# Novel tools for analysis and visualization of longitudinal data

**Sudeshna Das, PhD & Deborah Blacker, MD, ScD**

**Massachusetts Alzheimer's Disease Research Center**

**FALL ADC MEETING 2019**

# Unusual ~~Novel~~ tools for analysis and visualization of longitudinal data

**Sudeshna Das, PhD & Deborah Blacker, MD, ScD**

**Massachusetts Alzheimer's Disease Research Center**

**FALL ADC MEETING 2019**

# ~~Novel~~ Tools for analysis and visualization of longitudinal data

**Sudeshna Das, PhD & Deborah Blacker, MD, ScD**

**Massachusetts Alzheimer's Disease Research Center**

**FALL ADC MEETING 2019**

# Background & Objectives

- MADRC) has acquired longitudinal UDS data on ~1600 participants for >10 years
- Data & Clinical Core collaborate on multiple research agenda, such as:
  - Selecting samples for specific studies
  - Understanding heterogeneity in cohort
- Case Study:
  - Develop tool to select study participants with specific cognitive trajectories
  - Explore relationship of CDR and neuropsychological tests
  - Characterize heterogeneity of cognitive trajectories

# Example from biomarker study

- Select plasma samples from subjects with pre-specified trajectories:
  1. Normal cognition over time
  2. Pre-clinical MCI: normal subjects that decline to MCI
  3. MCI stable: amnestic MCI that remain stable
  4. MCI decliners: amnestic MCI that decline to dementia
  5. Dementia with AD pathology
- Samples were ***manually*** annotated with these 5 categories
  - ~30 subjects were selected for each category
- Our goal
  - Visualize the trajectories to help Identify potential inconsistencies or outliers
  - Develop classification algorithms to automatically label subjects

# Visualization of longitudinal cognitive within each manually assigned category

**Normal**

9513 N : N : N : N : N : N : N
9792 N : N : N : N : N : N : N : N : N : N : N
9989 N : N : N : N : N : N : N
10081 N : N : N : N : N : N : N : N : N : N : N
9930 N : N : N : N : N : N : N : N : N : N : N
9551 N : N : N : SCC : N : N : N : N : N : N
9803 N : N : N : N : N : N : N : N : SCC : N
10090 N : N : N : N : N : N : N : N : N : N : N
9606 N : N : N : N : N : N : N : N : N : N
9524 SCC : N : N : N : N : N : N : N : N : N : SCC
9990 N : N : N : N : N
9998 SCC : N : N : SCC : N : SCC : N : N : N : N
10040 N : N : N : N : N : N
10158 N : N : N : N
9626 N : N : N : N : N : N
10140 N : N : N : N : N : N : N : N : N : N
10135 N : N : N : N : N : SCC : N : SCC : N : N
9540 N : N : N : N : N : N : N

**Pre-clinical MCI**

10025 N : N : N : N : SCC : N : SCC : MCI_AD : MCI_AD : AD
10110 N : SCC : SCC : SCC : MCI_AD : DO : MCI_AD : SCC : AD : AD
9619 N : N : N : MCI_AD : MCI_AD
9970 N : MCI_O : SCC : MCI_O : MCI_AD
9815 N : N : N : MCI_AD : MCI_AD : MCI_AD ←
9629 N : N : N : N : N : N : SCC : SCC : MCI_AD
9557 N : MCI_O : N : SCC : MCI_AD
10147 N : N : N : N : SCC : MCI_O : MCI_AD
10348 N : N : SCC : N : SCC : SCC : MCI_AD : MCI_AD
10392 N : MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD : AD

Example:
Subject with 3 visits with N and 3 MCI_AD Dx

N: normal
SCC: subjective cognitive concern,
MCI_AD: amnestic MCI,
MCI_O: non-amnestic MCI
AD: Alzheimer dementia

