

# Data Analysis for Alzheimer's Disease

Presented by **Yueqi Ren**  
Graduate Student in the Stark Lab and Shahbaba Group

Summer 2022

# Overview of Topics

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# Overview of Topics

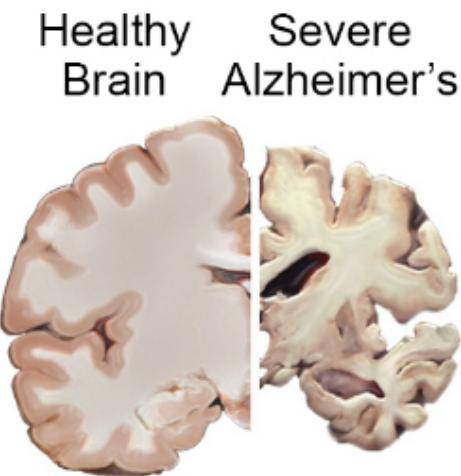
Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

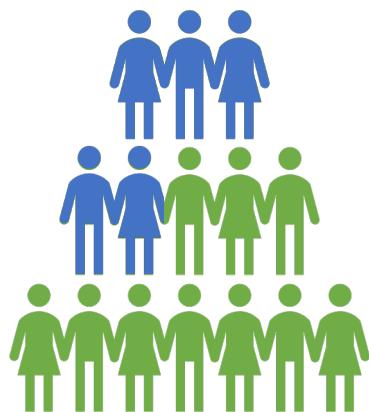
- NACC data set
- What my work focuses on

# What is late-onset Alzheimer's Disease (LOAD)?



LOAD is the dominant form of AD,  
affecting people  $\geq 65$  years old

*NIH National Institute on Aging*



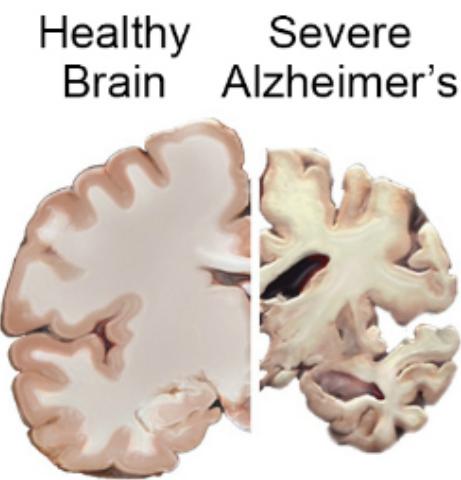
1/3 American seniors pass away  
with AD or dementia

*Alzheimer's Association Facts &  
Figures, 2021*



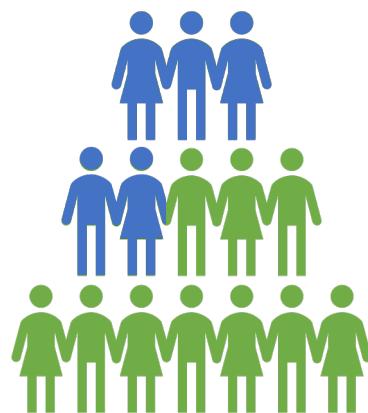
No cure for AD exist, but  
treatment is possible

# What is late-onset Alzheimer's Disease (LOAD)?



LOAD is the dominant form of AD,  
affecting people  $\geq 65$  years old

*NIH National Institute on Aging*



1/3 American seniors pass away  
with AD or dementia

*Alzheimer's Association Facts &  
Figures, 2021*



# Alzheimer's in Everyday Life



Breakout Groups!

- What do you know about AD?

# Alzheimer's in Everyday Life



Breakout Groups!

- What do you know about AD?
- What *symptoms* are associated with AD?

# Alzheimer's in Everyday Life

- What do you know about AD?
- What *symptoms* are associated with AD?
  - Forgetfulness
  - Misplacing items
  - Not recognizing friends or family
  - Being wrong about the day or year
  - Severe difficulty finding the right words
  - Becoming unable to navigate or drive around
  - ...

# Alzheimer's in Everyday Life



Breakout Groups!

- What do you know about AD?
- What *symptoms* are associated with AD?
- What are some risk factors of AD?

# Alzheimer's in Everyday Life

- What do you know about AD?
- What *symptoms* are associated with AD?
- What are some risk factors of AD?
  - Age
  - Genetics
  - Sex
  - Comorbidities
  - Health disparities
  - ...

# How the definition of AD has evolved

-1980s

You belong to either one of the groups

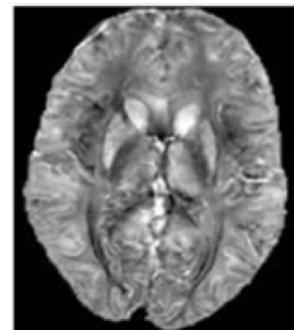
Healthy  
Cognition

Alzheimer's  
Disease

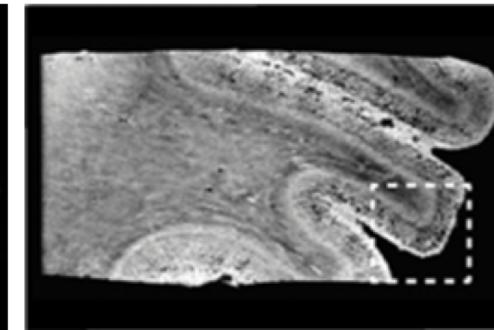
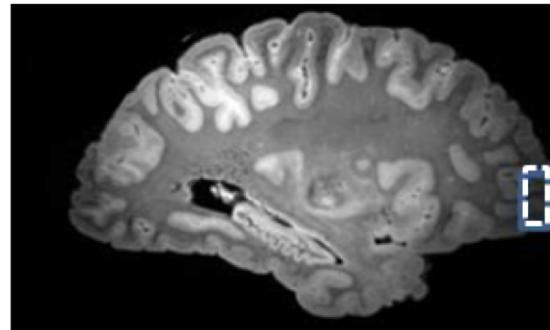
Why was the cutoff so binary?

