

Abstract

In this analysis, we examine the relationship between Alzheimer's disease (AD) and various clinical covariates. To this end, we perform logistic regression to determine which covariates of interest related to AD. In corroboration with past research, we find that the APOE4 allele frequency is most associated with Alzheimer's disease. Interestingly, after for highly influential points, we also find that years of education is significantly negatively associated with Alzheimer's disease.

Introduction

Alzheimer's Disease Overview¹:

- Neurodegenerative disease affecting the central nervous system (CNS)
- Affects 1 in 10 people over the age of 65
- 6th leading cause of death
- Rate of deaths due AD nearly doubled from 2000 to 2014, as a result of a rapidly aging Baby-Boomer generation
- Significant financial burden of ~ \$260 Billion on the U.S. healthcare system

Research Question:

- What are the clinical factors associated with Alzheimer's disease?

Methods

Our data is from the Allen Institute's "Aging, Dementia, and the TBI Study" research study². This dataset includes 107 randomly sampled deceased participants who had been diagnosed with dementia prior to death, based on the NINCDS-ADRDA criteria at enrollment and post-mortem neuropathology testing.

Variables of interest:

- Education, APOE4 allele presence, age, sex, age at first traumatic brain injury (TBI), longest loss of consciousness (LOC), number of TBI's with LOC, race, and AD status.

We address missing data using hot deck imputation, where missing values are imputed by drawing from similar cases at random³.

Logistic Regression model:

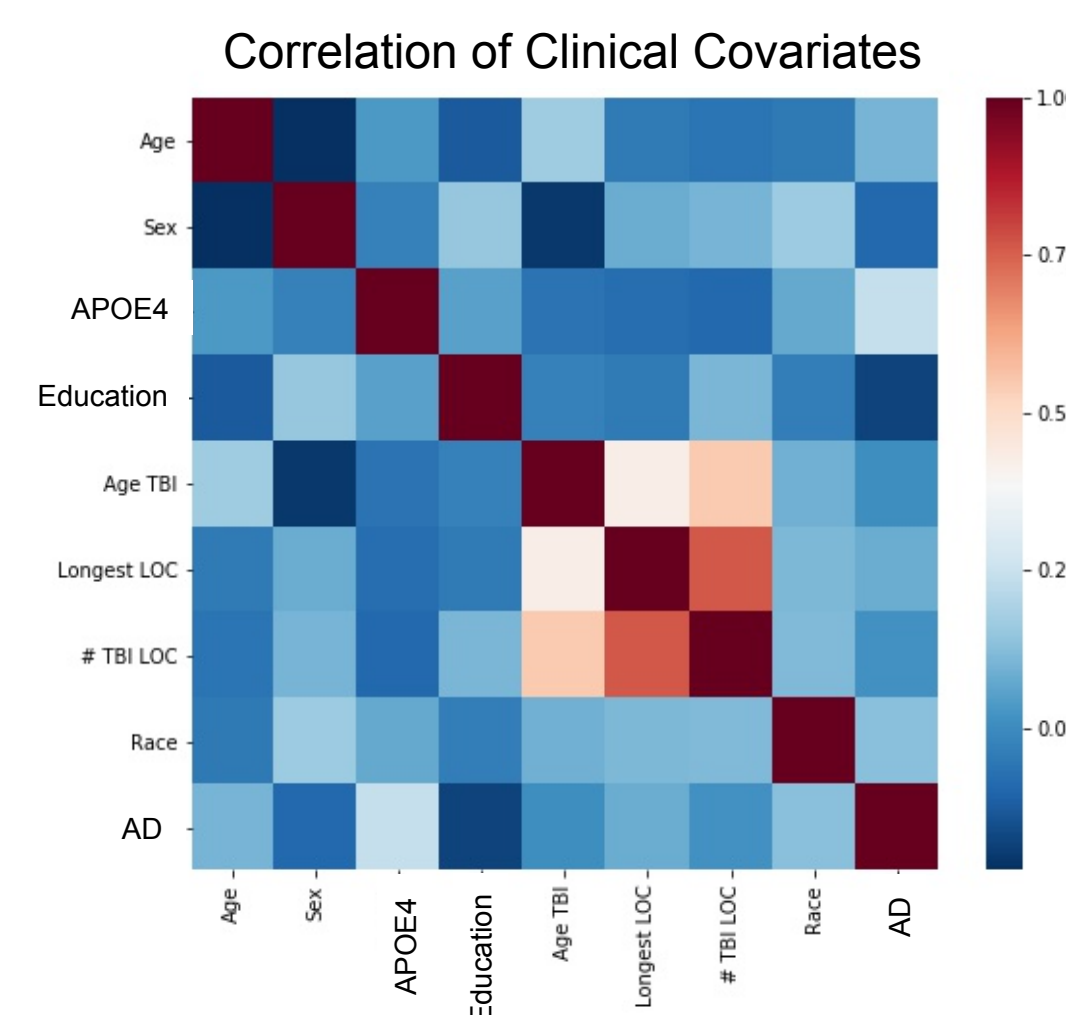
$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p$$

