

# Clinical and Imaging Analyses to Predict Alzheimer's Disease

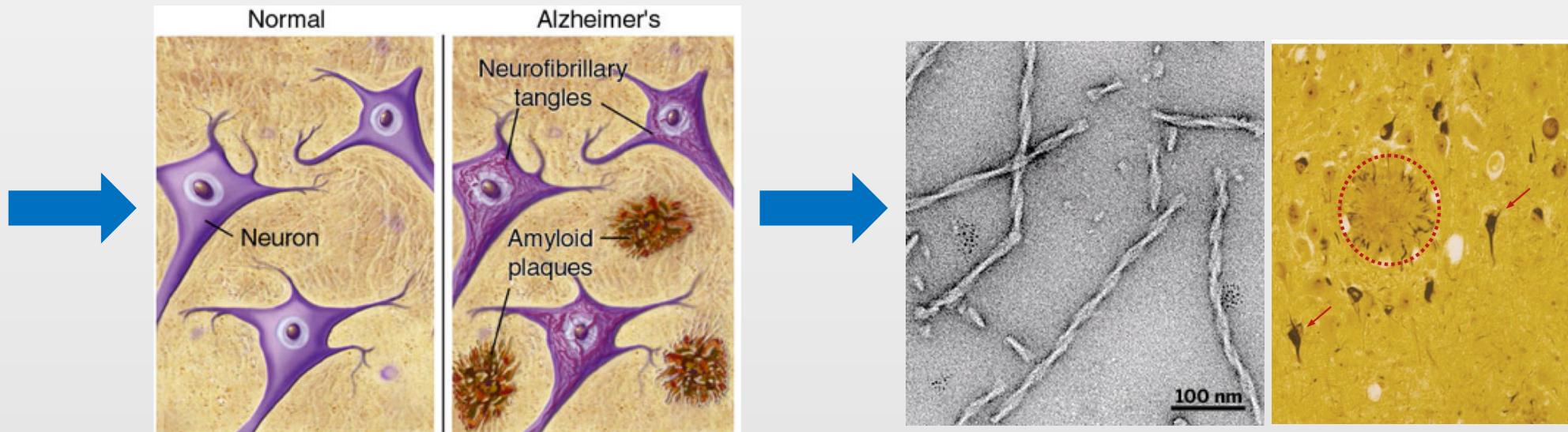
Shannon Ballard, Ph.D.  
Springboard Data Science Career Track

# The Big Picture Question:

**Can we predict whether an individual  
will be affected by a disease?**

# Alzheimer's Disease

Healthy Brain Severe AD



- Progressive loss of neurons and neural connections
- Characterized by tangles in neuron and amyloid plaques
- Affects 1 in 20 people by age 65 and 45-50% of population by age 85
- By 2050, 14 million Americans will have Alzheimer's Disease
- Signs include:
  - Loss of recent memory
  - Failure of language skills
  - Failure of visual-spatial orientation
  - Alteration in personality

## The Specific Questions:

Can the presence of Alzheimer's Disease be predicted from clinical and imaging measurements?

Can the stage of Alzheimer's Disease be predicted from MRI images?

# Description of Data Collected:

## Cross-Sectional Study:

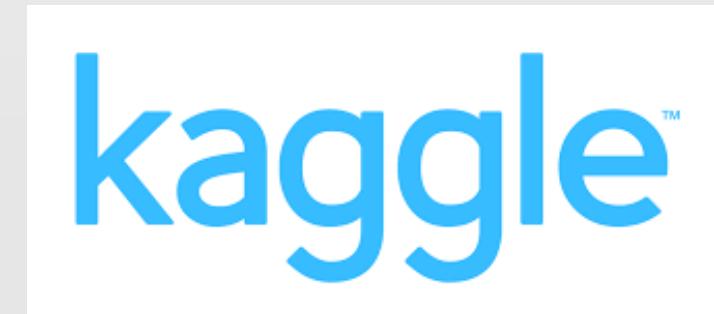
- 436 participants
- Young, Middle Aged, and Older Adults
- Non-demented or Demented
- Clinical, Demographic, and Imaging measurements

## Longitudinal Study:

- 373 participants
- Non-demented and Demented older adults
- Imaged at least 2 times in subsequent visits over at least 1 year
- Clinical, Demographic, and Imaging measurements

