Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ 4/5/2024

ANYUE RUAN Date

Correlation between biomarkers associated with Alzheimer's disease and the cognitive practice effect

By

Anyue Ruan
MSPH
Department of Biostatistics and Bioinformatics

John Hanfelt
Committee Chair

Felicia Goldstein
Committee Member

Correlation between biomarkers associated with Alzheimer's disease and the cognitive practice effect

By

Anyue Ruan
Bachelor of Science
University of Maryland, College Park
2020

Thesis Committee Chair: John Hanfelt, P.H.D

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Department of Biostatistics and Bioinformatics
2024

Abstract

Correlation between biomarkers associated with Alzheimer's disease and the cognitive practice effect

By Anyue Ruan

Alzheimer's disease (AD) is an enormous challenge to society and there is a need to improve the understanding of its pathogenesis and to develop effective interventions. Mild cognitive impairment (MCI) is an important prognostic indicator for the development of Alzheimer's disease and provides a window for early intervention. Cognitive tests play a key role in assessing cognitive decline, and biomarkers provide additional information to understand disease progression. This study utilized data from the Emory Healthy Brain Study (EHBS) to explore the correlation between biomarkers and cognitive test scores, focusing on practice effects. Considering demographic factors and short-term training effects, the study used a standardized regression (SRB)-based approach to predict cognitive scores at follow-up visits. Results showed negative correlations between most cognitive tests, cerebrospinal fluid (CSF) protein levels, and cognitive test scores, indicating significant cognitive decline. Participants with poorer initial cognitive performance showed greater changes in scores between visits, suggesting a more pronounced effect on cognition over time. However, the analysis encountered limitations due to the upper limit of scores on the cognitive tests, which may not adequately reflect the cognitive abilities of high performers. Despite these limitations, the findings of this study emphasize the importance of early detection and intervention for cognitive decline, especially for individuals exhibiting MCI biomarkers. Future work includes continually refining the selection of cognitive tests through factor analysis to improve the methodological robustness. In conclusion, this study provides valuable insights into the correlation between AD biomarkers and cognitive functioning, emphasizing the need for comprehensive assessment and individualized interventions in countering AD progression.

Correlation between biomarkers associated with Alzheimer's disease and the cognitive practice effect

By

Anyue Ruan
Bachelor of Science
University of Maryland, College Park
2020

Thesis Committee Chair: John Hanfelt, P.H.D

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Department of Biostatistics and Bioinformatics
2024

Introduction

Alzheimer's disease (AD) is a serious challenge to contemporary society, with a profound impact on both individual well-being and the healthcare field more broadly. This irreversible neurodegenerative disease manifests itself as a persistent decline in cognitive ability, affecting important aspects of memory, reasoning, and everyday functioning. As the global population ages, the prevalence of Alzheimer's disease is on a significant upward trend, creating an urgent need to unravel the complex mechanisms underlying its pathogenesis. This urgent need has prompted researchers and healthcare professionals to actively work towards a deeper understanding of the disease and the development of effective interventions.

The core of these efforts is to identify and screen for mild cognitive impairment (MCI), an important prognostic indicator for the onset of Alzheimer's disease. MCI represents a pivotal moment in the disease's trajectory, marking a significant decline in cognitive ability but not yet at a level of severity associated with full-blown Alzheimer's disease. Recognizing MCI is critical because it has the potential to be a great opportunity for early intervention, potentially mitigating or preventing further progression of Alzheimer's disease.

In light of this, neuropsychological testing has emerged as an indispensable tool for delving into the subtle nuances of cognitive decline. These assessments, evaluating memory reasoning, language, and other domains, furnish invaluable insights into an individual's cognitive capabilities. The progression of Alzheimer's disease is associated with changes in brain volume and specific proteins. By examining the correlation between these biomarkers and the outcomes of cognitive tests, while also accounting for practice effects, it becomes possible to formulate methods for earlier and more cost-effective prediction and diagnosis of Alzheimer's disease. Systematically measuring these domains empowers scientists to diagnose cognitive dysfunction

and chart its progression. Furthermore, cognitive tests assume a pivotal role in the realm of research, aiding scientists in unraveling the complexities inherent to Alzheimer's disease and facilitating the development of targeted interventions.