# How the definition of AD has evolved

*In vivo* MRI

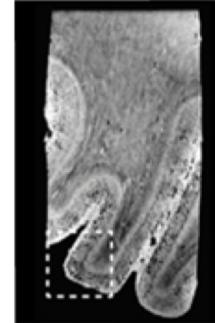


*Post mortem* MRI



Co-registration and analysis pipeline

MRI



Iron



Myelin



Tau

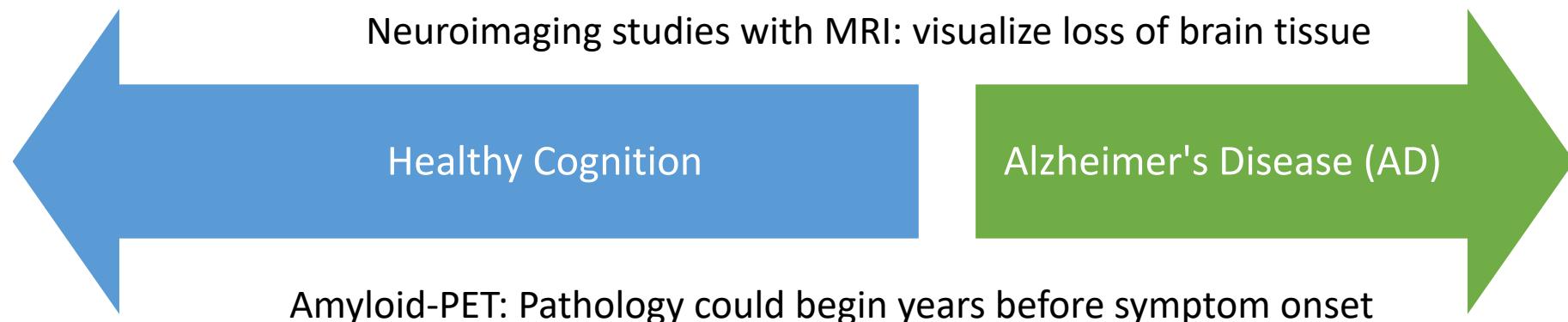


Abeta

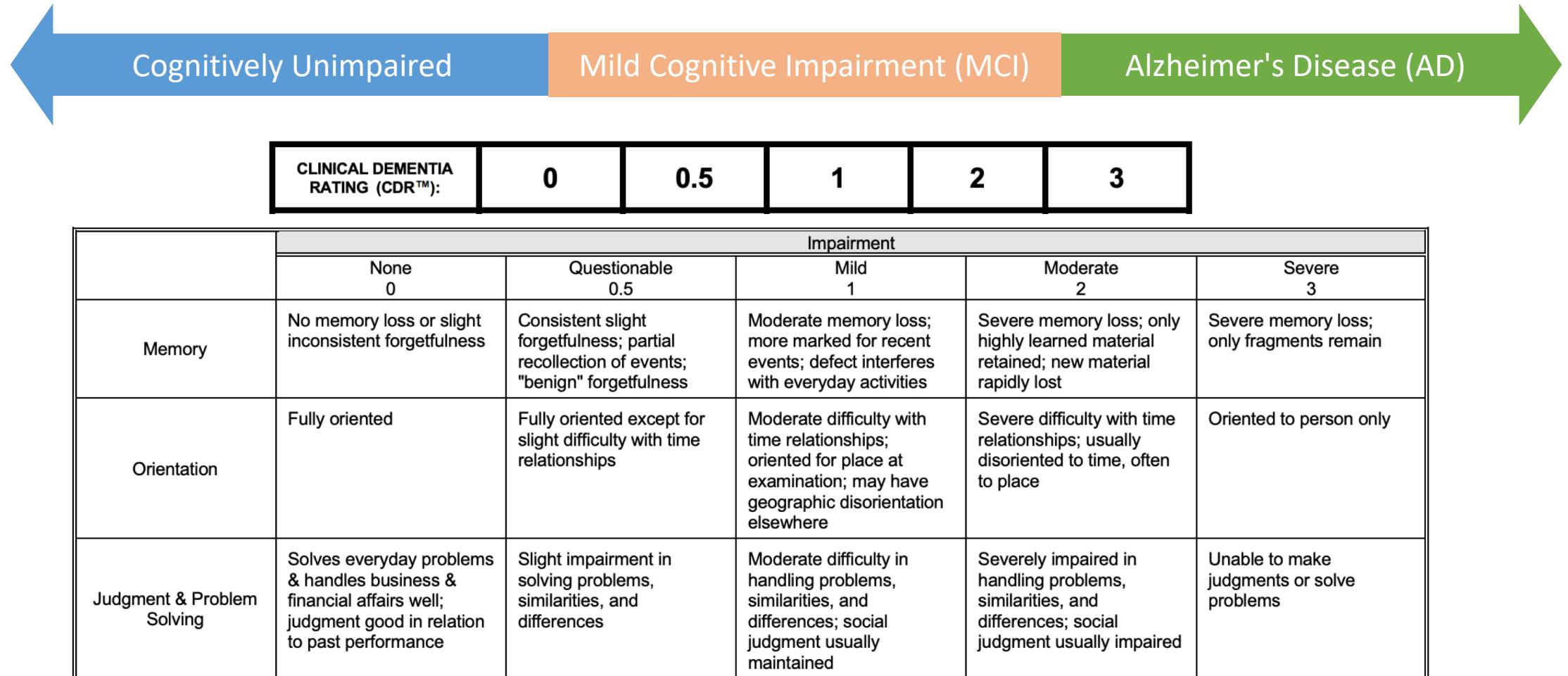