## Imaging Study:

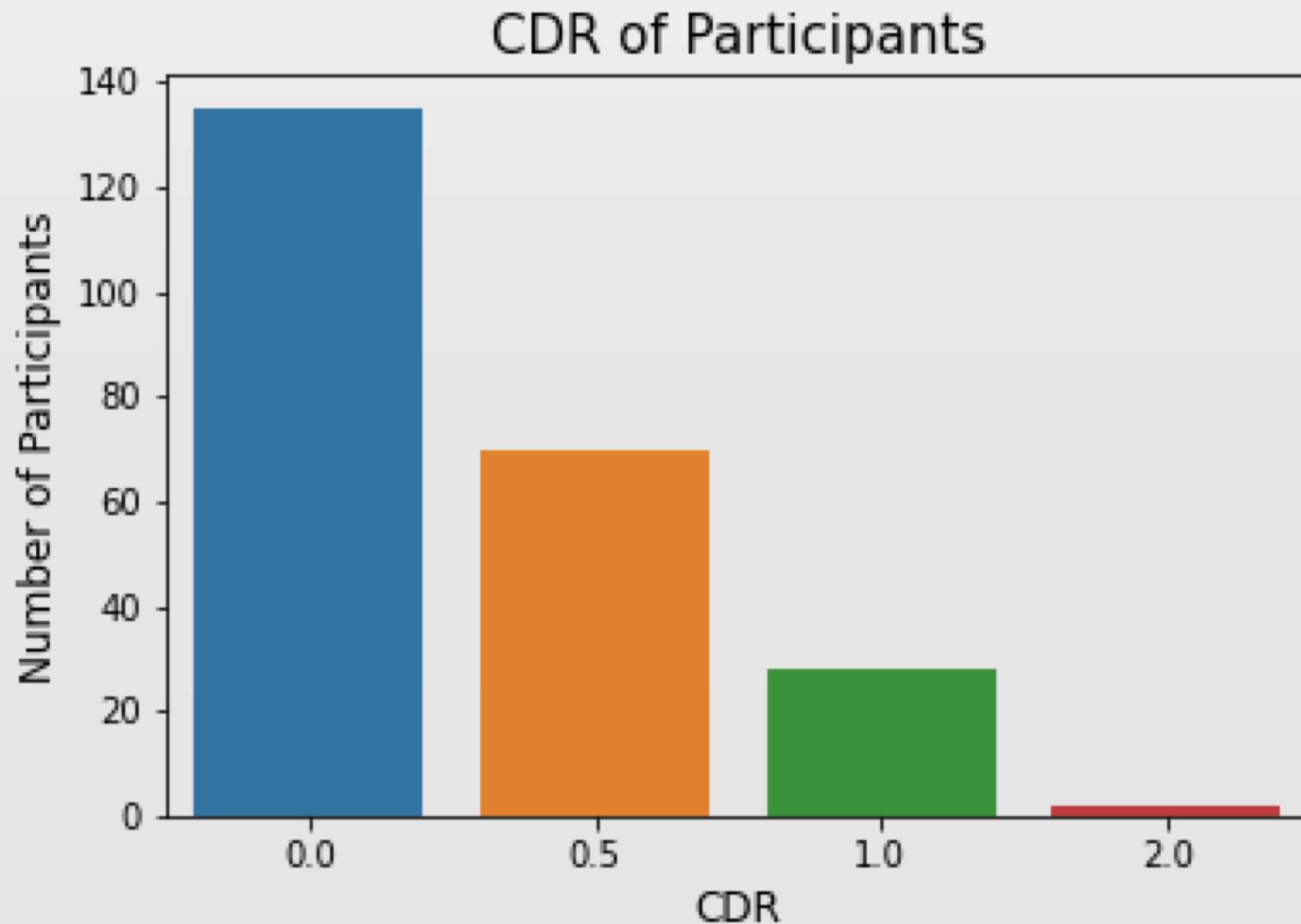
- 6,400 MRI Images
- Non-demented, Very Mild Demented, Mild Demented, Moderate Demented Adults