**Colored Cognitive Chains**

# Visualization of longitudinal cognitive within each manually assigned category

**MCI Stable**

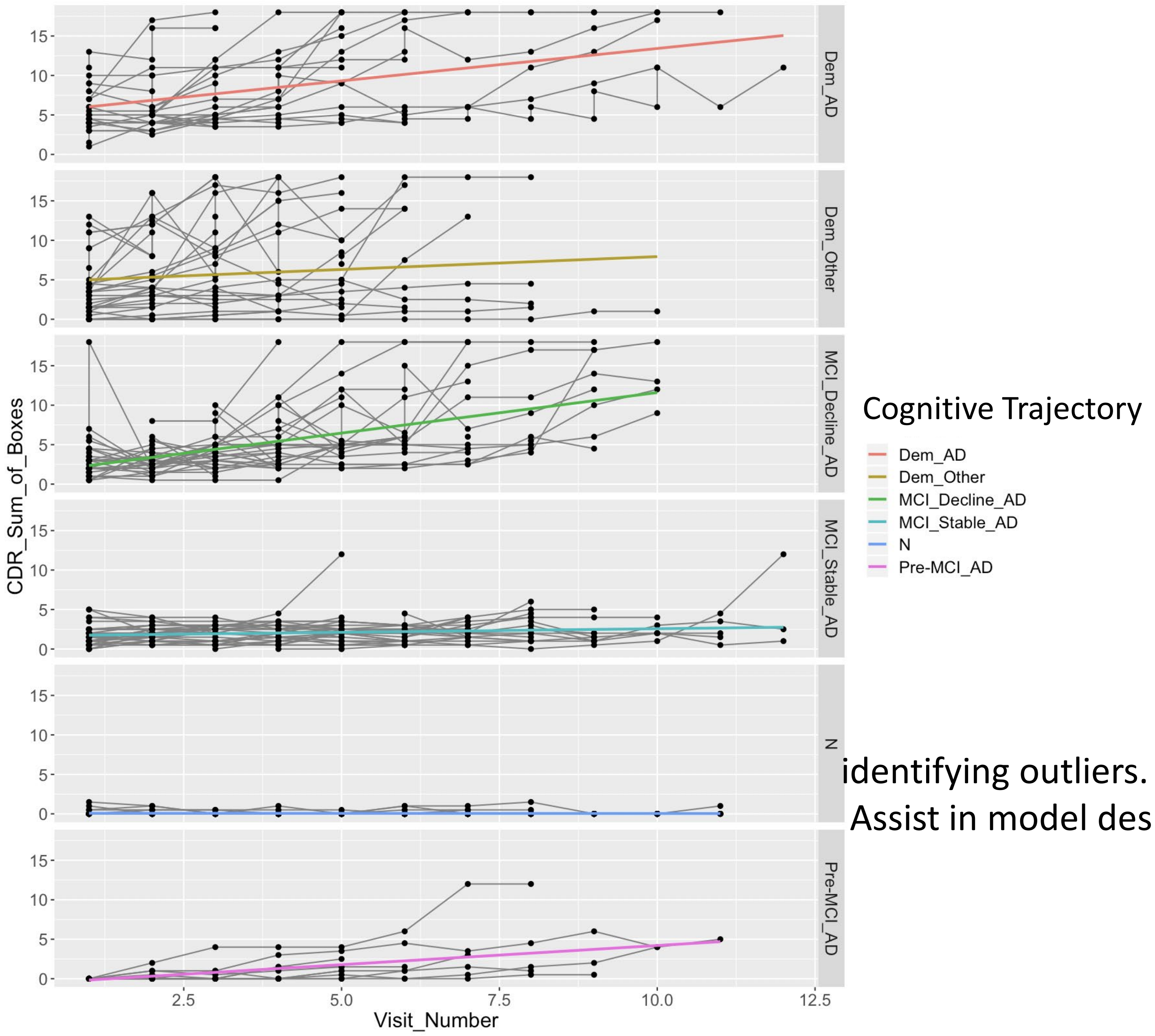
9679 AD : AD : MCI_AD : MCI_AD : MCI_AD : MCI_O : MCI_O : MCI_AD : SCC : MCI_AD : SCC
9421 MCI_O : MCI_AD : MCI_AD : MCI_AD : MCI_AD
9999 MCI_AD : MCI_AD : MCI_AD : SCC : SCC : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
9502 SCC : MCI_O : MCI_AD : MCI_O : MCI_O : SCC
10339 MCI_AD : SCC : SCC : SCC : SCC : SCC : MCI_AD : SCC
9910 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD
9344 MCI_O : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_O : MCI_AD : MCI_AD : MCI_AD
9503 SCC : MCI_AD : MCI_AD : MCI_AD : MCI_O : SCC : SCC : MCI_AD : MCI_AD
9974 SCC : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
9635 MCI_AD : MCI_AD : MCI_AD : DO : MCI_AD : MCI_AD : MCI_O
9843 N : MCI_O : MCI_AD : SCC : MCI_O : MCI_AD : MCI_O : MCI_AD : MCI_AD
9354 MCI_AD : MCI_AD : MCI_AD : MCI_AD : SCC
10499 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
10512 MCI_AD : MCI_AD : MCI_O : MCI_AD : MCI_AD : MCI_AD : MCI_AD
16986 MCI_AD : MCI_AD : MCI_AD
9589 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
9578 MCI_O : MCI_AD : MCI_AD : MCI_O : MCI_O : MCI_O : MCI_O : MCI_AD
9326 MCI_AD : MCI_AD : MCI_AD
10122 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
10005 N : MCI_AD : SCC : SCC : SCC : MCI_AD : MCI_AD : MCI_AD
9396 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD
9534 AD : MCI_AD : MCI_AD : SCC : MCI_AD : MCI_AD : SCC : SCC : SCC : MCI_AD : MCI_AD
9560 AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD : AD
9641 MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD
10401 MCI_AD : MCI_AD : MCI_AD : SCC : MCI_AD : MCI_AD
10546 MCI_AD : MCI_AD : MCI_AD : MCI_AD
10099 MCI_O : MCI_AD : MCI_AD : AD : DO
9530 MCI_O : MCI_O : MCI_O : MCI_O : MCI_O : MCI_O : MCI_AD : MCI_AD
9682 SCC : MCI_O : SCC : MCI_O : SCC : SCC : MCI_AD : N : SCC : MCI_AD : AD : AD
10485 MCI_O : MCI_AD : MCI_AD : MCI_AD
10462 MCI_O : SCC : MCI_AD : MCI_O : MCI_O : MCI_AD