Background

The diagnosis of Mild Cognitive Impairment (MCI) involves the utilization of comprehensive cognitive assessments, with the aim of not only identifying existing impairments but also predicting future transition to MCI. This research undertakes a comprehensive investigation into MCI, employing the expansive dataset housed within the Emory Healthy Brain Study (EHBS) database. The EHBS to date had registered 2,500 individuals aged between 50 and 75 years, devoid of any diagnosis pertaining to Alzheimer's Disease, mild cognitive impairment, or other memory-related conditions. The EHBS study entails biennial in-person sessions, incorporating a range of evaluations such as neuropsychological examinations, cardiovascular assessments, retinal and cerebral imaging, biospecimen procurement encompassing blood, cerebrospinal fluid, and gut microbiome samples, alongside additional assessments (Goetz et al., 2019).

The EHBS database incorporates diverse cognitive tests encompassing assessments including overall cognitive status (Montreal Cognitive Assessment (MoCA)), language (Multilingual Naming Test(MiNT), Letter and Animal Fluency), visuospatial ability (Judgment of Line Orientation (JOLO)), verbal memory (Rey Auditory Verbal Learning Test (RAVLT)), and visual memory (Rey Complex Figure Test (RCFT)). These tests collectively contribute to a comprehensive understanding of cognitive function and impairment. The dataset comprises a

rich tapestry of multiple distinct cognitive test scores, providing a nuanced perspective on cognitive abilities and potential areas of impairment.

Beyond cognitive assessments, this research incorporates additional layers of information by obtaining Magnetic Resonance Imaging (MRI) of the brain and cerebrospinal fluid (CSF) protein data sourced from participants, which function as valuable biomarkers in the context of MCI research. In the current study, we examined brain volume measurements of various regions, and CSF levels of $A\beta_{42}$, total Tau (tTau), phospho181-Tau (pTau), and the ratio of tTau to $A\beta_{42}$.

Method

Hammers et al.'s standardized regression-based (SRB) prediction equation was cited in the study. Hammers 's research encompassed various commonly administered cognitive tests, such as the Hopkins Verbal Learning Test—Revised, the Brief Visuospatial Memory Test—Revised, and the Controlled Oral Association Test. These assessments were conducted twice within a week on community-dwelling older adults, some of whom were considered cognitively intact and some of whom were diagnosed with amnesic MCI. Approximately one year later, the same cohort underwent a third assessment using the same battery of tasks. The equation not only integrated baseline scores and demographic characteristics into their SRB equations for predicting test scores at the one-year mark but also factored in short-term practice effects specific to the tests and individuals observed between baseline and the one-week administration of these tasks. By incorporating diagnostic status, such as cognitively intact and MCI into the complex SRB calculations, Hammers et al.'s prediction equations were designed to be applicable to an independent sample of older adults with either normal cognition or MCI (Hammers et al., 2021).

In our recent research, it became apparent that our data collection framework differed significantly from that of Hammers's study. Unlike Hammers's study, which drew upon data collected within an approximately one-week and one-year timeframe, the EHBS database conducted tests and gathered data on a biennial basis. Consequently, this necessitated adjustments to the equation utilized in our research framework.

In the process of standardizing the practice effect within a cognitive testing framework, a regression model is employed to quantify and assess the variations in participants' cognitive test scores between two consecutive visits. This model, expressed as

$$T_2 - T_1 = \beta_0 + \beta_1(T_1 - \overline{T_1}) + \beta_2 * \text{Demographic},$$

aims to elucidate the relationship between the observed cognitive test score differences ($T_2 - T_1$) and various factors, including the average cognitive score at the initial visit $\overline{T_1}$, demographic characteristics denoted by Demographic, and model parameters β_0, β_1, and β_2.