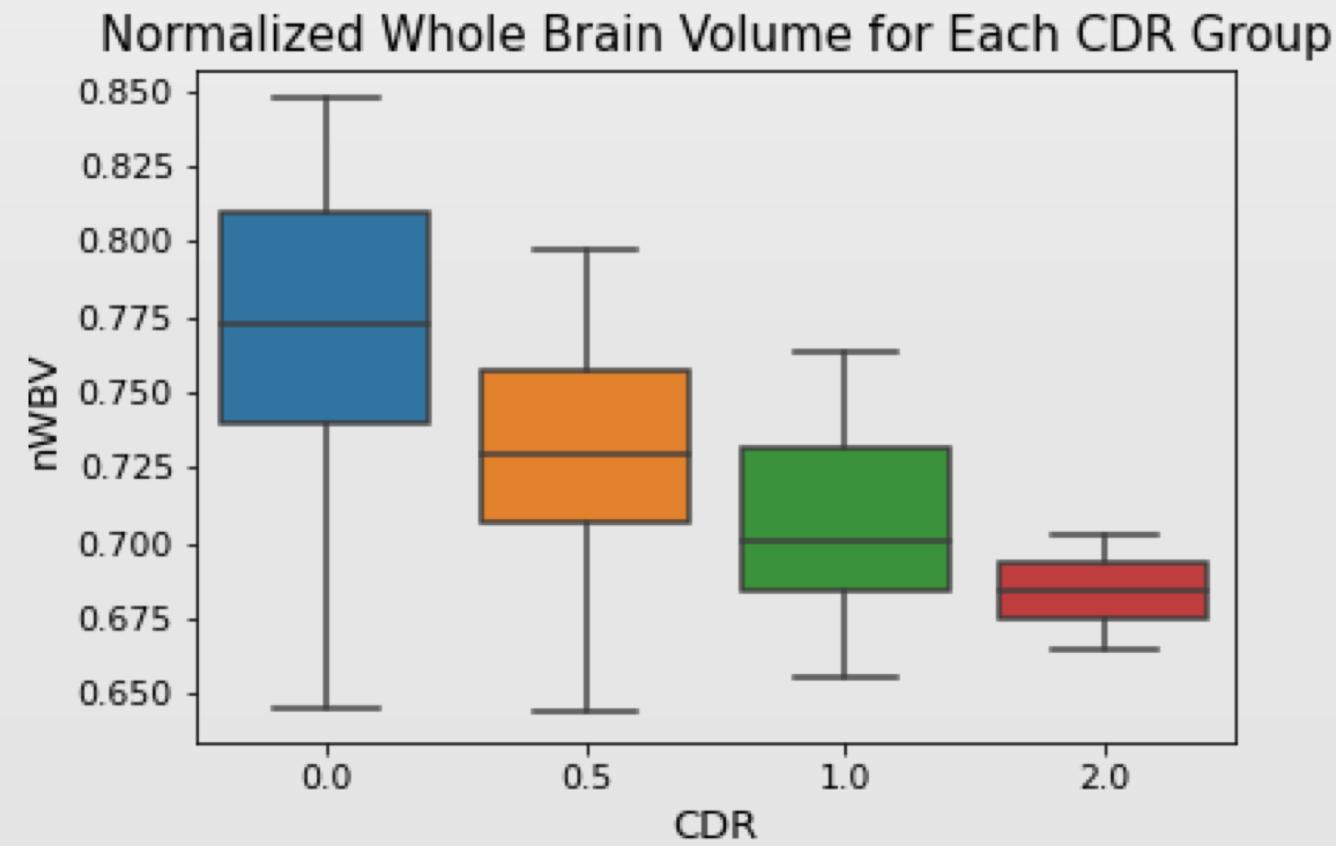
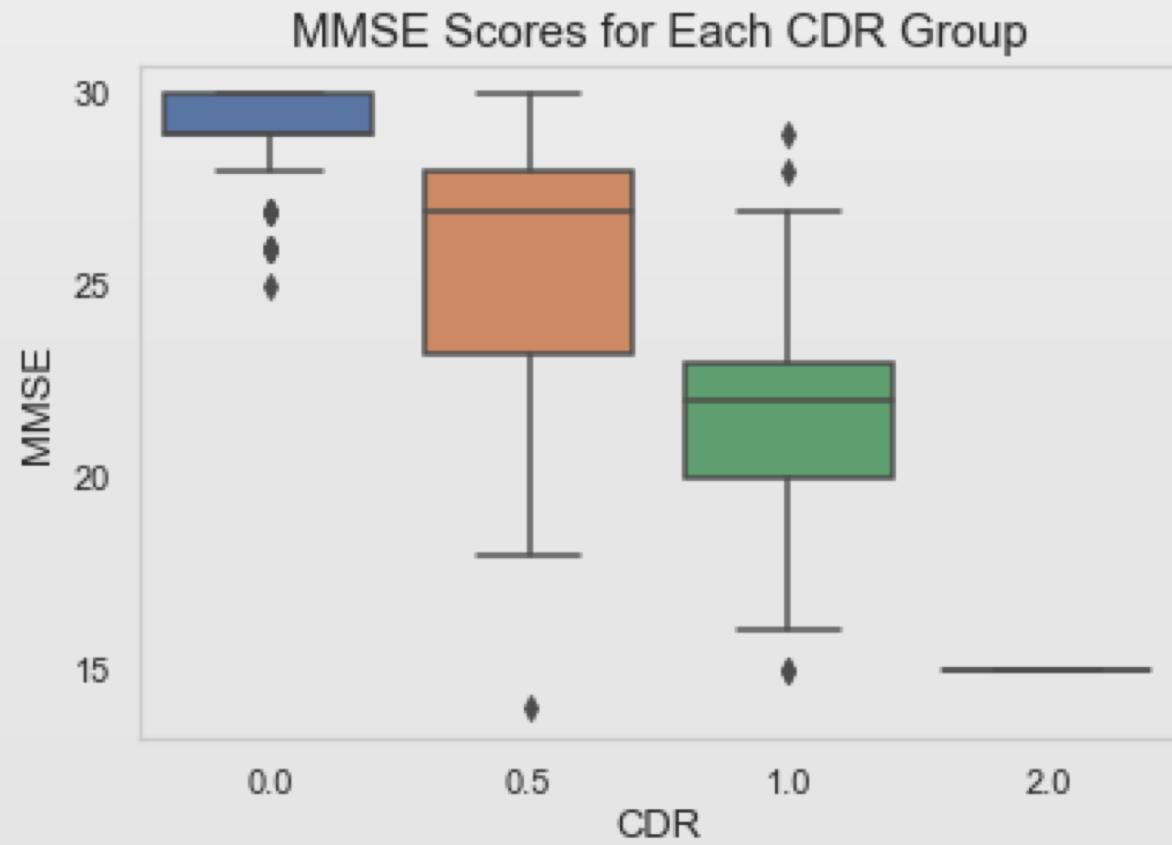
# Description of Data Collected – Cross-Sectional Study:

Variable	Details	Type	
ID	Image identification number	Continuous	
M/F	Gender of patient (M=male and F=female)	Categorical	
Hand	Patient handedness (right or left)	Categorical	
Age	Age of patients (in years)	Continuous	
+	Educ	Level of education (1 = lowest and 5 = highest)	Categorical
+	SES	Socioeconomic status (1 = highest status, 5 = lowest status)	Categorical
+	MMSE	Mini mental state examination (see Table 2)	Continuous/Categorical
CDR	Clinical dementia rating (0 = no dementia, 0.5 = very mild dementia, 1 = mild dementia, 2 = moderate dementia, 3 = severe dementia)	Categorical	
eTIV	Estimated total intracranial volume (cm <sup>3</sup> )	Continuous	
nWBV	Normalized whole brain volumes	Continuous	
ASF	Atlas scaling factor - comparison of eTIV based on normalization of head size	Continuous	
Delay	Delay time before imaging in real time	Continuous	

# Data Analysis Cross-Sectional Study: Initial Findings



# Data Analysis Cross-Sectional Study: Initial Findings

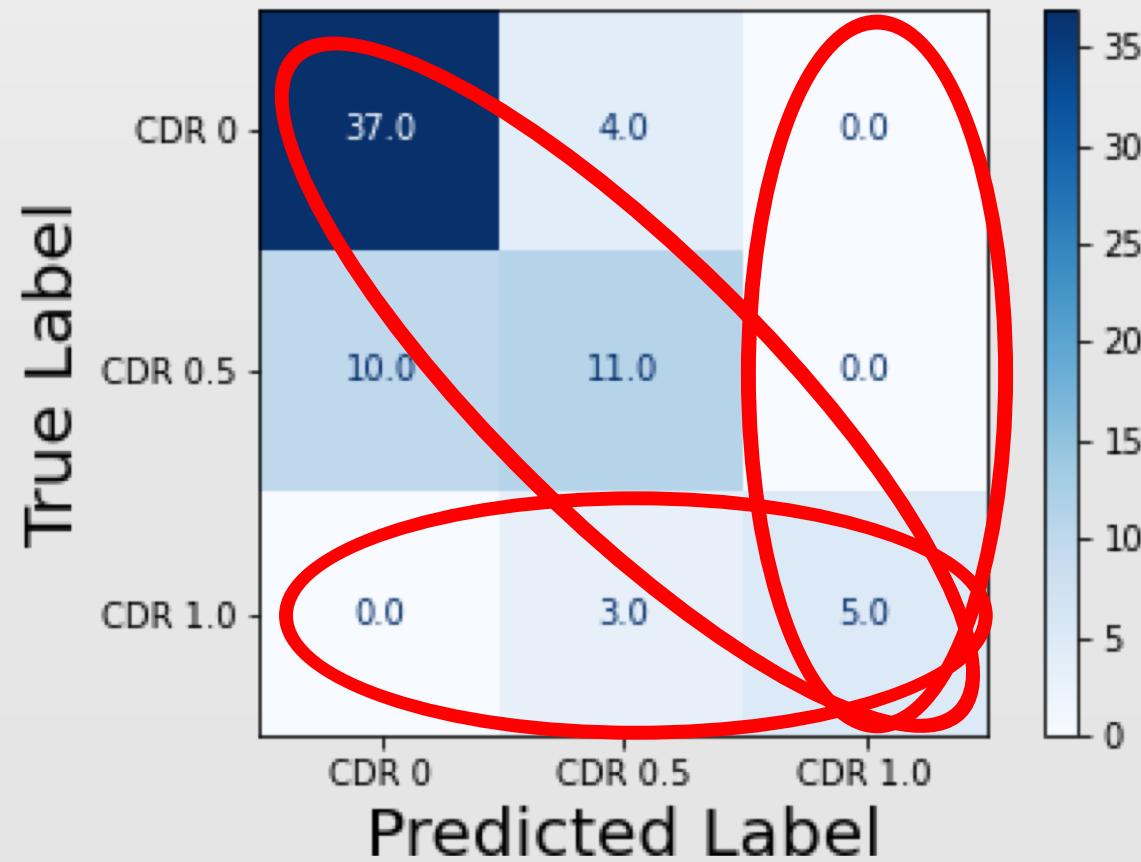


# Data Analysis: Generating Models to Predict Alzheimer's

- Assume that features other than 'CDR' will account for the presence or absence of Alzheimer's
- 4 Models Tested:
  - K-nearest neighbors
  - Random Forest
  - One vs. Rest (Logistic Regression)
  - One vs. Rest (Multi-Layer Perceptron)

# Data Analysis Cross-Sectional Study: Choosing the Model

KNN Model - Confusion Matrix



- **Higher Precision** - What proportion of positive identifications was actually correct?
- **Higher Recall** - What proportion of actual positives was identified correctly?
- **Higher F1 Score** - Combines Precision and recall to measure a model's accuracy
- **Higher Accuracy** - What percentage of the predictions were correct?

