between $100,000 and $149,999, 12.5% had a household income >$150,000, and 5% of participants preferred not to answer.

The screening process for participants in Project AMBER ensures that all individuals included in this analysis were cognitively high functioning (i.e., have no signs of MCI or dementias) and have no other severe medical problems or psychopathology (see Table 1). Participants were incentivized to participate in AMBER with monetary compensation (baseline compensation: $100 for cannabis users, $50 for non-users). The forty participants selected for this analysis were chosen based on the availability of plasma aliquots, completeness of cognitive data, and being 65 years old or older.

Measures

Rey Auditory Verbal Learning Task (RAVLT)

The RAVLT assesses verbal learning and memory (Vakil & Blachstein, 1993). The researcher played a recorded list of 15 words that are spaced one second apart. Participants were then asked to repeat all the words that they could remember. This process was repeated five times. Then, the participants were presented with a second auditory list of fifteen words and were given one attempt to repeat back as many words as possible from this second list. Immediately after, they were asked to recall as many words from the first list as possible. After a twenty-minute delay, the participants were once again asked to recall as many words as possible from the first list. Lastly, the participants were asked to listen to a final list of fifteen words and were asked to indicate if the words were on the first list. For the purposes of this study, the first 5-trial-sum (5TS) score was used in analysis because it quantifies learning across the first five trials (Morati et al., 2017). The scoring scale ranges from 0-75 points. Scores were calculated using RStudio.

Flanker Task

The Flanker task measures participant attention and inhibitory control (Kopp et al., 1996). The participants were required to focus on a given stimulus, an arrow, while inhibiting attention to similar stimuli flanking the main stimulus on both sides. The middle stimulus could have either been pointing in the same direction as the flankers (congruent) or in the opposite direction (incongruent). The participants were asked to indicate the direction of the middle arrow as quickly as possible while ignoring the flanking stimuli. The task consisted of twenty consecutive trials. Scoring depended on both accuracy and reaction time. For the purposes of this study, the computed Flanker score was used because it is a two-vector sum of all trials that represents overall inhibitory control; the scoring scale ranges from 0-10 points (NIH Toolbox, 2021). The Flanker task was administered and scored using NIH Toolbox® (iPad version).

Picture Sequence Memory Task (PSMT)

The PSMT is a measure of episodic memory (Dikmen et al., 2014). Participants were shown increasingly lengthy series of illustrated objects/activities presented on an iPad while corresponding audio-recorded phrases played as pictures appeared in story format. After all illustrations were presented, they were shuffled and moved to the bottom of the screen, and participants were asked to order the pictures in the sequence in which they were presented. They were given two presentations and attempts to order a sequence of pictures; the first sequence had 15 pictures and the second increased to 18 pictures. The computed PSMT scores were used for the purposes of this study because it is a standardized theta score between the two trials and provides an estimate of episodic memory ability; the scoring scale ranges from 200-700 points (NIH Toolbox, 2021). The PSMT was scored and administered using NIH Toolbox® (iPad version).

Neurofilament Light (NfL) Protein Assay

Concentrations of NfL were measured using the UmanDiagnostics NF-Light™ Serum enzyme-linked immunosorbent assay (ELISA) according to the ELISA kit instructions. A prior study that utilized this serum assay determined that it has an analytical sensitivity of 78.0pg/mL (Kuhle et al., 2016). The manufacturer reports that this assay has an intra-precision of 4.3% (Quanterix, 2018).

Procedures

Project AMBER baseline appointments are run at the Center for Innovation and Creativity (CINC) at CU Boulder. The appointments are structured so that after completing a blood alcohol content breathalyzer and a urine-based toxicology drug screen, participants read

and sign an informed consent document. Next, they have blood drawn; the blood is spun down, separated into plasma aliquots, and stored at -80ºC. Then they are asked to complete surveys (via iPad) regarding demographics, mental and physical health, and substance use. AMBER participants then begin cognitive testing, starting with RAVLT, then the Flanker, and lastly PSMT. Participant data are stored in REDCap (survey responses, RAVLT) and NIH Toolbox® (Flanker, PSMT).

Data Analysis

This project reports the cross-sectional association between cognitive task performance and NfL concentrations in blood plasma. The concentration of NfL in plasma was expected to predict the outcome of cognitive performance. Covariates of interest included age and gender. Associations were explored via multiple regression, and all statistical analyses were conducted using RStudio.

For statistical analysis, RAVLT data were extracted from REDCap (Research Electronic Data Capture), and 5-trial-sum scores were calculated using RStudio. Computed scores for the Flanker and PSM tasks were extracted from NIH Toolbox® and imported into RStudio. The NfL plasma assay data were converted from absorbance wavelengths to concentration values represented in picograms per milliliter (pg/mL) using Microsoft Excel and imported into RStudio.