Breaking down the components, T_1 represents the cognitive test score at the first visit, $\overline{T_1}$ denotes the average cognitive score during the initial visit, and T_2 signifies the cognitive test score at the second visit. Additionally, $(T_2 - T_1)'$ stands for the predicted cognitive test score difference based on the regression model.

The standardized regression-based (SRB) metric, defined as

$$\text{SRB} = \frac{T_2 - T_1 - (T_2 - T_1)'}{\text{RMSE}},$$

emerges as a crucial output of this model. If SRB exceeds zero, it signifies that participants have exhibited a practice effect larger than what would be anticipated based on their demographic characteristics. Essentially, SRB acts as a standardized indicator, offering a quantified prediction of the cognitive test score difference between the initial and subsequent visits.

A subsequent step involves correlating the standardized prediction of cognitive test score differences (SRB) with biomarkers. This correlation analysis was conducted using the Pearson method, allowing for the inclusion of missing data in the assessment. The method was chosen due to its capability to handle missing data effectively, ensuring that our analysis remains robust even in the presence of incomplete information. However, it is noted that an impediment arises in the form of approximately 80% missing values for the Rey Complex Figure Test (RCFT) Emory Scoring. Consequently, it becomes unreliable to include RCFT Emory Scoring in our correlation analysis due to the overwhelming number of missing values. Without a sufficient amount of data for this parameter, attempting to establish correlations could potentially mislead our interpretation of the relationship between cognitive test scores and biomarkers. Therefore, we made the decision to exclude the RCFT.

Results

Initially, EHBS had a total of 1985 participants. Of these, 1034 participants experienced two or more visits during the study period. In addition, there were 802 participants who had less than 2.5 years between their first and second visits. This subset of 802 individuals was specifically selected for inclusion in the correlation analysis.

In examining the demographic distribution within the EHBS cohort, female participants accounted for a significant majority, comprising 66.3% of the total population. Regarding age, the average age at the first visit was 63.1 years, with a standard deviation of 6.6 years. The age range spanned from 45.2 years to 79.6 years, reflecting a diverse spectrum of ages represented within the study (Table 1).

Analyzing the racial composition of the EHBS participants revealed further diversity within the cohort. Among the racial categories, Asian individuals represented 1.6%, followed by Black individuals at 17.2%. Native American participants accounted for 0.25%, while individuals identifying as "Other" comprised 1.01%. The largest racial demographic within the study was White individuals, constituting 79.4% of the total population. The mean (SD) education level was 16.7 (2.1) years, with a range from 10-2- years. This span showcases the diversity in educational backgrounds within the participant pool, with individuals having completed varying durations of formal education (Table 1).

Table 1:

Table 1 Descriptive Statistics for All Variables of Interest

Variable	Level	N = 1985	%
Gender	Female	1315	66.2
	Male	670	33.8
Primary_Race	American Indian or Alaska Native	12	0.6
	Asian	31	1.6
	Black or African American	342	17.2
	Native Hawaiian or Other Pacific Islander	5	0.3
	Other Race	20	1.0
	Caucasian or White	1575	79.3

Variable	Level	N = 1985	%
Age_at_visit	Mean	63.11	-
	Median	63.40	-
	Minimum	45.20	-
	Maximum	79.60	-
	Std Dev	6.64	-
	Missing	0	-
Education	Mean	16.70	-
	Median	16	-
	Minimum	10	-
	Maximum	20	-
	Std Dev	2.08	-
	Missing	0	-

Pearson analysis was performed to analyze the correlation between Standardized prediction of cognitive test score difference between visit one and two (SRB) and AD's biomarkers. After analyzing the correlation between SRB scores for each cognitive test and AD biomarkers, it became evident that verbal fluency and CSF protein exhibit the most robust negative correlation among the variables considered. The shade of blue in their correlation suggests a particularly strong inverse relationship, such that a higher practice effect is associated with lower biomarker level. Following closely in terms of negative correlation strength is delayed verbal recall (RAVLT 6) (Figure 1).

