

Predicting Cognitive Decline

Group 8:

Brett Castro, Luke Napolitano, Sam
Remmey, Lucy Shichman, Sanjana
Srivastava



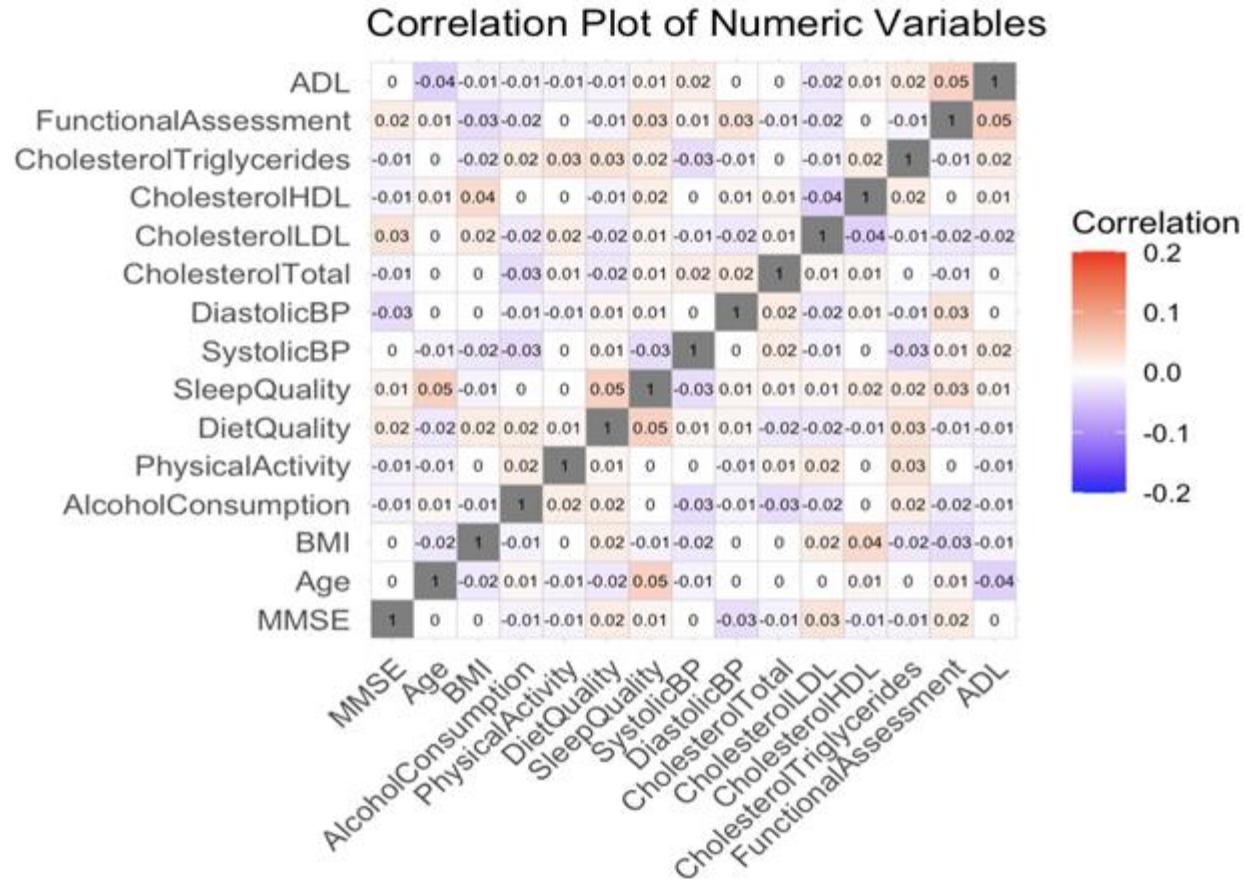
Introduction

- Alzheimer's Disease: Progressive neurodegenerative disease that deteriorates memory and other important mental functions
 - Highly prevalent in patients 85+
 - Earlier detection allows for earlier treatment
 - Clinical diagnostic tools are expensive and invasive
- Research Questions
 - How can we use health indicators to predict cognitive decline?
 - Mini-Mental State Exam (MMSE) via MLR
 - Diagnosis via Logistic Regression
- Dataset summary:
 - Health information of 2149 patients
 - Symptoms, demographics, medical assessments, and medical history
 - 14 numerical, 22 categorical variables