To reduce the effects of any outliers, all calculated numeric data points (cognitive scores and NfL concentrations) were winsorized to two standard deviations above or below the mean. Scores that were above/below this cut-off point were replaced by the value of the cut-off (i.e., the value of exactly two standard deviations above or below the mean). Winsorization was used for

outliers to retain a sample size of 40 while correcting for extreme scores and NfL concentrations. All analyses were completed after winsorization. A visual comparison of original data to winsorized data can be seen in Figures 1-3.

For statistical analysis, first, descriptive statistics on all tasks were examined to characterize the overall cognitive functioning of the sample. Second, bivariate correlations between all tasks, NfL concentrations, and demographic data were calculated. Finally, three separate multiple regression analyses were estimated to investigate associations between NfL concentration and cognitive performance on each task, controlling for covariates of age and gender.

Results

Concentrations of NfL in blood plasma ranged from 8.58 pg/mL to 48 pg/mL (R = 39.42 pg/mL) with a mean of 20.95 pg/mL (s = 9.82, s^2 = 96.46, SE = 1.55). Three NfL concentration values were abnormally high and were winsorized down. The protein assay for NfL had a relative error within 3% of the standards and a percent recovery also within 3%. All %CV values were <20%, indicating that duplicates were highly similar to each other and that there was a high reliability within the assay.

Performance on the Rey Auditory Verbal Learning Task (RAVLT) ranged from 33 points to 61.15 points (R = 28.15). The mean RAVLT score was 45.53 with a standard deviation of 7.31 (s^2 = 53.47, SE = 1.16). Two scores for the RAVLT were greater than two standard deviations above the mean and were thus winsorized down.

Flanker task performance ranged from 6.51 points to 8.74 points (R = 2.23) with a mean of 7.63 points (s = 0.534, s^2 = 0.285, SE = 0.084). This shows that participants performed

similarly well on the Flanker task. One score for the Flanker was winsorized up, and one score was winsorized down.

Performance on the Picture Sequence Learning and Memory Task (PSMT) ranged from a score of 322.20 points to 592.02 points (R = 269.82). The mean score was 457.40 with a standard deviation of 67.35 (s^2 = 4526, SE = 10.65). The large spread of the data indicates that some participants performed very well on the PSMT, while others performed very poorly. One PSMT score was winsorized up, and one score was winsorized down.

For bivariate relationships between all data points of interest (performance on all tasks, NfL concentration, age, and gender), a correlation analysis was conducted using R Studio. Results from this bivariate correlation are depicted in Table 2. Age was positively correlated with NfL and negatively correlated to all of the cognitive tasks except for PSMT. Gender only had a relationship with the RAVLT in the positive direction. The PSMT and RAVLT shared a strong negative correlation. The Flanker task and RAVLT were moderately positively correlated. The PSMT was negatively correlated with the concentration of NfL in blood plasma.

Multiple linear regression was used to test if concentration of NfL in blood serum, age, and gender significantly predicted performance on the RAVLT. The overall regression was statistically significant (R^2 = 0.273, F (3, 36) = 4.501, p = 0.009). Consistent with the bivariate correlation findings, age and gender significantly predicted RAVLT scores, but NfL concentrations did not (Table 2). A regression table for performance on the RAVLT compared to age, gender, and NfL concentrations can be found in Table 3.

The linear model examining associations between NfL concentration, age, and gender on the Flanker task resulted in no significant associations, and the overall regression was not

significant (R^2 = 0.105, F (3, 36) = 1.411, p = 0.255). A regression table for this analysis can be found in Table 4.

In the model predicting computed PSMT scores using the concentration of NfL, age, and gender, NfL was a significant negative predictor. Specifically, as the concentration of NfL in blood plasma increased by one pg/mL, the PSMT score decreased by 20.19 points. Consistent with the hypothesis, performance on the PSMT got worse as NfL concentration increased. Age and gender were not statistically associated with PSMT scores (Table 2), and the set of predictors accounted for a marginal amount of the variability in PSMT scores, resulting in an insignificant overall regression (R^2 = 0.160, F (3, 36) = 2.285, p = 0.095). A regression table for this analysis of PSMT performance can be found in Table 5.

Discussion

In this investigation of the association of NfL concentration with cognitive performance in healthy older adults, it was hypothesized that as the concentration of NfL in plasma increased, cognitive performance would decrease on the RAVLT, PMST, and the Flanker task. The hypothesis was only partially supported: out of all three of the cognitive tasks, NfL concentration was significantly associated in the predicted direction with PSMT scores but not the RAVLT or Flanker task.