Visualizing entire longitudinal history can assist in labeling

**MCI Decliners**

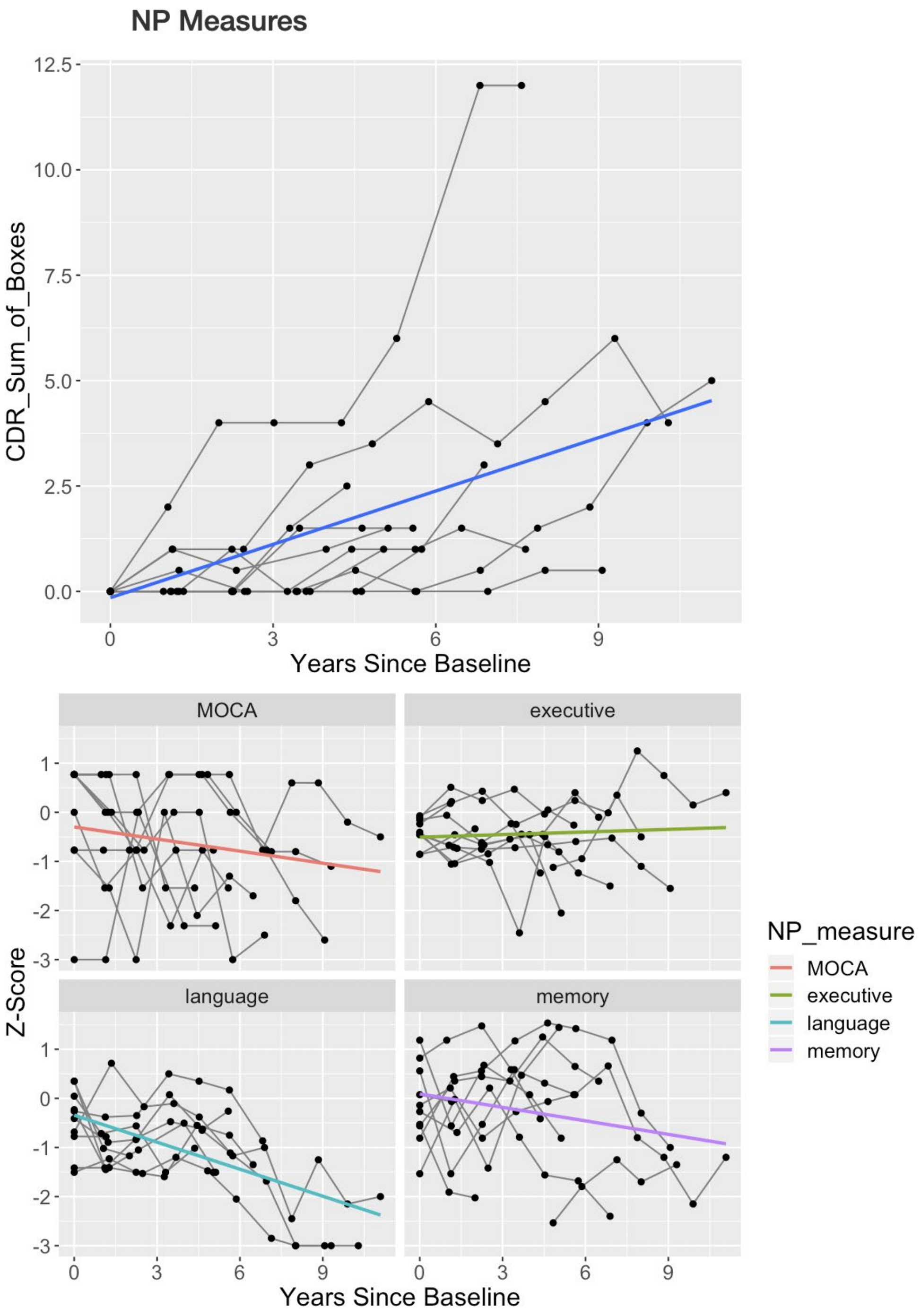
10523 MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD
9360 MCI_AD : MCI_AD : AD : AD : AD : AD : AD : AD : AD
10246 MCI_AD : MCI_AD : AD : AD : AD : AD
9432 MCI_AD : MCI_AD : MCI_AD : AD : AD : AD : AD : AD : AD : AD
9481 MCI_AD : MCI_AD : AD : MCI_AD : AD : AD : AD : AD : AD : AD
9482 SCC : MCI_AD : AD
9451 MCI_O : MCI_AD : MCI_AD : MCI_AD : AD : AD
9677 SCC : DO : MCI_AD : MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD
10196 MCI_AD : AD
16384 MCI_AD : AD : AD : AD
9404 MCI_AD : MCI_AD : MCI_AD : AD
10021 MCI_AD : MCI_O : AD
9468 SCC : MCI_AD : AD : AD : AD : AD : AD : AD : AD : AD
9571 MCI_AD : MCI_AD : AD : AD : AD : AD
9400 AD : MCI_O : MCI_AD : AD : MCI_AD : MCI_AD : AD
9522 MCI_AD : MCI_AD : MCI_O : MCI_O : MCI_O : MCI_O : AD : DO : DO
9374 MCI_AD : MCI_AD : MCI_AD : MCI_AD : AD : AD : AD : AD : AD
9653 SCC : MCI_AD : MCI_AD : MCI_O : MCI_O : MCI_O : MCI_AD : AD : AD : AD
9916 MCI_AD : AD
10179 MCI_AD : MCI_AD : AD : AD : AD
10534 MCI_AD : AD : AD : AD : AD : AD
16659 MCI_AD : MCI_AD : AD : MCI_AD : AD
16542 MCI_O : DO : DO : DO
9905 MCI_O : MCI_AD : AD : AD : AD : AD : AD : AD
10397 MCI_O : MCI_AD : MCI_AD : MCI_AD : AD : AD : AD
9906 MCI_O : MCI_O : DO : DO : DO : DO : DO
9913 MCI_O : MCI_O : AD : AD : AD : AD : AD
9681 MCI_O : MCI_O : MCI_O : SCC : SCC : SCC : SCC : AD