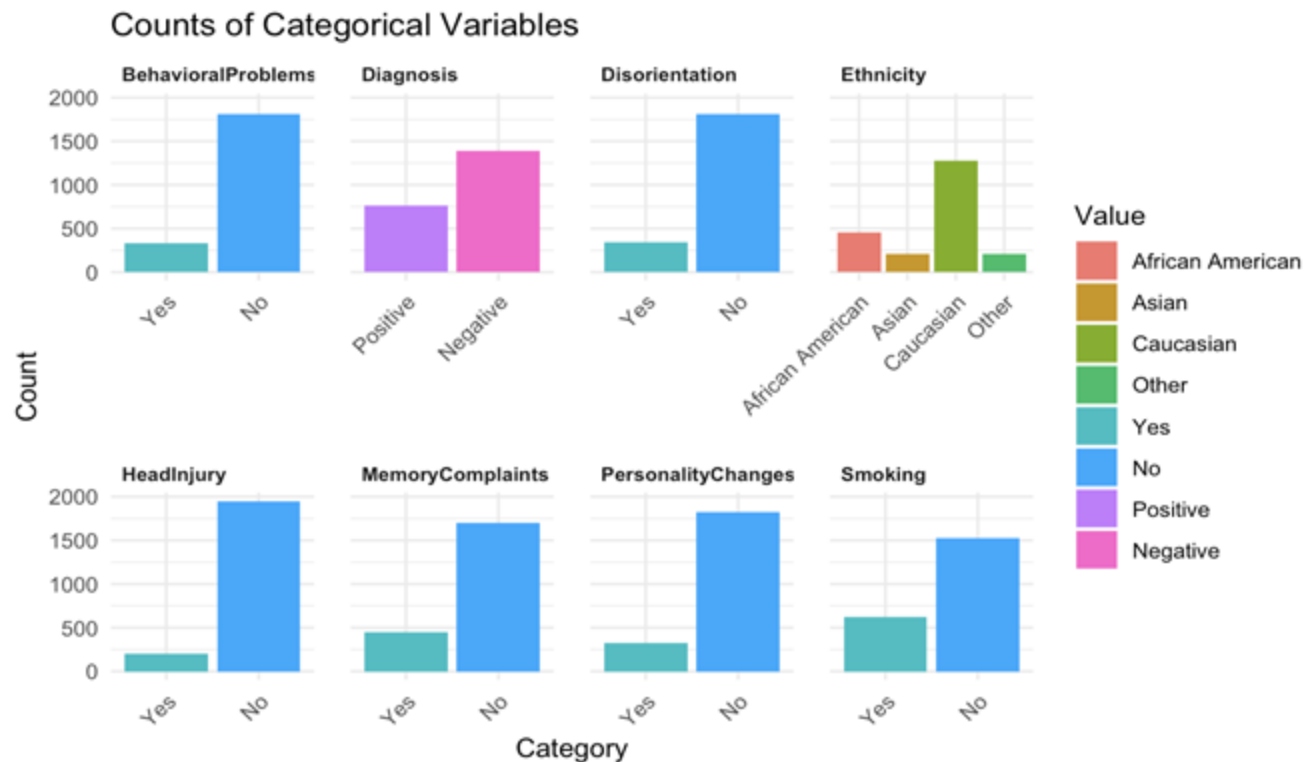


Exploratory Data Analysis - Numerical

- Low correlations
- Strongest correlation: 0.05



Exploratory Data Analysis - Categorical



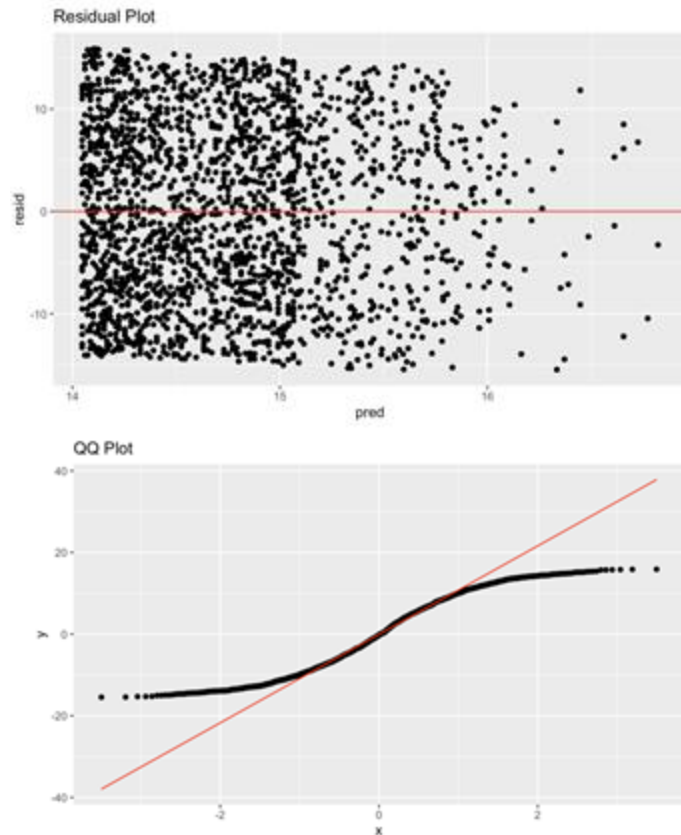
- Key categorical variables
- More negative diagnoses
- Majority white patients

MLR: How can we use health indicators to predict MMSE score?