Hypothesis Test:

$$H_0: b_i = 0 \quad H_1: b_i \neq 0$$

The assumptions for Logistic Regression:

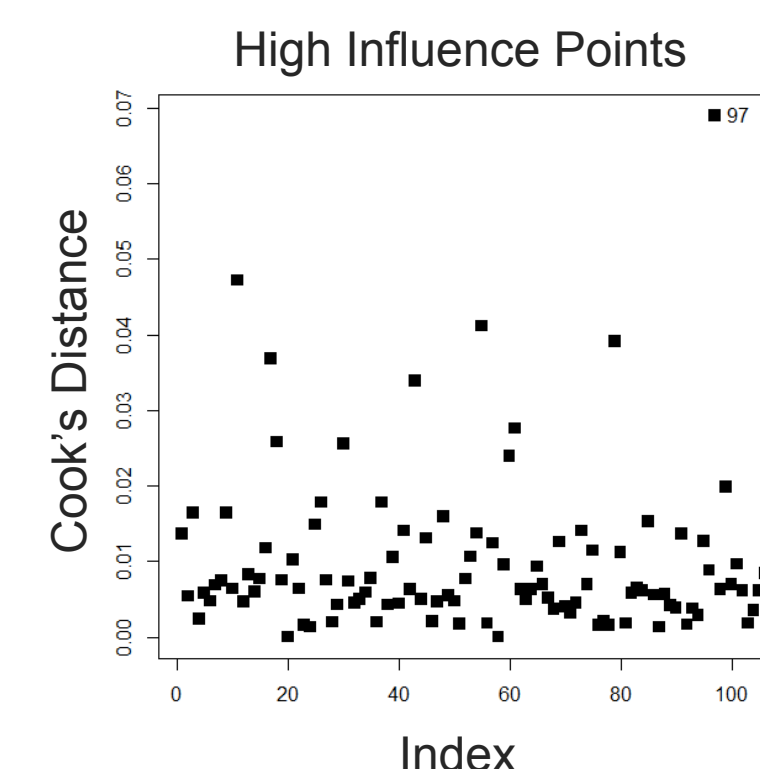
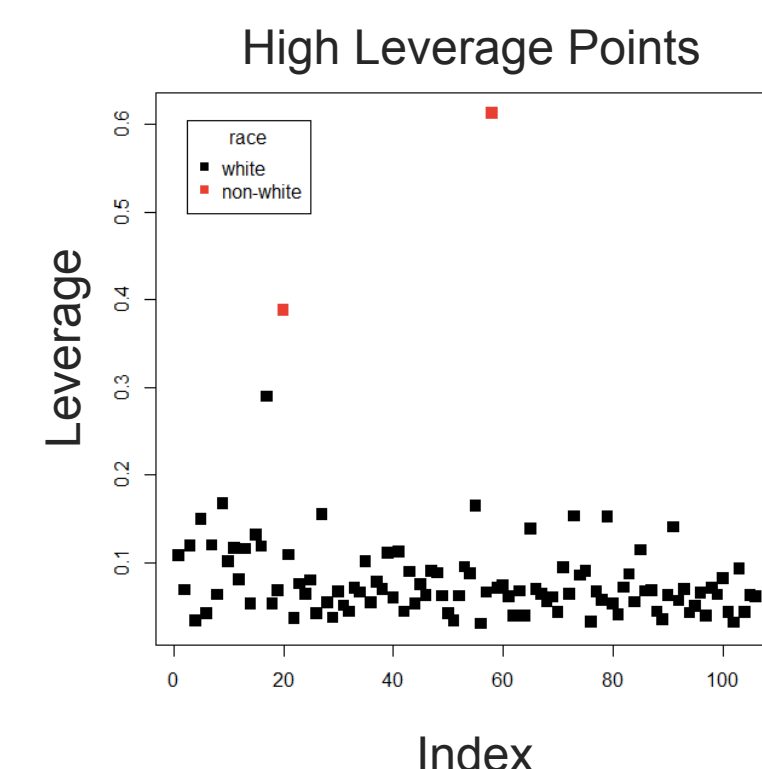
- Independent observations
- Little collinearity between variables (correlation matrix below)
- Linearity of independent variables to log odds