# Visualization of CDR Sum of Boxes for pre-defined categories




identifying outliers.
Assist in model design

# Further explore pre-MCI AD




**Neuropsych Test**

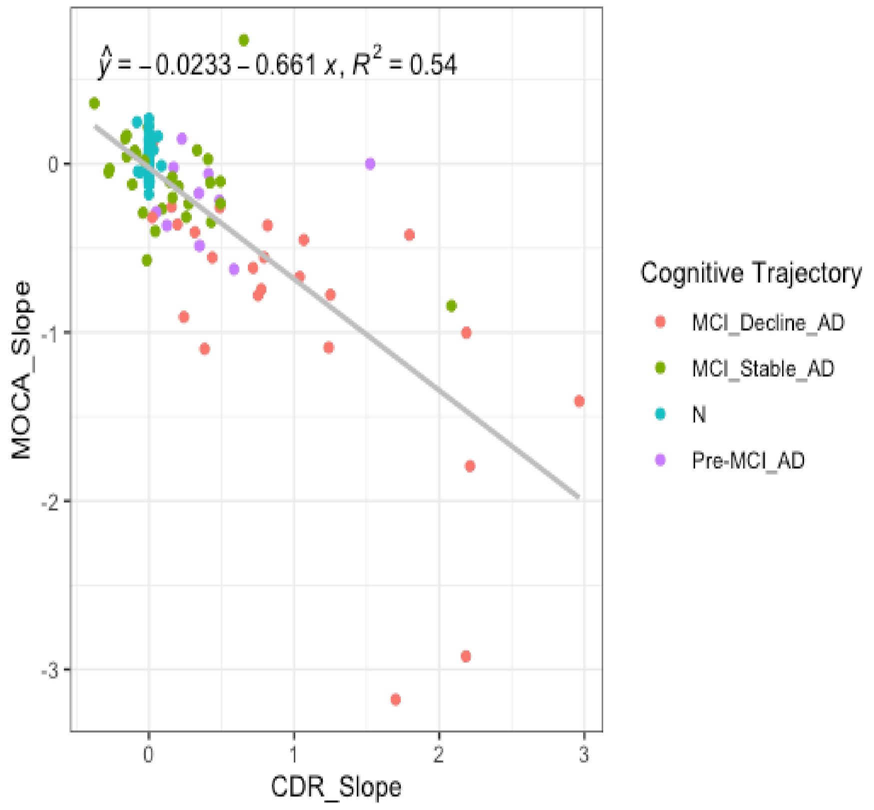
Z-scores using NACC norms with age, sex, education adjustments were computed.

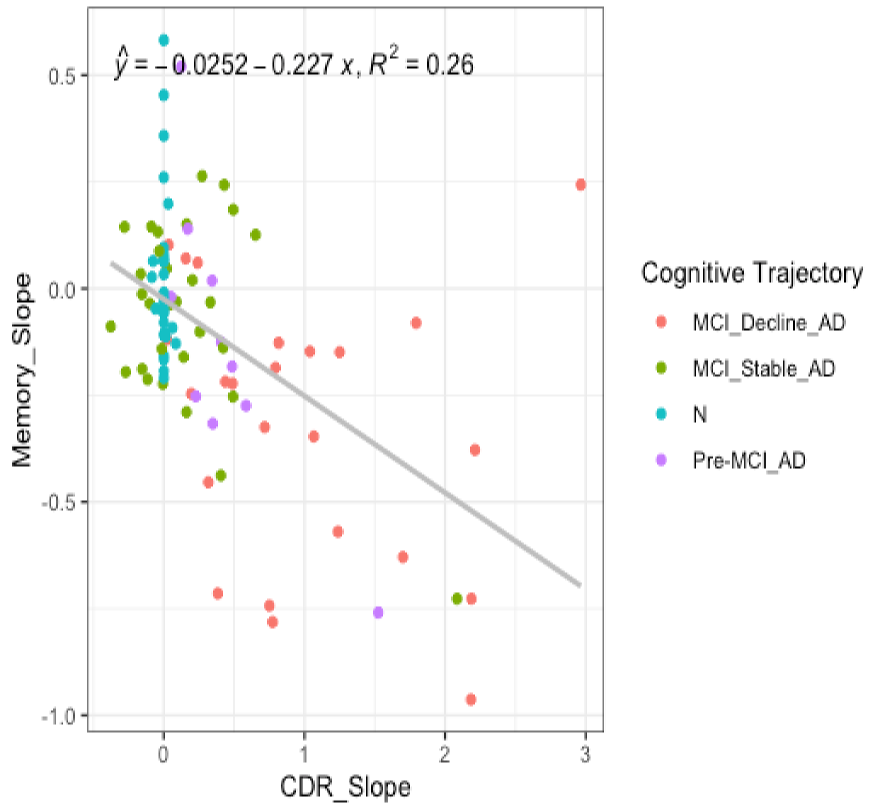
**Executive**: Trails B and digits backward