We rely on postmortem analysis of brain tissue to confirm AD diagnosis

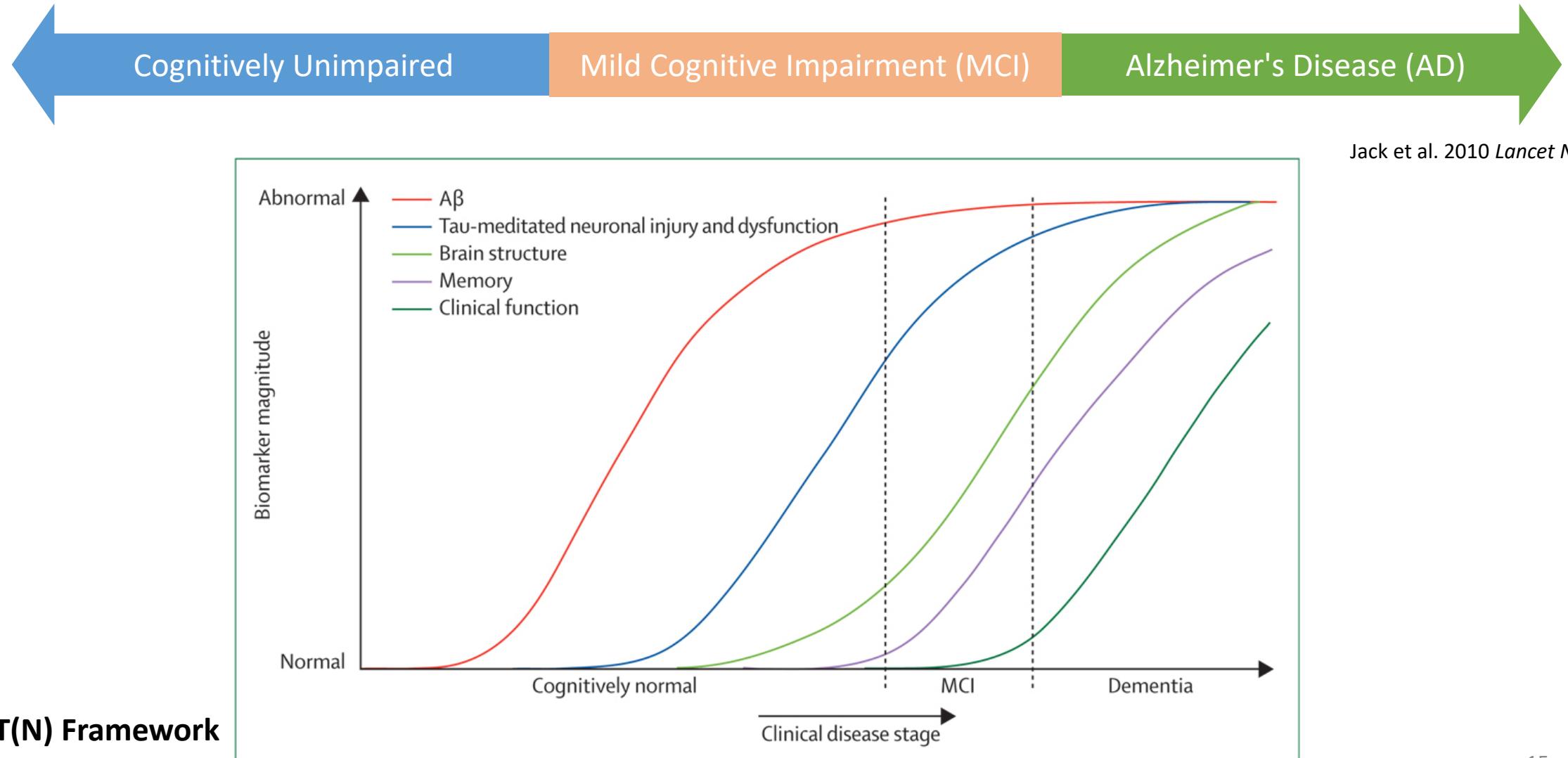
# The Alzheimer's continuum



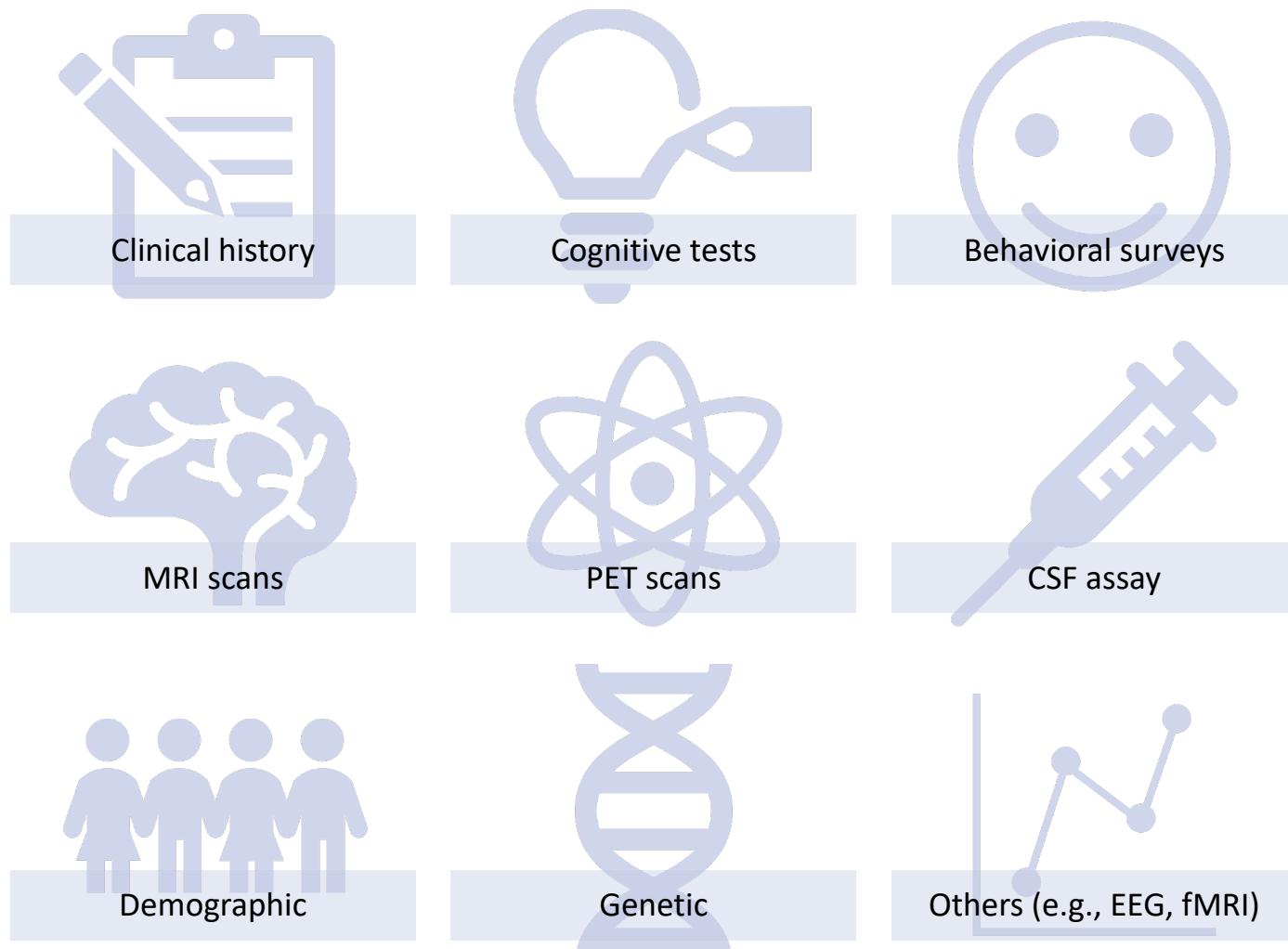
# The Alzheimer's continuum



# The Alzheimer's continuum

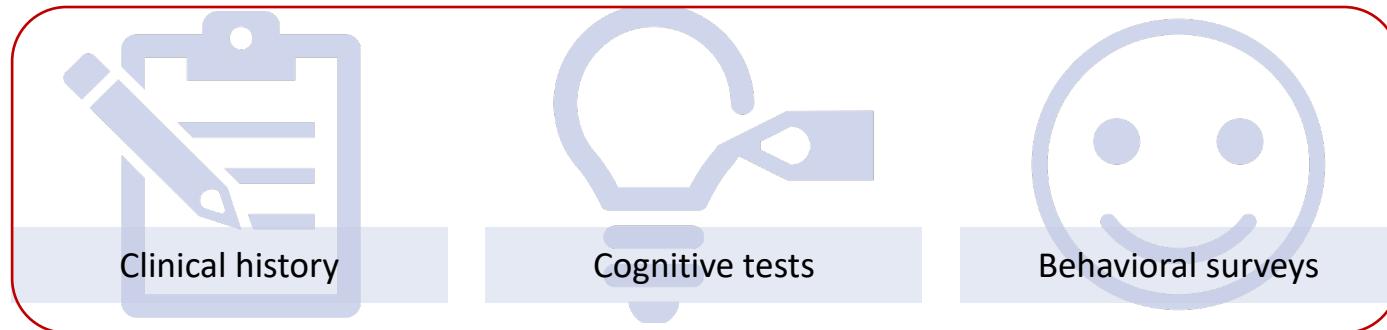