It is possible that the association between NfL and PSMT scores was stronger due to the way the brain processes pictures in working memory. It has been postulated that cognitive decline associated with normal aging is due to the loss of volume in the prefrontal cortex (PFC) (Kanne et al., 1998). The PFC is important for many cognitive processes but is especially crucial in deriving meaning from visual stimuli (Freedman et al., 2001), and the PSMT is the only task

that includes visual stimuli associated with meaning (the visual stimuli in the Flanker task are not associated with semantics). Processing of visual stimuli is superior to that of auditory stimuli, so much so that it is hypothesized that there is a more complex, longer neural circuitry for visual information than that of auditory information (Cohen et al., 2009). If more processing is occurring for visual information, then it is feasible to assume that more axons are involved in visual processing than auditory processing. By this logic, it suggests the possibility that the PSMT involves more axonal communication than the RAVLT and Flanker task, and therefore may be more associated with NfL concentration in blood plasma for this reason. Theoretically speaking, the more axons that information must travel through, the more power there is to detect axonal damage via a blood biomarker.

Interestingly, there was a strong negative correlation between performance on the PSMT (a visual working memory task) and the RAVLT (a verbal/auditory working memory task), meaning that a high score on one task is associated with a low score on the other. This is contradictory to prior research that has found a strong positive correlation between performance on these same two tasks (Loring et al., 2018). While it is difficult to explain this finding conceptually, it may be due to statistical artifacts. There was more variability in the PSMT than in the RAVLT in this study's sample, so perhaps more effort was put in by participants on the first task they were given (RAVLT) than the last task (PSMT) leading to this difference in performance.

There was also a significant correlation between the Flanker task and RAVLT, but this relationship was positive, meaning that as Flanker scores increased so did RAVLT scores. However, the Flanker task does not share a correlation with the PSMT, suggesting that perhaps there was a statistical artefact in the PSMT scores that led to this lack of correlation with the

Flanker and the strong negative association with RAVLT. This could also further support the suggestion that there could have been less participant effort by some put into the last cognitive task administered (PSMT).

Consistent with prior literature (Bridel et al., 2019), age was significantly positively correlated with plasma NfL concentration and supports the idea of axonal degradation being a normal aspect of aging. Performance on the Flanker and RAVLT both showed a negative correlation with age, which demonstrates the notion that cognitive ability declines with age (Oschwald et al., 2018). The absence of a PMST correlation with age is curious but could perhaps also be due to statistical artifacts.

It was also found that RAVLT scores were significantly correlated with gender, with females performing better than males. Findings on the association of gender with RAVLT performance are actually mixed in literature. Some studies find no correlation between gender and RAVLT performance (Ferreira Correia & Campagna Osorio, 2014; de Sousa Magalhães et al., 2012), while others find the same correlation as in this study such that females perform better on the RAVLT than males (Ricci et al., 2022). Clearly, more work on the association of gender to RAVLT performance is needed.

Limitations

A large limitation in this study is the size of the sample under analysis. For a discovery-based biomarker study, it is recommended that a sample of 100 subjects be used (Skates et al., 2013), yet due to financial constraints and the analysis of NfL not being a planned biomarker for Project AMBER, the sample only consisted of 40 participants. It is recommended that this study be conducted again with a larger sample size to see if findings are replicable.

Also, it must be acknowledged that participants in Project AMBER, and therefore those included in this study, must suffer from depression, anxiety, chronic pain, and/or trouble sleeping in order to be eligible to participate. This limits this analysis because the findings may only be generalizable to older individuals who suffer from one or more of these ailments.

Another concern is that this sample of individuals lacked diversity. The vast majority of participants were white, college educated, and middle-class. In order to get a better gauge of the effect of NfL in predicting cognitive decline associated with normal aging in the general population, this study would have to be reproduced with a more diverse sample.

It has recently been found that NfL can be influenced by other factors other than axonal damage and aging, including body mass index (BMI), renal function, blood volume, and increased concentration of high-density lipoprotein (Koini et al., 2021). While Koini et al. stipulate that age is by far the strongest factor influencing NfL concentrations in blood plasma (2021), for this current study to be completely sound, the confounds mentioned above should ideally be controlled for.

This analysis only included one timepoint rather than being a longitudinal study. It is possible that NfL could detect negative changes in cognition overtime in older adults and could therefore be able to serve as a risk indicator for neurodegeneration, but a longitudinal study over the course of many years would have to be conducted to see if this is a valid hypothesis.

Implications and Future Directions

Despite its limitations, this analysis implies that NfL may have the ability to predict cognitive decline, specifically in relation to performance on the PSMT. However, more studies

must be conducted to test if the negative significant correlation between PSMT and NfL can be replicated in larger samples, especially since statistical artifacts may be present in this data.

The PSMT is a relatively new cognitive task created by the National Institutes of Health in 2014 (Dikmen et al., 2014). The newness behind this task results in a lack of data comparing PMST performance to NfL, so the findings in this paper are novel (as far as the author is aware). It is highly encouraged that more research be done on the relationship between NfL and the PSMT. If the correlation in this analysis can be replicated, then it is possible that NfL has efficacy in predicting cognitive change.