**Language**: Animals & Vegetables

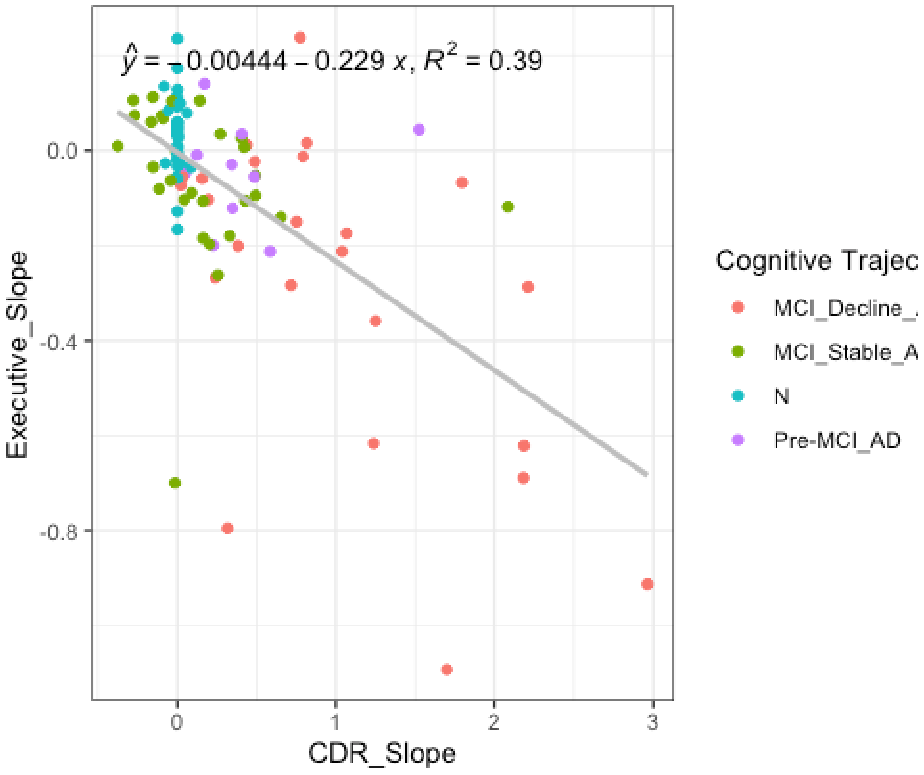
**Memory**: Craft immediate and delayed paraphrase