# Types of Data for Evaluating AD

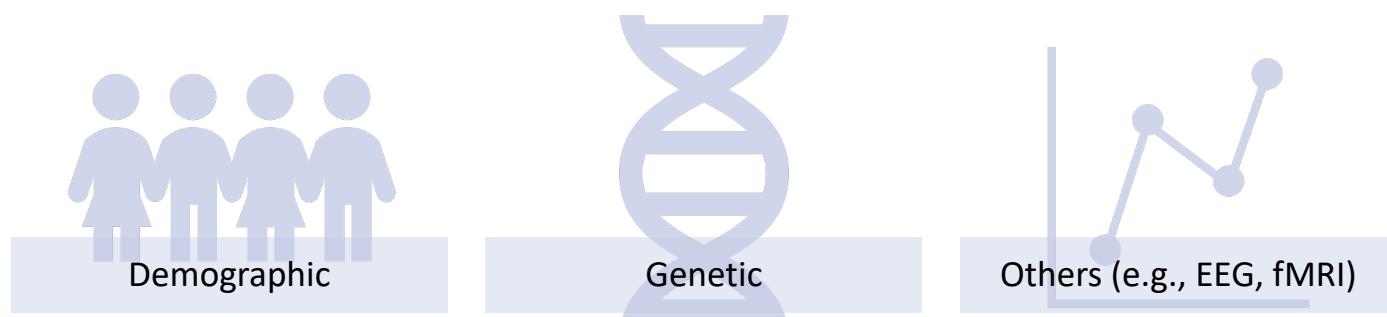


# Types of Data for Evaluating AD

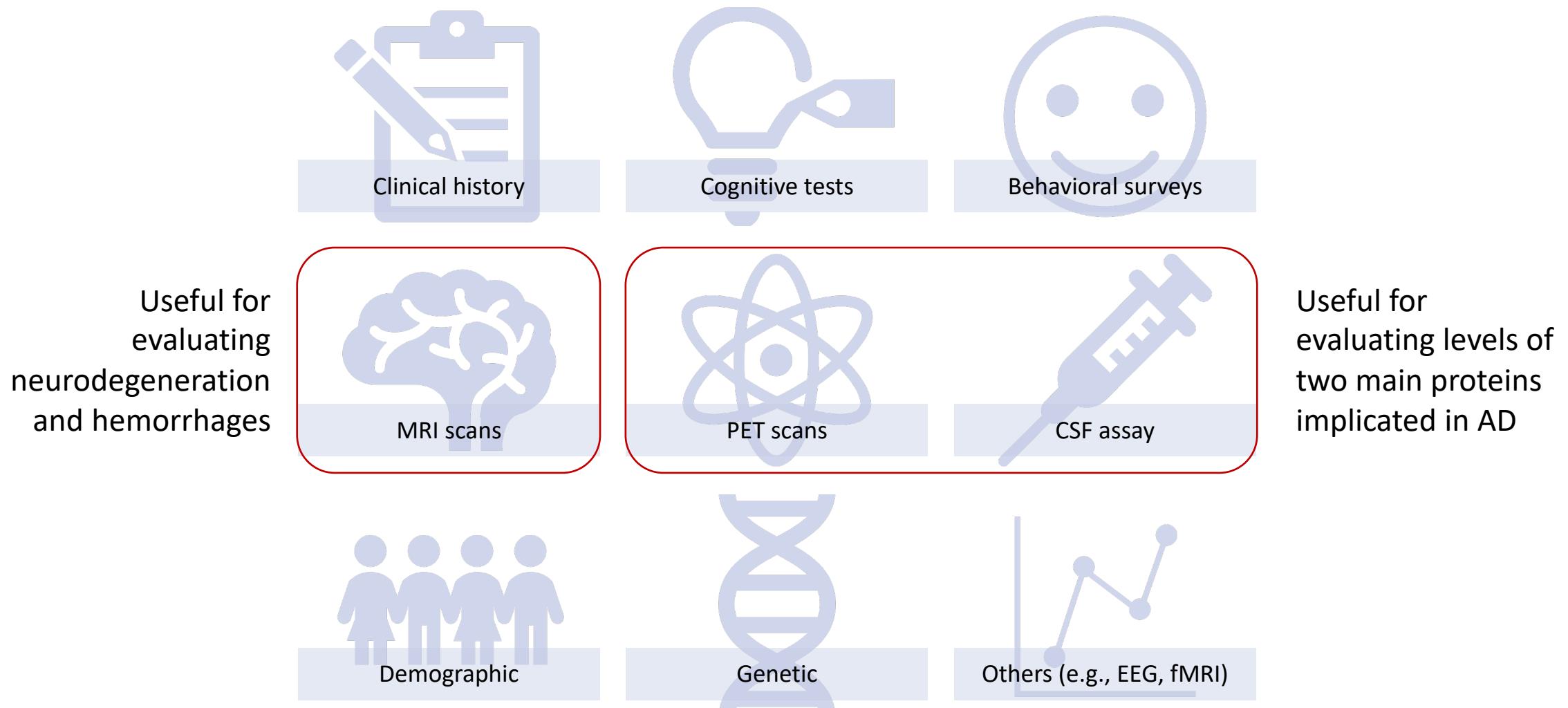
Obtained by the physician and/or neuropsychologist



Many tests are compared to the “normative” values found from healthy subjects