It would be beneficial to the growing field of neurodegenerative biomarkers if this study was replicated with the addition of other common cognitive tests used to assess severity of dementias and mild cognitive impairment (MCI), such as eye-tracking tasks (Bueno et al., 2019), the Numberlink puzzle game for decision-related cognitive assessments (Nef et al., 2020), and pattern recognition tasks (Ding et al., 2015). Addition of these tasks in analysis could aide in determining to what extent NfL could predict cognitive decline (if at all) and would help to identify what aspects of cognition NfL could predict change in (if any).

It could also be advantageous if this study was replicated with two different groups: a subset of the sample that consists of cognitively healthy older adults and a subset that is actively suffering from MCIs and dementias. With the ability to compare NfL levels between cognitively healthy older adults and older adults actively struggling with cognition, there could possibly be a stronger link between aging, cognition, and NfL concentrations identified.

Overall, this study took a small step in identifying the efficacy of NfL in blood plasma as a biomarker for cognitive decline. While the findings in this study were somewhat equivocal, it

is important to compare findings in this analysis to other research on the topic to compile a full picture of how NfL is related to cognition.

Works Cited

Aschenbrenner, A. J., Gordon, B. A., Fagan, A. M., Schindler, S. E., Balota, D. A., Morris, J. C., & Hassenstab, J. J. (2020). Neurofilament light predicts decline in attention but not episodic memory in preclinical Alzheimer's disease. *Journal of Alzheimer's Disease, 74(4)*, 1119-1129.

Bridel, C., Van Wieringen, W. N., Zetterberg, H., Tijms, B. M., Teunissen, C. E., Alvarez-Cermeño, J. C., ... & NFL Group. (2019). Diagnostic value of cerebrospinal fluid neurofilament light protein in neurology: a systematic review and meta-analysis. *JAMA neurology*, *76*(9), 1035-1048. DOI: 10.1001/jamaneurol.2019.1534

Benedet, A. L., Ashton, N. J., Pascoal, T. A., Leuzy, A., Mathotaarachchi, S., Kang, M. S., ... & Alzheimer's Disease Neuroimaging Initiative. (2019). Plasma neurofilament light associates with Alzheimer's disease metabolic decline in amyloid-positive individuals. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 11(1)*, 679-689.

Bueno, A. P. A., Sato, J. R., & Hornberger, M. (2019). Eye tracking–The overlooked method to measure cognition in neurodegeneration?. *Neuropsychologia*, *133*, 107191. DOI: 10.1016/j.neuropsychologia.2019.107191

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: the forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General*, *132*(2), 310. DOI: 10.1037/0096-3445.132.2.310

Cohen, M. A., Horowitz, T. S., & Wolfe, J. M. (2009). Auditory recognition memory is inferior to visual recognition memory. *Proceedings of the National Academy of Sciences, 106*(14), 6008-6010. DOI: 10.1073/pnas.0811884106

Dikmen, S. S., Bauer, P. J., Weintraub, S., Mungas, D., Slotkin, J., Beaumont, J. L., ... & Heaton, R. K. (2014). Measuring episodic memory across the lifespan: NIH toolbox picture sequence memory test. *Journal of the International Neuropsychological Society, 20*(6), 611-619.

Ding, W., Ding, L. J., Li, F. F., Han, Y., & Mu, L. (2015). Neurodegeneration and cognition in Parkinson's disease: a review. *Eur Rev Med Pharmacol Sci*, *19*(12), 2275-81.

Ferreira Correia, A., & Campagna Osorio, I. (2014). The rey auditory verbal learning test: Normative data developed for the venezuelan population. *Archives of Clinical Neuropsychology*, *29*(2), 206-215.

Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in neurosciences, 31*(7), 361-370. DOI: 10.1016/j.tins.2008.04.001

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 291*(5502), 312-316. DOI: 10.1126/science.291.5502.312

Gaetani, L., Blennow, K., Calabresi, P., Di Filippo, M., Parnetti, L., & Zetterberg, H. (2019). Neurofilament light chain as a biomarker in neurological disorders. *Journal of Neurology, Neurosurgery & Psychiatry, 90(8)*, 870-881.

Harada, C. N., Love, M. C. N., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in geriatric medicine, 29(4)*, 737-752.

Kanne, S. M., Balota, D. A., Storandt, M., McKeel, D. W., & Morris, J. C. (1998). Relating anatomy to function in Alzheimer's disease: neuropsychological profiles predict regional neuropathology 5 years later. *Neurology*, *50*(4), 979-985. DOI: 10.1212/WNL.50.4.979

Karoly, H.C., Skrzynski, C.J., Moe, E., Bryan, A.D., & Hutchison, K.E. (2021). Associations between inflammatory biomarkers, gray matter, neurofilament light and cognitive performance in healthy older adults. *Frontiers in Aging Neuroscience, 13*. DOI: 10.3389/fnagi.2021.719553

Koini, M., Pirpamer, L., Hofer, E., Buchmann, A., Pinter, D., Ropele, S., ... & Khalil, M. (2021). Factors influencing serum neurofilament light chain levels in normal aging. *Aging (Albany NY), 13(24)*, 25729.