- Building the model:
 - checking assumptions
 - log transformations
 - trimming influential/leverage points
 - stepAIC
 - lasso

$$\text{MMSE} = 21.3271 - 1.5233(\log_DiastolicBP) + 0.7412(I_PersonalityChangesYes) + 1.0159(I_DisorientationYes)$$

- P-value = 0.03919, Adjusted $R^2 = 0.002528$, Residual SE = 8.533
- Cross-validation
 - 5 folds, 10 repetitions
 - RMSE = 8.539161, $R^2 = 0.004267931$
- Prediction Interval
 - Given a diastolic BP of 72, no personality changes, and no disorientation, we are 95% confident that MMSE score is between -211.6598 and 34.95187
 - actual: 21.46353



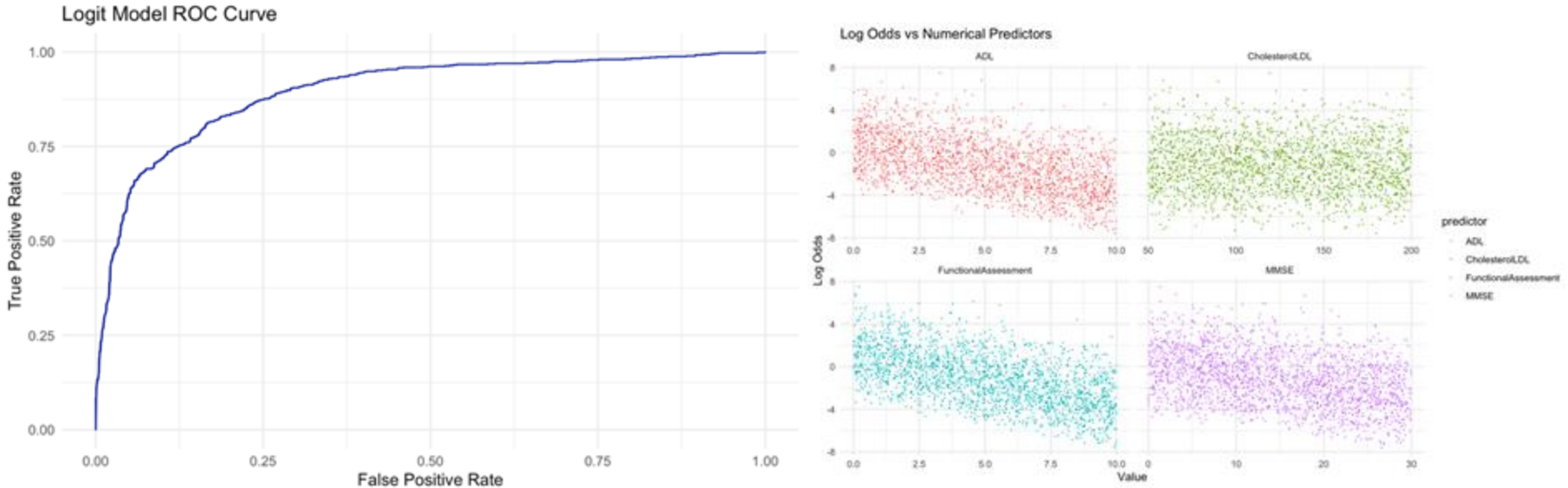
Logistic Regression: How can we use health indicators to predict Alzheimer's diagnosis?

- Building the model
 - Including MMSE, trying to predict diagnosis
 - No transformations
 - Reducing model to only include predictors significant at $\alpha=0.05$
 - stepAIC (AIC:1835.6 -> AIC:1615), reducing information loss
 - Lasso: kept same variables as stepAIC model

Log odds = 4.268 - 0.003(CholesterolLDL) - 0.107(MMSE) - 0.445(FunctionalAssessment) +
2.595(I_MemoryComplaintsYes) + 2.487(I_BehavioralProblemsYes) - 0.416(ADL)

- AUC = 0.903, Accuracy = 0.841, F-1 score = 0.798
- Cross-validation
 - 10 folds
 - AIC = 1617.2
- Prediction
 - Based on this model, the predicted probabilities of a positive diagnosis are **0.396** and **0.929** for two different patients who have an actual diagnosis of negative and positive, respectively.

Logistic Regression: How can we use health indicators to predict Alzheimer's diagnosis?



AUC = 0.903, Accuracy = 0.841, F-1 score = 0.798

Conclusion & Future Work

- We found that the logistic model predicting diagnosis was far more useful than the multi-linear regression model predicting MMSE score
- Other explorations:
 - predicting MMSE:
 - Random Forest: 8.207 RMSR, 9.17% variance explained
 - predicting diagnosis:
 - Support Vector Machine: Overall accuracy (94%)
- Developing non-invasive alzheimer's diagnostic modeling:
 - Machine Learning with clinical data
 - NLP with speech testing data



Research Citation: Moradi, E., Prakash, M., Hall, A., Solomon, A., Strange, B., Tohka, J., & Alzheimer's Disease Neuroimaging Initiative (2024). Machine learning prediction of future amyloid beta positivity in amyloid-negative individuals. *Alzheimer's research & therapy*, 16(1), 46. <https://doi.org/10.1186/s13195-024-01415-w>

Research Citation: Amini S, Hao B, Yang J, et al. Prediction of Alzheimer's disease progression within 6 years using speech: A novel approach leveraging language models. *Alzheimer's Dement*. 2024; 1-9. <https://doi.org/10.1002/ak.13886>

Thank You! Questions?