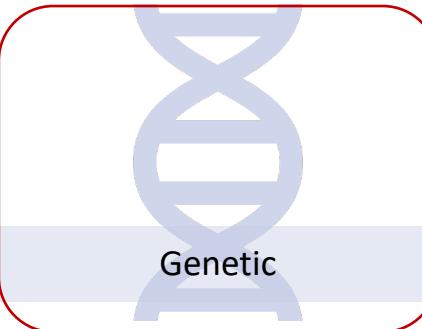
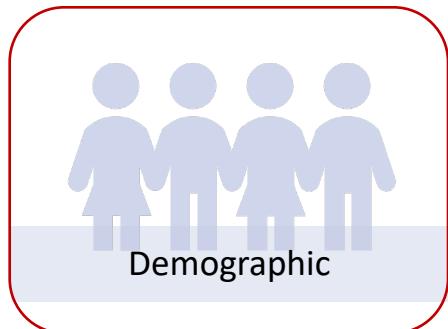
# Types of Data for Evaluating AD



# Types of Data for Evaluating AD



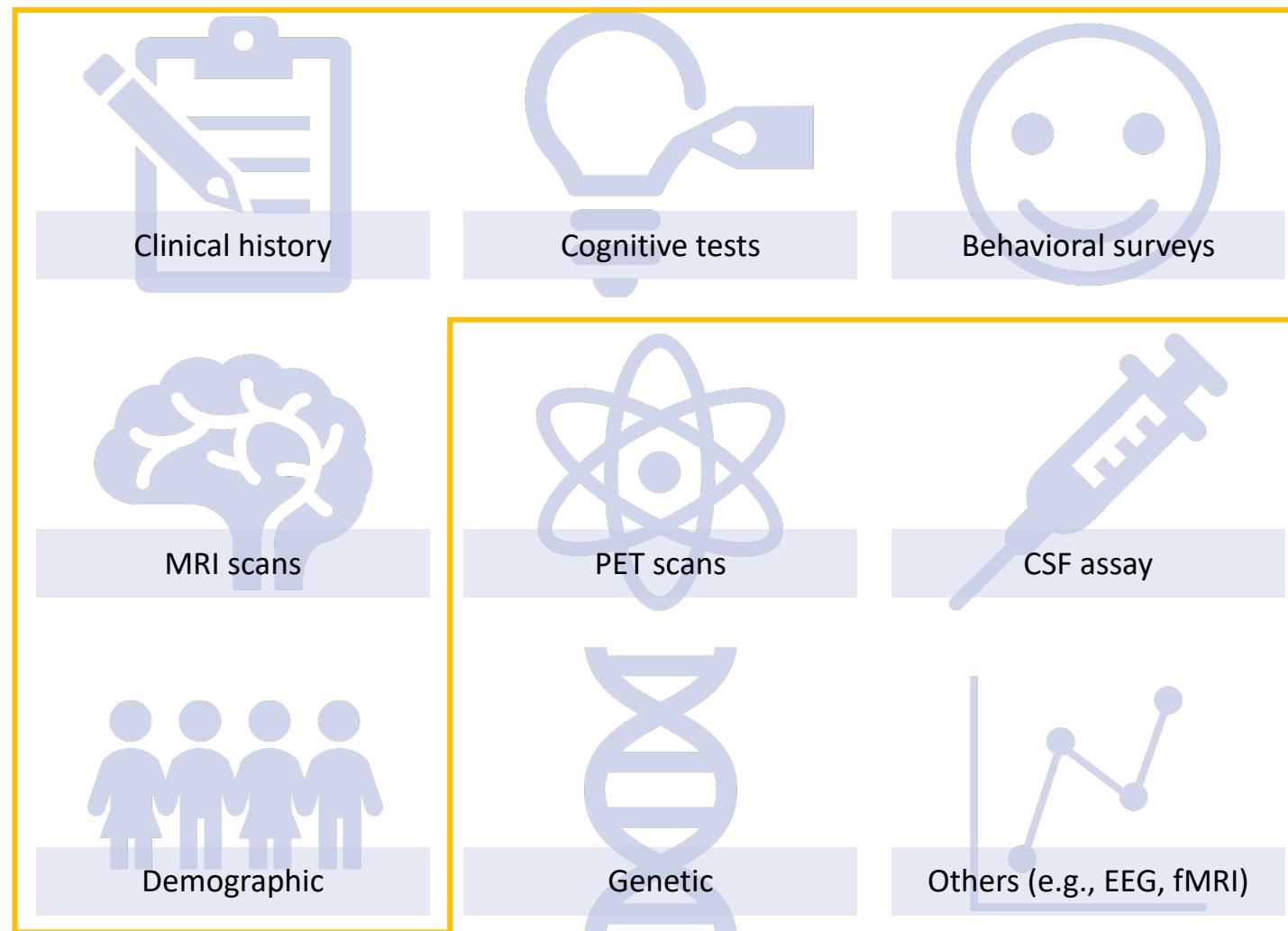
Age is strongly associated with AD; education may be protective