Kopp, B., Rist, F., & Mattler, U. W. E. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*(3), 282-294. https://doi.org/10.1111/j.1469-8986.1996.tb00425.x

Kuhle, J., Barro, C., Andreasson, U., Derfuss, T., Lindberg, R., Sandelius, Å., ... & Zetterberg, H. (2016). Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clinical Chemistry and Laboratory Medicine (CCLM), 54(10), 1655-1661.

Ladang, A., Kovacs, S., Lengelé, L., Locquet, M., Reginster, J. Y., Bruyère, O., & Cavalier, E. (2022). Neurofilament light chain concentration in an aging population. *Aging Clinical and Experimental Research*, 34(2), 331-339.

Loring, D. W., Bowden, S. C., Staikova, E., Bishop, J. A., Drane, D. L., & Goldstein, F. C. (2019). NIH toolbox picture sequence memory test for assessing clinical memory function: Diagnostic relationship to the rey auditory verbal learning test. *Archives of Clinical Neuropsychology*, *34*(2), 268-276. DOI: 10.1093/arclin/acy028

Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: contributions from diffusion tensor imaging. *Neuropsychology review*, *19*, 415-435.

de Sousa Magalhães, S., Fernandes Malloy-Diniz, L., & Cavalheiro Hamdan, A. (2012). Validity convergent and reliability test-retest of the rey auditory verbal learning test. *Clinical Neuropsychiatry*, *9*(3).

Moradi, E., Hallikainen, I., Hänninen, T., Tohka, J., & Alzheimer's Disease Neuroimaging Initiative. (2017). Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage: Clinical*, *13*, 415-427.

Nef, T., Chesham, A., Schütz, N., Botros, A. A., Vanbellingen, T., Burgunder, J. M., ... & Urwyler, P. (2020). Development and evaluation of maze-like puzzle games to assess cognitive and motor function in aging and neurodegenerative diseases. *Frontiers in aging neuroscience*, *12*, 87. DOI: 10.3389/fnagi.2020.00087

NIH Toolbox. (2021, May 25). *NIH Toolbox Scoring and Interpretation Guide for the iPad*. NIH Toolbox: Motor, Emotion, Sensation, Cognition.

Oschwald, J., Guye, S., Liem, F., Rast, P., Willis, S., Röcke, C., ... & Mérillat, S. (2019). Brain structure and cognitive ability in healthy aging: a review on longitudinal correlated change. *Reviews in the Neurosciences*, *31*(1), 1-57. DOI: 10.1515/revneuro-2018-0096

Peters, A. (2002). The effects of normal aging on myelin and nerve fibers: a review. *Journal of neurocytology*, *31*(8-9), 581-593.

Quanterix. (2018). NF-light™ Serum ELISA (RUO). Ultra-sensitive biomarker detection | Quanterix. Retrieved January 8, 2023, from https://www.quanterix.com/wp-content/uploads/2021/10/NF-light-serum-ELISA-brochure.pdf

Ricci, M., Ruggeri, M., Gnisci, C., Pizzoni, L., Gerace, C., & Blundo, C. (2022). Improving Amnesia Diagnostic Accuracy with RAVLT Single Scores and Composite Indices: Italian Normative Data. Archives of Clinical Neuropsychology, 37(8), 1749-1764.

Skates, Steven J., Michael A. Gillette, Joshua LaBaer, Steven A. Carr, Leigh Anderson, Daniel C. Liebler, David Ransohoff et al. "Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies." *Journal of proteome research* 12, no. 12 (2013): 5383-5394. DOI:

Travica, N., Berk, M., & Marx, W. (2022). Neurofilament light protein as a biomarker in depression and cognitive function. *Current Opinion in Psychiatry*, *35*(1), 30-37. DOI: 10.1097/YCO.0000000000000756

Vakil, E., & Blachstein, H. (1993). Rey auditory-verbal learning test: structure analysis. *Journal of clinical psychology, 49*(6), 883-890. DOI: 10.1002/1097-4679(199311)49:6<883::AID-JCLP2270490616>3.0.CO;2-6

Additional Bibliography

Barro, C., Chitnis, T., & Weiner, H. L. (2020). Blood neurofilament light: a critical review of its application to neurologic disease. *Annals of clinical and translational neurology, 7(12),* 2508-2523.

Disanto, G., Barro, C., Benkert, P., Naegelin, Y., Schädelin, S., Giardiello, A., ... & Vehoff, J. (2017). Serum neurofilament light: a biomarker of neuronal damage in multiple sclerosis. *Annals of neurology, 81(6),* 857-870.