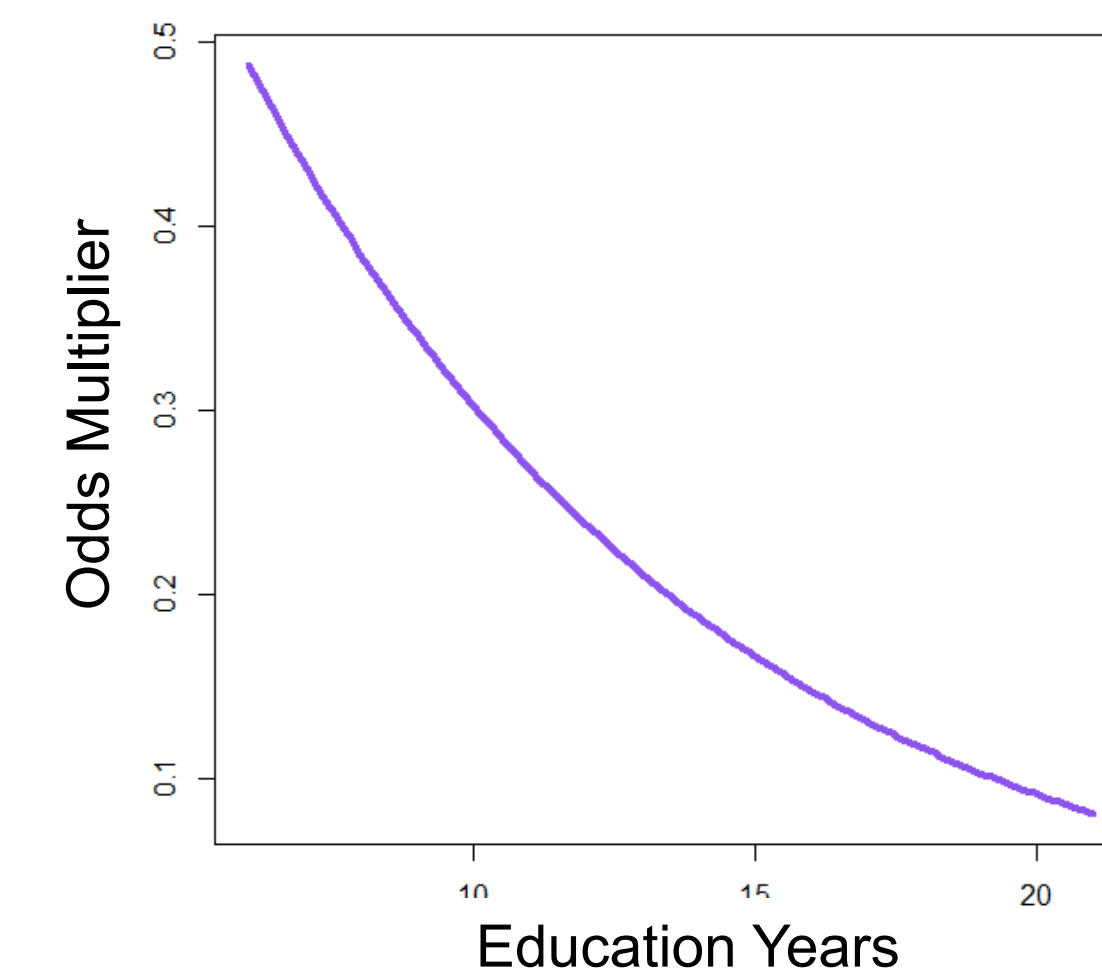
Results

	Covariate	Estimate	p-value
Model 1 (Full Data)	Education	-0.12	0.08
	APO E4 Allele	1.2	0.021*
Model 2 (No influential points)	Education	-0.14	0.045*
	APO E4 Allele	1.27	0.016*
ANOVA comparing models p-value = 0.04			

- These p-values are calculated from the Likelihood Ratio Test
- APO E4 allele increases the odds of AD, more years of education decreases the odds of AD⁴
- One year increase in education results in a decrease in odds of developing Alzheimer's by 11% ($1 - e^{-0.12}$)



Odds Multiplier Graph



The plotted odds multiplier graph shows the effect of increasing years of education on the odds of developing Alzheimer's Disease,

Discussion

Conclusions:

- Years of education and the presence of the APO E4 Allele are significant covariates associated with Alzheimer's Disease

Implications:

- These results show the importance APOE allele analysis
- Further studies may look deeper into the effect of education on Alzheimer's Disease development – potentially preventative treatment to AD
- The two high leverage points are subjects whose race was Non-White, which may indicate that the covariates interact differently for these patients.

Limitations of Study:

- The analysis presented here does not include neuropathology data, which may be informative

References

1. Latest Alzheimer's Facts and Figures. (2016, March 29). Retrieved March 09, 2018, from <https://www.alz.org/facts/>
2. Aging, Dementia and Traumatic Brain Injury Study. (n.d.). Retrieved March 09, 2018, from <http://aging.brain-map.org/overview/home>
3. Crammer, S. et. al. *hot.deck*. (2016, January 04)
4. Cobb, J. L. et. Al (1995, September). The effect of education on the incidence of dementia and Alzheimer's disease in the Framingham Study. <http://www.ncbi.nlm.nih.gov/pubmed/7675231>