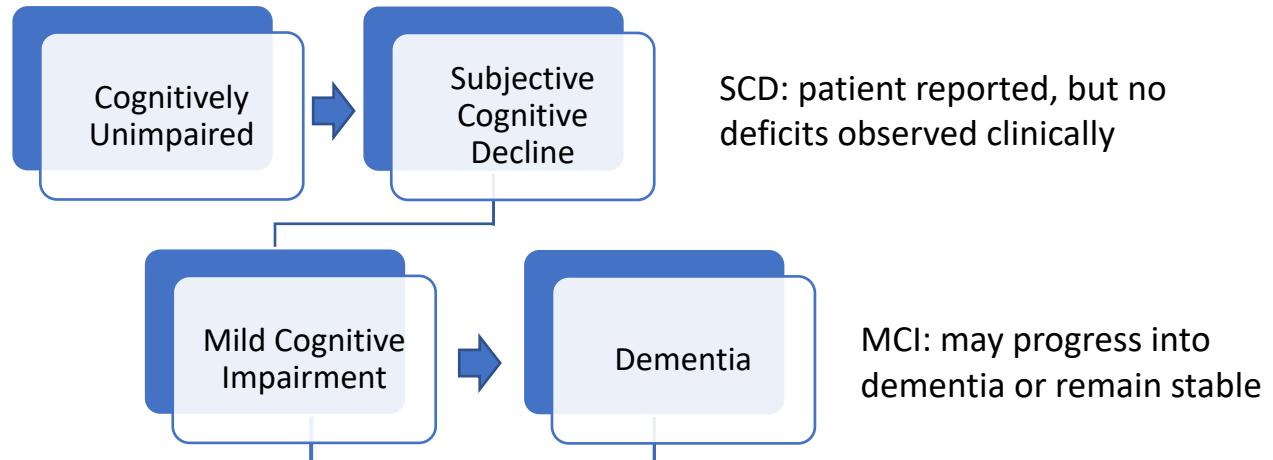
Some forms of AD are autosomal dominant (early onset); for most cases, no single gene is responsible, but **ApoE** has been shown to be strongly linked to AD

# Types of Data for Evaluating AD

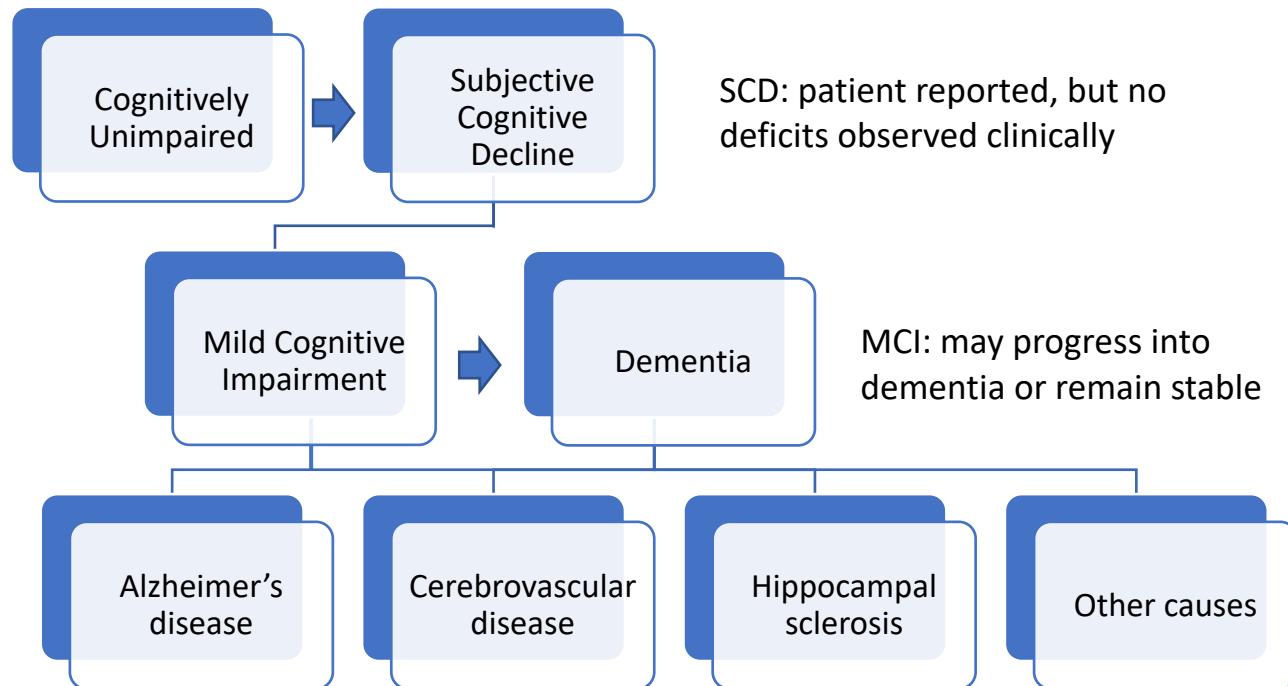
Our focus will  
be on these  
types of data