Gaiottino, J., Norgren, N., Dobson, R., Topping, J., Nissim, A., Malaspina, A., ... & Kuhle, J. (2013). Increased neurofilament light chain blood levels in neurodegenerative neurological diseases. *PloS one, 8(9),* e75091.

Idland, A. V., Sala-Llonch, R., Borza, T., Watne, L. O., Wyller, T. B., Brækhus, A., ... & Fjell, A. M. (2017). CSF neurofilament light levels predict hippocampal atrophy in cognitively healthy older adults. *Neurobiology of aging, 49,* 138-144.

Nyberg, L., Lundquist, A., Nordin Adolfsson, A., Andersson, M., Zetterberg, H., Blennow, K., & Adolfsson, R. (2020). Elevated plasma neurofilament light in aging reflects brain white-matter alterations but does not predict cognitive decline or Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 12(1),* e12050.

Panek, W. K., Gruen, M. E., Murdoch, D. M., Marek, R. D., Stachel, A. F., Mowat, F. M., ... & Olby, N. J. (2020). Plasma neurofilament light chain as a translational biomarker of aging and neurodegeneration in dogs. *Molecular neurobiology, 57(7),* 3143-3149.

Skillbäck, T., Farahmand, B., Bartlett, J. W., Rosén, C., Mattsson, N., Nägga, K., ... & Zetterberg, H. (2014). CSF neurofilament light differs in neurodegenerative diseases and predicts severity and survival. *Neurology, 83(21),* 1945-1953.

Table 1

Inclusion/Exclusion Criteria for Participant Eligibility in Project AMBER (CUChange)

Inclusion	Exclusion
60+ years of age	> 0 blood alcohol level at screening (to sign consent form
Able to provide informed consent	Report of other drug use (opiate, cocaine, methamphetamine) in the past 90 days or fail urine screen for any of these drugs
cannabis users: must not be using cannabis more than seven times per month in the past 6 months and should not have used more than once per month in the past 6 months to relieve symptoms (depression, anxiety, chronic pain, and/or trouble sleeping) **cannabis non-users**: must NOT be interested in using cannabis but have at least one of the following complaints: depression, anxiety, chronic pain, trouble sleeping	Diagnosis of psychosis (past or current)
Postmenopausal (female)	Current use of antipsychotic medications
Can read and write English at an 8th grade level if not a native English speaker	Score above 8 on the Alcohol Use Identification Test (AUDIT) (indicates heavy drinking)
	History of vertigo or falls within the past six months
	Unwilling/unable to get in and out of the mobile laboratory van

Note. Adapted from AMBER 20-0422 Protocol (IRB Appvd 02.08.23) Table 3 created by Chasmine Malabanan (PRA, CUChange).

Table 2

Bivariate Correlation Analysis Between Variables and Covariates

	1	2	3	4	5	6
5 Trial Sum RAVLT Score	—					
Computed Flanker Score	0.289*	—				
Computed PSMT Score	-0.539***	0.040	—			
Concentration of NfL in Plasma	0.022	-0.120	-0.262*	—		
Age	-0.272*	-0.272*	0.161	0.345*	—	
Gender	0.427***	-0.167	-0.206	0.055	0.024	—

Note. For the purposes of this bivariate correlation, male was coded as 0 and female as 1 (i.e., a positive gender correlation indicates females performed better).

* $p < 0.05$, two-tailed. ** $p < 0.01$, two-tailed. *** $p < 0.005$, two-tailed.

Table 3

Regression Analysis: Prediction of RAVLT Performance by NfL Concentration, Age, and Gender

Effect	Estimate (b)	SE	95% CI		t	p	
			LL	UL			
Intercept	79.627	14.888	49.434	109.821	5.349	5.16e-6	***
NfL [pg/mL]	0.081	0.113	-0.148	0.310	0.719	0.477	
Age	-0.452	0.213	-0.885	-0.019	-2.117	0.041	*
Gender	6.191	2.056	-10.36	-2.022	-3.012	0.004	**

Note. Male was coded as 0 and female as 1 (i.e., a positive gender estimate indicates females performed better).

* $p < 0.05$, two-tailed. ** $p < 0.01$, two-tailed. *** $p < 0.001$, two-tailed.

Table 4

Regression Analysis: Prediction of Flanker Task Performance by NfL Concentration, Age, and Gender

Effect	Estimate (b)	SE	95% CI		t	p	
			LL	UL			
Intercept	9.731	1.205	7.286	12.174	8.075	1.35e-9	***
NfL [pg/mL]	-0.002	0.009	-0.021	0.016	-0.231	0.818	
Age	-0.027	0.017	-0.062	-0.008	-1.567	0.049	*
Gender	0.184	0.166	-0.521	0.154	-1.103	0.277	

Note. Male was coded as 0 and female as 1 (i.e., a positive gender estimate indicates females performed better).