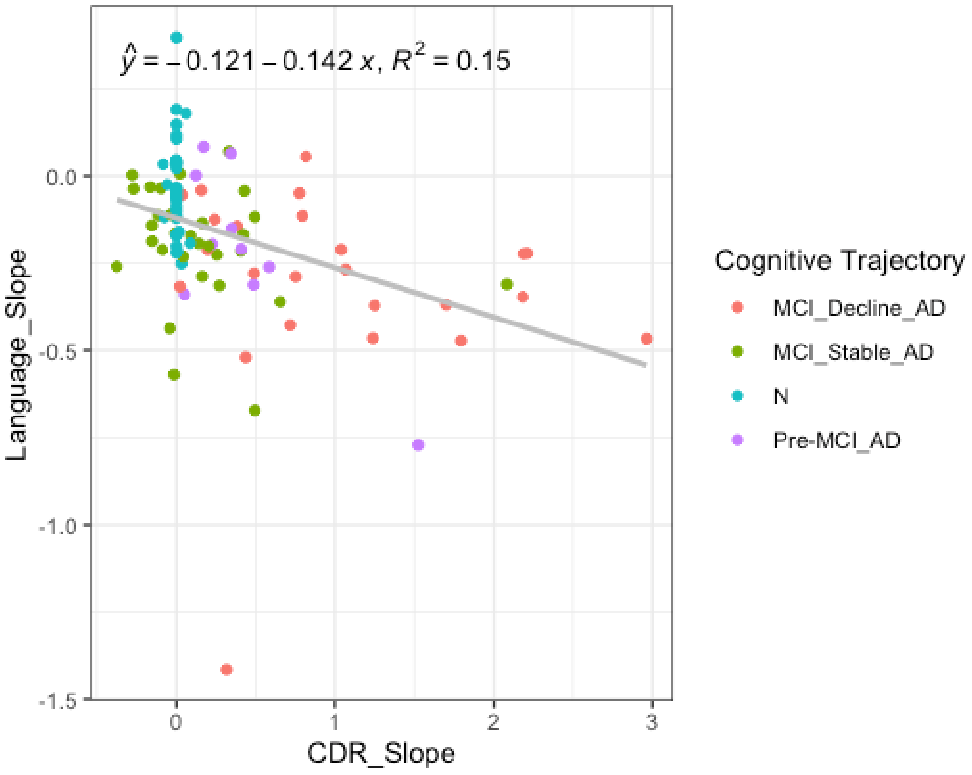
# How do slopes of CDR SOB trajectory relate to slopes of Neuro-Psych?