# Data Analysis Cross-Sectional Study: Choosing the Model

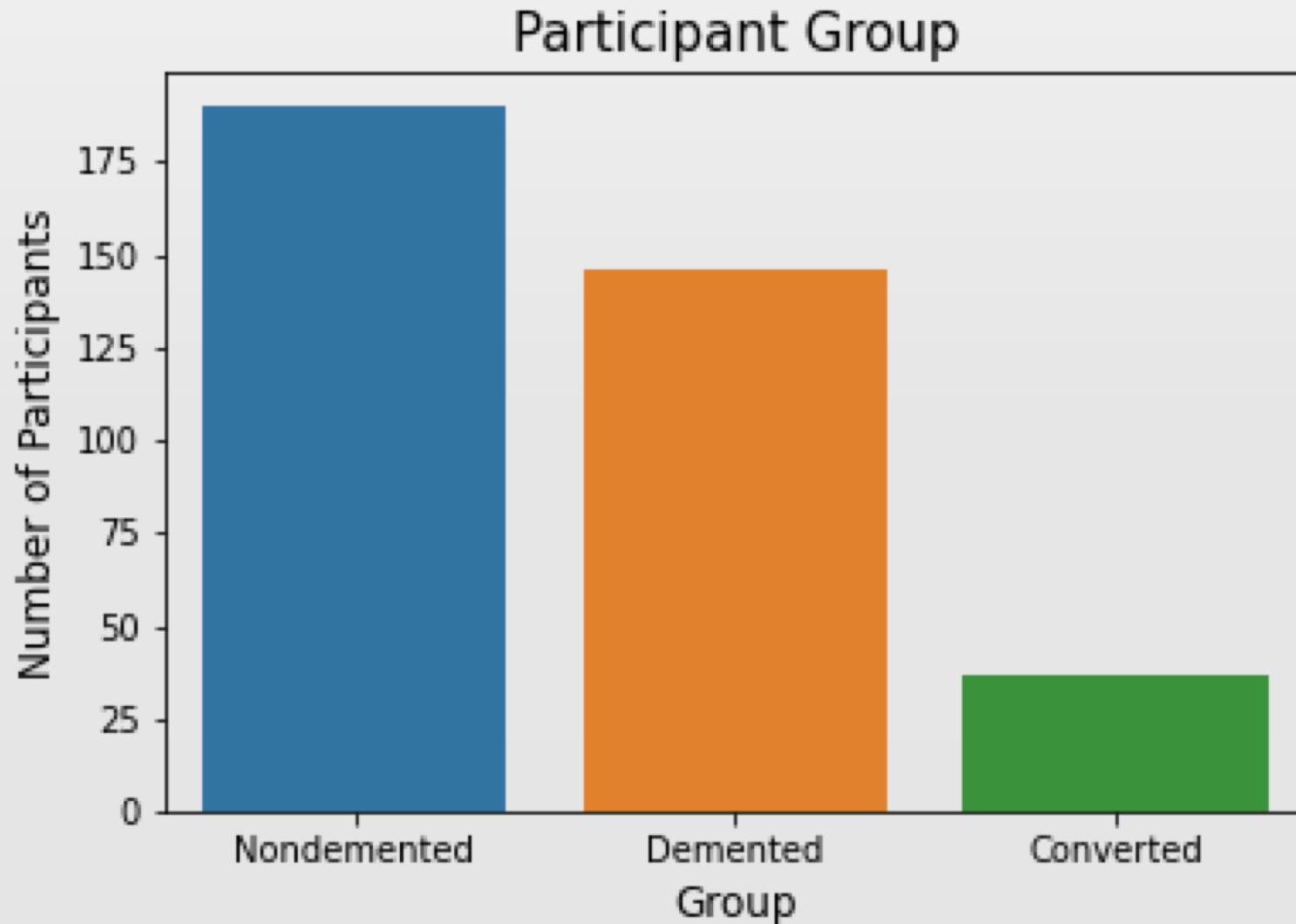
Model	Group	Precision	Recall	F1 score	Accuracy
KNN	CDR 0	0.79	0.9	0.84	
	CDR 0.5	0.61	0.52	0.56	0.76
	CDR 1	1	0.62	0.77	
Random Forest	CDR 0	0.82	0.88	0.85	
	CDR 0.5	0.48	0.48	0.48	0.69
	CDR 1	0.4	0.25	0.31	
OVR(logistic regression)	CDR 0	0.79	0.9	0.84	
	CDR 0.5	0.5	0.43	0.46	0.7
	CDR 1	0.6	0.38	0.46	
OVR(Multi-Layer Perceptron)	CDR 0	0.89	0.8	0.85	
	CDR 0.5	0.57	0.76	0.65	0.76
	CDR 1	0.8	0.5	0.62	