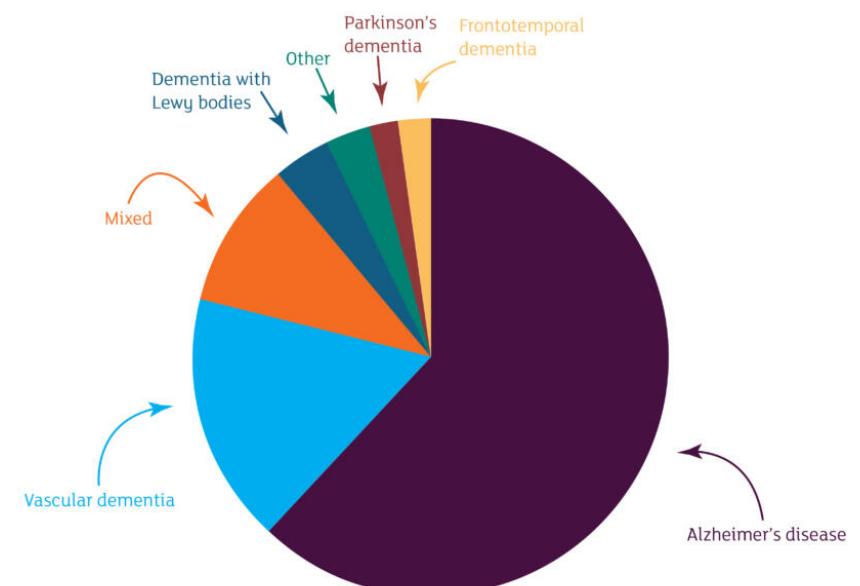
# Alzheimer's Continuum: Clinical Focus



# Alzheimer's Disease and Related Dementias



Causes of dementia



Causes of cognitive impairment are heterogeneous and not mutually exclusive

- Some patients have more than one underlying neuropathology

Alzheimer's Research UK

# Summary

- ✓ Prevalence of AD
- ✓ How AD presents clinically
- ✓ Definition of AD
- ✓ Risk factors for AD
- ✓ Types of data available to clinicians and researchers
- ✓ Relationship of AD and related dementias

# Any Questions?

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# Next Up

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# Important AD Topics

## Improving diagnostic accuracy

- To assist physicians
- To screen more patients given limited access to experts

## Understanding relationships

- Some tests are more expensive and less accessible
- What tests should be conducted more routinely?

## Earlier detection of AD

- Prior to obvious symptoms
- Better clinical trials

# What skills will you learn?

- One sample statistics
  - Inference for 1 sample
    - Hypothesis test and confidence intervals for the mean
- Two sample statistics
  - Inference for 2 samples
    - Hypothesis test and confidence intervals for comparing means, correlation, and simple linear regression

# What else will you learn?

- Multiple regression
- Logistic regression for binary outcomes
- Resampling methods
  - Bootstrapping
- How to think like a data scientist!

# Using these skills, you can answer key Qs

## Improving AD diagnosis

- What variables are most informative of diagnosis?

## Understanding AD pathology

- How are important clinical features related to each other?

# Any Questions?

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# Next Up

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# NACC Data Set



National Alzheimer's  
Coordinating Center  
(NACC)

- **Uniform Data Set (UDS):** demographic, behavioral, neuropsychological testing, clinical information
- **sMRI:** image-derived features from structural MRI scans

<https://naccdata.org/>

# Clinical Data

- NACC ID: unique identifier for each subject
- DIAGNOSIS: clinician diagnosis of either
  - Normal cognition (0)
  - Mild cognitive impairment (1)
  - Dementia due to AD (2)

# Clinical Data

Feature Name	Definition	Data Type	Possible Values
AGE	Subject's age (years)	Numeric	18 - 120
EDUC	Subject's number of years of education	Numeric	0 - 36
FEMALE	Subject's sex	Categorical	0 = male 1 = female
BPSYS	Subject blood pressure (sitting), systolic	Numeric	70 - 230
BPDIAS	Subject blood pressure (sitting), diastolic	Numeric	30 - 140
HEIGHT	Subject's height (inches)	Numeric	36.0 - 87.9
WEIGHT	Subject's weight (lbs)	Numeric	50 - 400
HRATE	Subject resting heart rate (pulse)	Numeric	33 - 160

# Behavioral Survey Data

Feature Name	Definition	Data Type	Possible Values
NACCGDS	Total Geriatric Depression Scale (GDS) Score	Numeric	0 - 15
CDRGLOB	Global Clinical Dementia Rating (CDR) Score	Numeric	<p>0.0 = No impairment 0.5 = Questionable impairment 1.0 = Mild impairment 2.0 = Moderate impairment 3.0 = Severe impairment</p>