In contrast, the remaining cognitive tests appear to have comparatively milder or no discernible correlation with the biomarkers. The colors corresponding to these associations are generally lighter, hinting at a weaker or negligible connection. This nuanced understanding of the correlation heatmap enables us to identify and highlight the varying degrees of associations between verbal fluency, CSF protein, RAVLT 6, and the other cognitive tests with precision.

Figure 1:

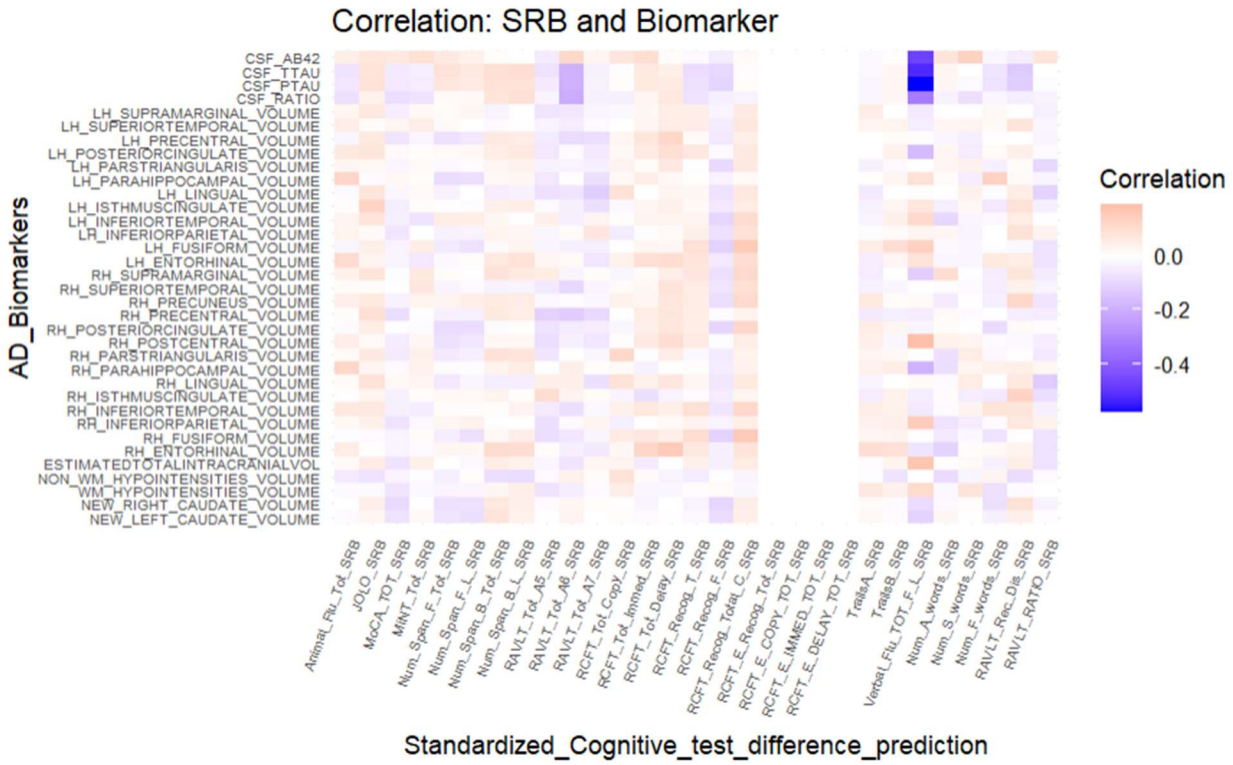

Figure 1: In the correlation heatmap, the intensity of color serves as an indicator of the strength of correlation. Specifically, darker hues signify a more pronounced correlation, with red denoting a positive association and blue indicating a negative one.