# Description of Data Collected – Longitudinal Study:

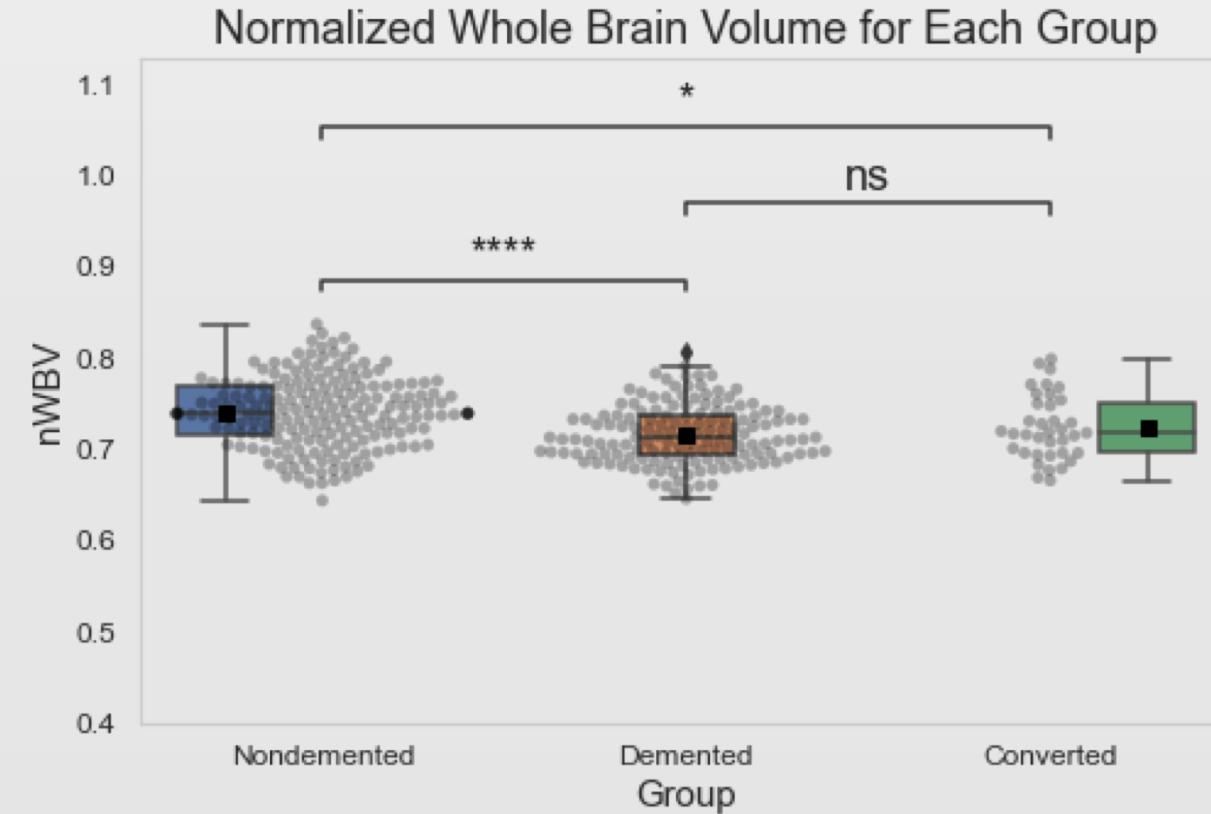
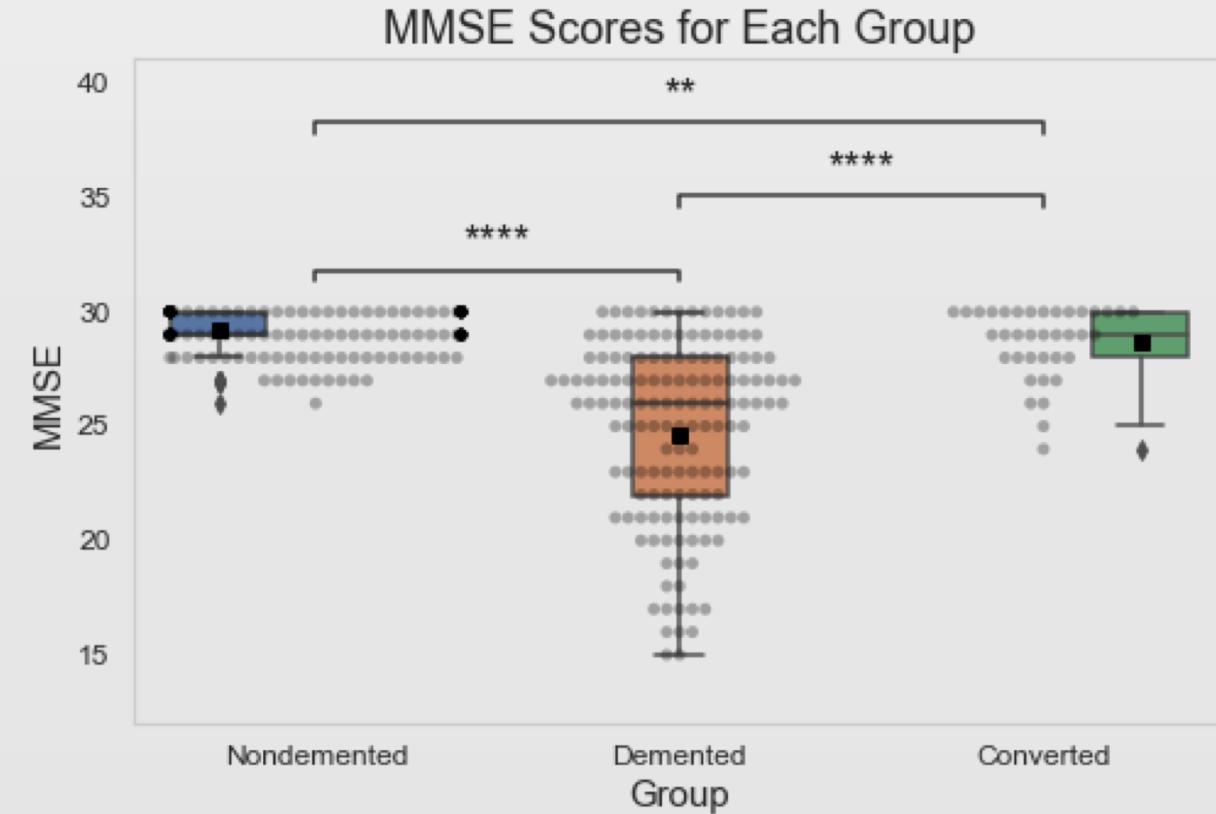
Variable	Details	Type
Subject ID	Subject Identification number	Continuous
MRI ID	MRI identification number	Continuous
Group	Demented, Nondemented, or Converted (developed disease during study)	Categorical
Visit	Which visit image was taken	Categorical
MR Delay	Number of days since first image	Continuous
M/F	Gender of patient (M=male and F=female)	Categorical
Hand	Patient handedness (right or left)	Categorical
Age	Age of patients (in years)	Continuous
EDUC	Number of years of education	Categorical
SES	Socioeconomic status (1 = highest status, 5 = lowest status)	Categorical
MMSE	Mini mental state examination (see Table 2)	Categorical/Continuous
CDR	Clinical dementia rating (0 = no dementia, 0.5 = very mild dementia, 1 = mild dementia, 2 = moderate dementia, 3 = severe dementia)	Categorical
eTIV	Estimated total intracranial volume (cm <sup>3</sup> )	Continuous
nWBV	Normalized whole brain volumes	Continuous
ASF	Atlas scaling factor - comparison of eTIV based on normalization of head size	Continuous