# Behavioral Survey Data

Feature Name	Definition	Data Type	Possible Values
TRAVEL	In the past four weeks, did the subject have any difficulty or need help with: Traveling out of the neighborhood, driving, or arranging to take public transportation	Numeric	0 = Normal 1 = Has difficulty, but does by self 2 = Requires assistance 3 = Dependent 8 = Not applicable (e.g., never did)
REMDATES	In the past four weeks, did the subject have any difficulty or need help with: Remembering appointments, family occasions, holidays, medications	Numeric	0, 1, 2, 3, 8
PAYATTN	In the past four weeks, did the subject have any difficulty or need help with: Paying attention to and understanding a TV program, book, or magazine	Numeric	0, 1, 2, 3, 8
MEALPREP	In the past four weeks, did the subject have any difficulty or need help with: Preparing a balanced meal	Numeric	0, 1, 2, 3, 8
EVENTS	In the past four weeks, did the subject have any difficulty or need help with: Keeping track of current events	Numeric	0, 1, 2, 3, 8
MEALPREP	In the past four weeks, did the subject have any difficulty or need help with: Preparing a balanced meal	Numeric	0, 1, 2, 3, 8
SHOPPING	In the past four weeks, did the subject have any difficulty or need help with: Shopping alone for clothes, household necessities, or groceries	Numeric	0, 1, 2, 3, 8
GAMES	In the past four weeks, did the subject have any difficulty or need help with: Playing a game of skill such as bridge or chess, working on a hobby	Numeric	0, 1, 2, 3, 8
STOVE	In the past four weeks, did the subject have any difficulty or need help with: Heating water, making a cup of coffee, turning off the stove	Numeric	0, 1, 2, 3, 8
SHOPPING	In the past four weeks, did the subject have any difficulty or need help with: Shopping alone for clothes, household necessities, or groceries	Numeric	0, 1, 2, 3, 8
BILLS	In the past four weeks, did the subject have any difficulty or need help with: Writing checks, paying bills, or balancing a checkbook	Numeric	0, 1, 2, 3, 8
TAXES	In the past four weeks, did the subject have any difficulty or need help with: Assembling tax records, business affairs, or other papers	Numeric	0, 1, 2, 3, 8

# Behavioral Survey Data

Feature Name	Definition	Data Type	Possible Values
APPSEV	Appetite and eating severity	Numeric	1 = Mild (noticeable, but not a significant change) 2 = Moderate (significant, but not a dramatic change) 3 = Severe (very marked or prominent; a dramatic change) 8 = Not applicable, no appetite or eating problems reported
NITESEV	Nighttime behaviors severity	Numeric	1, 2, 3, 8
MOTSEV	Motor disturbance severity	Numeric	1, 2, 3, 8
IRRSEV	Irritability or lability severity	Numeric	1, 2, 3, 8
DISNSEV	Disinhibition severity	Numeric	1, 2, 3, 8
APASEV	Apathy or indifference severity	Numeric	1, 2, 3, 8
ELATSEV	Elation or euphoria severity	Numeric	1, 2, 3, 8
ANXSEV	Anxiety severity	Numeric	1, 2, 3, 8
DEPDSEV	Depression or dysphoria severity	Numeric	1, 2, 3, 8
AGITSEV	Agitation or aggression severity	Numeric	1, 2, 3, 8
HALLSEV	Hallucinations severity	Numeric	1, 2, 3, 8
DELSEV	Delusions severity	Numeric	1, 2, 3, 8

# Neuropsychological Testing Data

Feature Name	Definition	Data Type	Possible Values
ANIMALS	Total number of animals named in 60 seconds	Numeric	0 - 77
TRAILA	Trail Making Test Part A - Total number of seconds to complete	Numeric	0 - 150
TRAILB	Trail Making Test Part B - Total number of seconds to complete	Numeric	0 - 300
DIGIF	Digit span forward trials correct	Numeric	0 - 12
MEMUNITS	Logical Memory IIA - Delayed - Total number of story units recalled	Numeric	0 - 25
NACCMMSE	Total Mini-Mental State Exam (MMSE) score	Numeric	0 - 30

# MRI Data

Feature Name	Measurement Type	Definition	Data Type
NACCICV	Regional gray matter volumes	Total intracranial volume (cc)	Numeric
CSFVOL	Regional gray matter volumes	Total brain cerebrospinal fluid volume (cc)	Numeric
LHIPPO	Regional gray matter volumes	Segmented left hippocampus volume (cc)	Numeric
RHIPPO	Regional gray matter volumes	Segmented right hippocampus volume (cc)	Numeric
FRCORT	Regional gray matter volumes	Segmented total frontal lobe cortical gray matter volume (cc)	Numeric
LPARCORT	Regional gray matter volumes	Segmented left parietal lobe cortical gray matter volume (cc)	Numeric
RPARCORT	Regional gray matter volumes	Segmented right parietal lobe cortical gray matter volume (cc)	Numeric
LTEMPCOR	Regional gray matter volumes	Segmented left temporal lobe cortical gray matter volume (cc)	Numeric
RTEMPCOR	Regional gray matter volumes	Segmented right temporal lobe cortical gray matter volume (cc)	Numeric
LCAC	Regional gray matter volumes	Left caudal anterior cingulate gray matter volume (cc)	Numeric
RCAC	Regional gray matter volumes	Right caudal anterior cingulate gray matter volume (cc)	Numeric
LENT	Regional gray matter volumes	Left entorhinal gray matter volume (cc)	Numeric
RENT	Regional gray matter volumes	Right entorhinal gray matter volume (cc)	Numeric
LPARHIP	Regional gray matter volumes	Left parahippocampal gray matter volume (cc)	Numeric
RPARHIP	Regional gray matter volumes	Right parahippocampal gray matter volume (cc)	Numeric
LPOSCIN	Regional gray matter volumes	Left posterior cingulate gray matter volume (cc)	Numeric
RPOSCIN	Regional gray matter volumes	Right posterior cingulate gray matter volume (cc)	Numeric

# Any Questions?

Background on Alzheimer's disease (AD)

Goals of the Course

Description of the Data

- NACC data set
- What my work focuses on

# Summary

- ✓ What is AD?
- ✓ What types of data do you have?
- ✓ What skills will you learn?
- ✓ What questions could you answer?

# Any Questions?

- ✓ What is AD?
- ✓ What types of data do you have?
- ✓ What skills will you learn?
- ✓ What questions could you answer?