In the context of the regression model, an insightful pattern emerges. Patients who perform well during their first visit tend to exhibit smaller differences in their cognitive test scores between visits. This suggests that the cognitive impact on these individuals is relatively modest. Conversely, patients who initially perform poorly exhibit more significant cognitive score variations, indicating a greater cognitive effect on this subgroup during the subsequent visit.

Discussion

The examination of practice effects in this study is limited due to the presence of ceiling effects of some cognitive tests. These score ceilings impose limits on the attainable scores, thereby creating a situation where participants, despite achieving higher scores, remain constrained by the ceiling score. This limitation has potential implications for the analysis, introducing a certain level of influence on the overall interpretation of the results. To be further detailed, cognitive tests are designed to assess various aspects of cognitive function, such as memory, attention, and problem-solving skills. However, the existence of score ceilings can impact the accuracy and depth of the analysis. Participants who excel in these cognitive tests may not reflect their true cognitive capabilities due to the imposed ceiling, leading to an underrepresentation of their actual cognitive performance. It is essential to consider the nuances associated with score ceilings in cognitive testing. While participants may exhibit superior cognitive abilities beyond the ceiling score, the test's design restricts the differentiation and measurement of such higher-level skills. This limitation is particularly noteworthy when attempting to draw meaningful conclusions from the analysis, as it may inadvertently downplay the true cognitive potential of certain individuals. The presence of score ceilings can also influence the overall statistical outcomes of a study. The aggregated data may reflect a skewed representation of participants' cognitive abilities, creating challenges in accurately characterizing the sample group. Therefore, caution and appropriate adjustments should be applied when interpreting results derived from cognitive tests with score ceilings. In conclusion, the analysis conducted faced limitations in comprehensiveness due to the inherent constraints posed by cognitive tests featuring score ceilings. Acknowledging these limitations is crucial for a nuanced understanding of the participants' cognitive abilities and ensures that the findings are interpreted

with a balanced perspective, considering the potential influence of score ceilings on the overall analysis.

Furthermore, it is noteworthy that all participants involved in this research were cognitive normal, signifying the potential absence of observable symptoms. However, based on their CSF levels, some are exhibiting indications of Alzheimer’s neuropathology, with the likelihood of progression into symptomatic Alzheimer's disease in the future. This particular aspect regarding the selection of participants distinguishes this study from others.

Our ongoing efforts include the continuation of factor analysis to refine and streamline the number of cognitive tests administered. Given that cognitive tests are typically categorized into various domains, employing factor analysis enables us to eliminate redundant or insensitive tests, thereby mitigating potential biases and ensuring a more focused assessment. This strategic approach aligns with our commitment to methodological rigor and precision in the pursuit of comprehensive cognitive evaluations.

References

Hammers DB; Porter S; Dixon A; Suhrie KR; Duff K; (n.d.). *Validating 1-year reliable change methods*. Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists. https://pubmed.ncbi.nlm.nih.gov/32885234/.

Tian, G., Hanfelt, J., Lah, J., & Risk, B. (2023a, March 1). *Mixture of regressions with multivariate responses for discovering subtypes in alzheimer's biomarkers with detection limits*. arXiv.org. https://arxiv.org/abs/2303.00715.

Goetz, M. E., Hanfelt, J. J., John, S. E., Bergquist, S. H., Loring, D. W., Quyyumi, A., Clifford, G. D., Vaccarino, V., Goldstein, F., Johnson 2nd, T. M., Kuerston, R., Marcus, M., Levey, A. I., & Lah, J. J. (2019, August 27). *Rationale and design of the Emory Healthy Aging and emory healthy brain studies*. Karger Publishers. https://karger.com/ned/article-abstract/53/3-4/187/211260/Rationale-and-Design-of-the-Emory-Healthy-Aging?redirectedFrom=fulltext.