# Data Analysis Longitudinal Study: Initial Findings



# Data Analysis Longitudinal Study: Initial Findings



# Data Analysis: Generating Models to Predict Alzheimer's

- Assume that features other than ‘Group’ will account for the presence or absence of Alzheimer’s
- 4 Models Tested:
  - K-nearest neighbors
  - Random Forest
  - One vs. Rest (Logistic Regression)
  - One vs. Rest (Multi-Layer Perceptron)

# Data Analysis Longitudinal Study: Choosing the Model

Model	Group	Precision	Recall	F1 score	Accuracy
KNN	Non-demented	0.88	1	0.93	
	Demented	0.94	0.94	0.94	0.9
	Converted	1	0.22	0.39	
Random Forest	Non-demented	0.89	0.98	0.93	
	Demented	0.95	1	0.97	0.9
	Converted	0.5	0.11	0.18	
OVR(logistic regression)	Non-demented	0.89	1	0.94	
	Demented	0.95	1	0.97	0.91
	Converted	1	0.11	0.2	
OVR(Multi-Layer Perceptron)	Non-demented	0.9	0.94	0.92	
	Demented	0.97	1	0.99	0.9
	Converted	0.5	0.33	0.4	

# Data Analysis: Important Features for Predicting Alzheimer's

## Cross-sectional Study:

- Mini-Mental State Examination
- Normalized Whole Brain Volume

## Longitudinal Study:

- Clinical Dementia Rating
- Mini-Mental State Examination

# Data Analysis: Imaging Study

Patient Category	Number of jpeg Files
Non-Demented	3200
Very Mild Demented	2240
Mild Demented	896
Moderate Demented	64