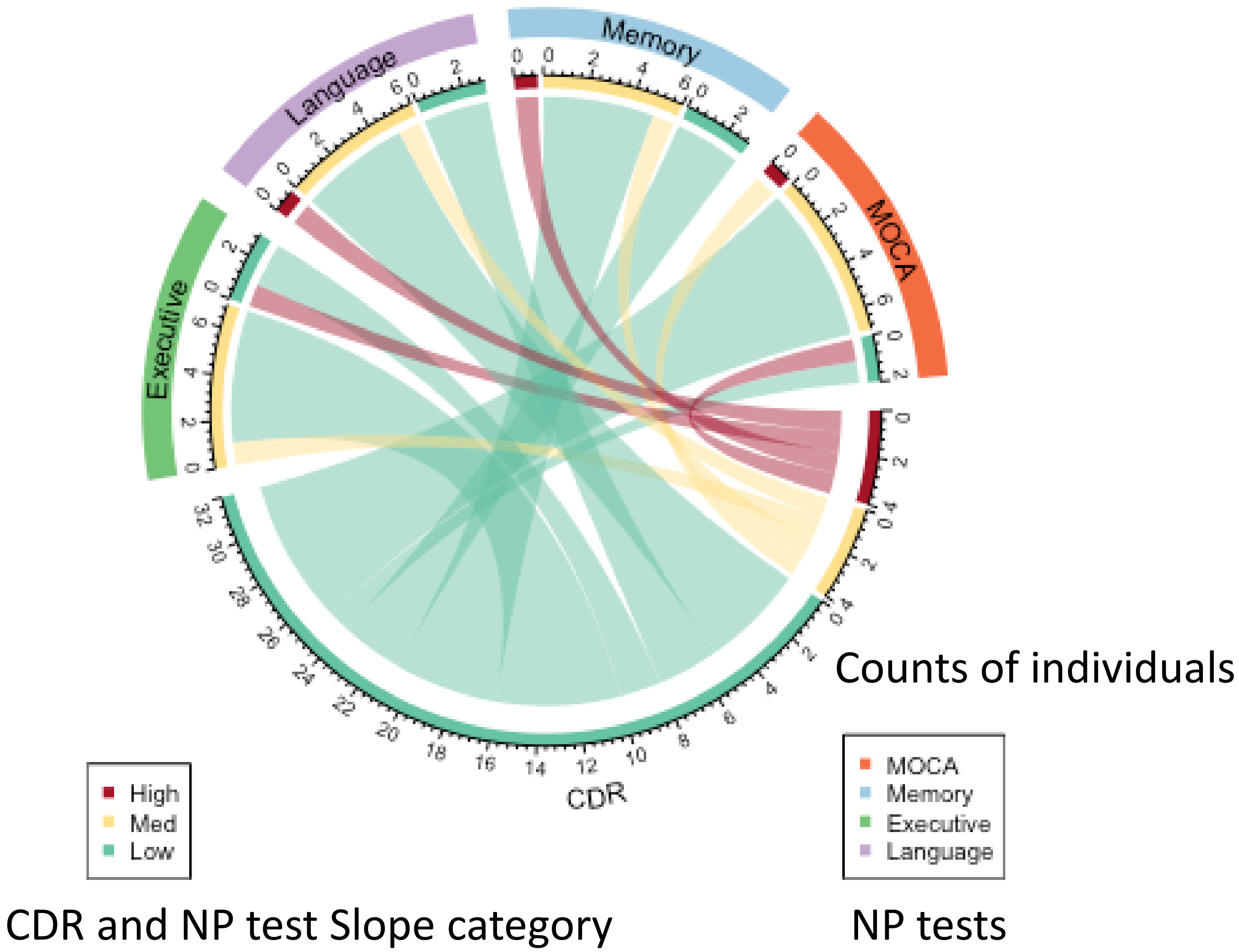



# How do slopes of CDR SOB trajectory relate to slopes of Neuro-Psych?