* $p < 0.05$, two-tailed. ** $p < 0.01$, two-tailed. *** $p < 0.001$, two-tailed.

Table 5

Regression Analysis: Prediction of PSMT Performance by NfL Concentration, Age, and Gender

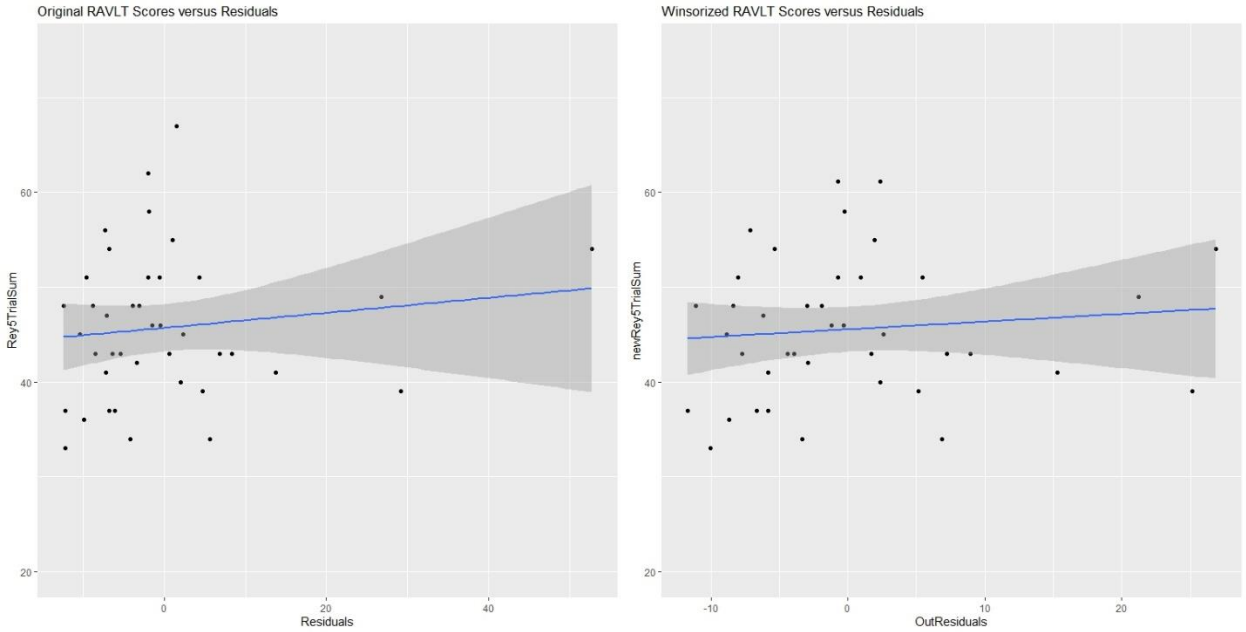
Effect	Estimate (b)	SE	95% CI		t	p	
			LL	UL			
Intercept	231.350	147.375	7.286	12.174	1.570	0.125	
NfL [pg/mL]	-20.193	1.117	-0.021	0.016	-1.963	0.050	*
Age	3.578	2.113	-0.062	-0.008	1.693	0.099	
Gender	-26.184	20.349	-0.521	0.154	1.278	0.209	

Note. Male was coded as 0 and female as 1 (i.e., a positive gender estimate indicates females performed better).

* $p < 0.05$, two-tailed. ** $p < 0.01$, two-tailed. *** $p < 0.001$, two-tailed.