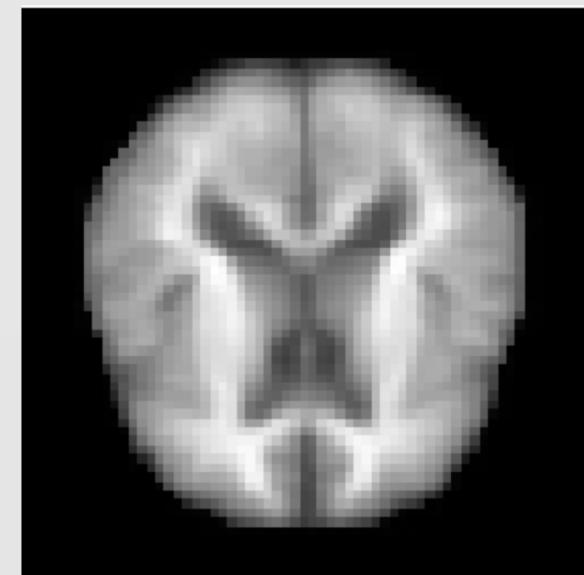
Non-demented



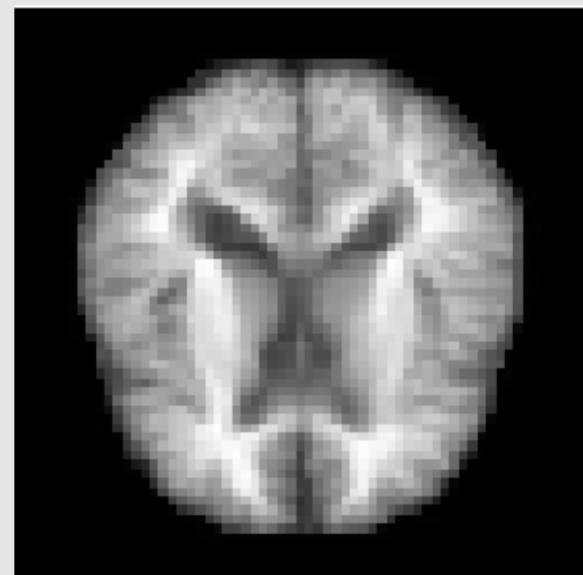
Very Mild Demented



Mild Demented

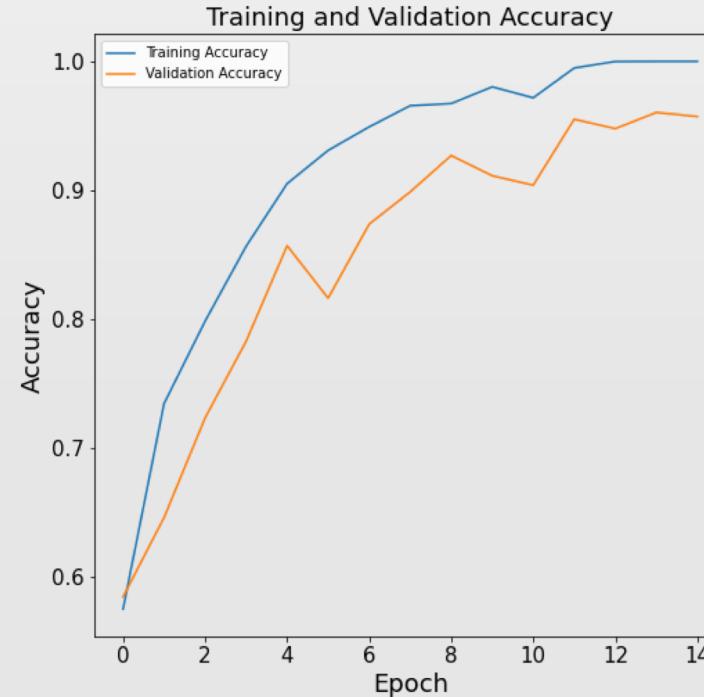


Moderate Demented



# Data Analysis: Generating Model to Predict Alzheimer's Stage

1 Model Tested: Convolutional Neural Network (CNN)



Validation accuracy: 96%

Group	Precision	Recall	F1 score	Accuracy
Non-Demented	0.97	0.97	0.97	
Very Mild Demented	0.94	0.96	0.95	0.96
Mild Demented	0.97	0.91	0.94	
Moderate Demented	1	1	1	

# Future Research and Recommendations

- **Cross-sectional Study** – KNN Model
  - More participants with CDR 2 and 3
- **Longitudinal Study** – One vs Rest (MLP) Model
  - More participants in Converted Group
- **Imaging Study** – CNN Model
  - More participants in Moderate Group
  - Follow up imaging
  - Test other models (Xception or ResNet50)
- Can be used by physicians , researchers, and patients to:
  - Predict presence of Alzheimer's Disease in patients
  - Predict stage of Alzheimer's Disease in patients
  - Use imaging to determine if drug(s) of interest can slow progression of Alzheimer's Disease