Figure 1

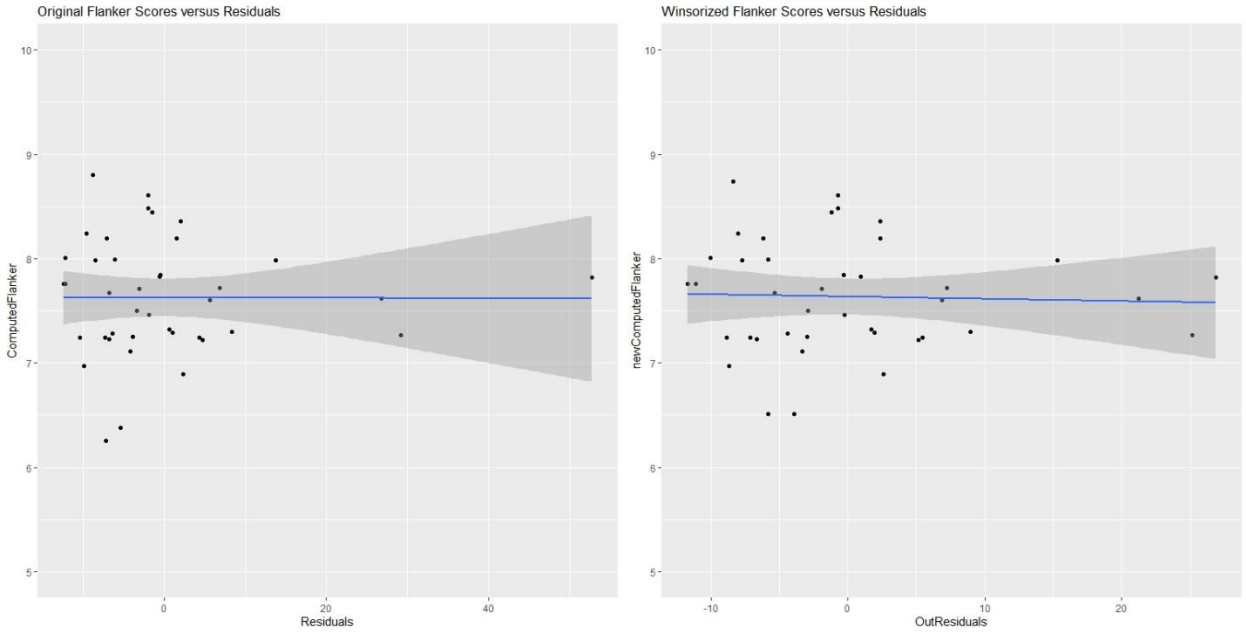
Rey Auditory Verbal Learning Task (RAVLT) Performance vs Residuals Before and After Winsorization



Note. Residuals were calculated using a regression with the outcome being the concentration of NfL with predictors of age and gender. “new” and “Out” preceding the variables on the axes of the rightmost graph indicate winsorization.

Figure 2

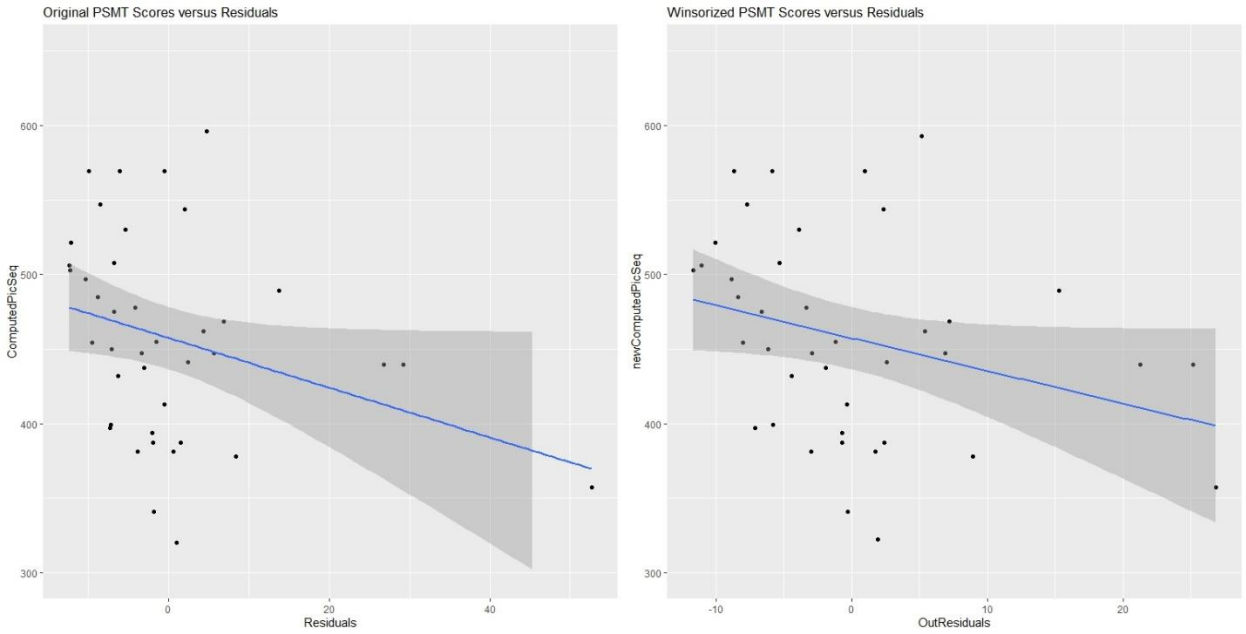
Flanker Task Performance vs Residuals Before and After Winsorization



Note. Residuals were calculated using a regression with the outcome being the concentration of NfL with predictors of age and gender. “new” and “Out” preceding the variables on the axes of the rightmost graph indicate winsorization.

Figure 3

Picture Sequence Memory Task (PSMT) Performance vs Residuals Before and After Winsorization



Note. Residuals were calculated using a regression with the outcome being the concentration of NfL with predictors of age and gender. “new” and “Out” preceding the variables on the axes of the rightmost graph indicate winsorization.