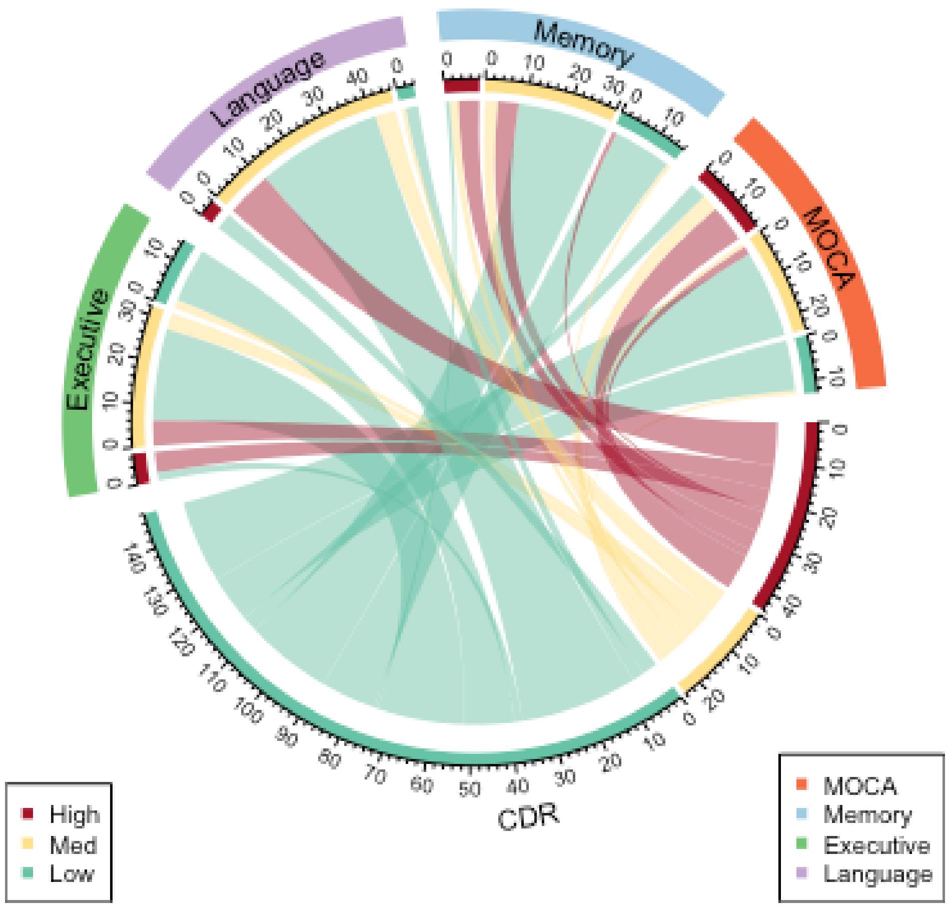



# How do slopes of CDR SOB trajectory relate to slopes of Neuro-Psych in *pre-MCI*?



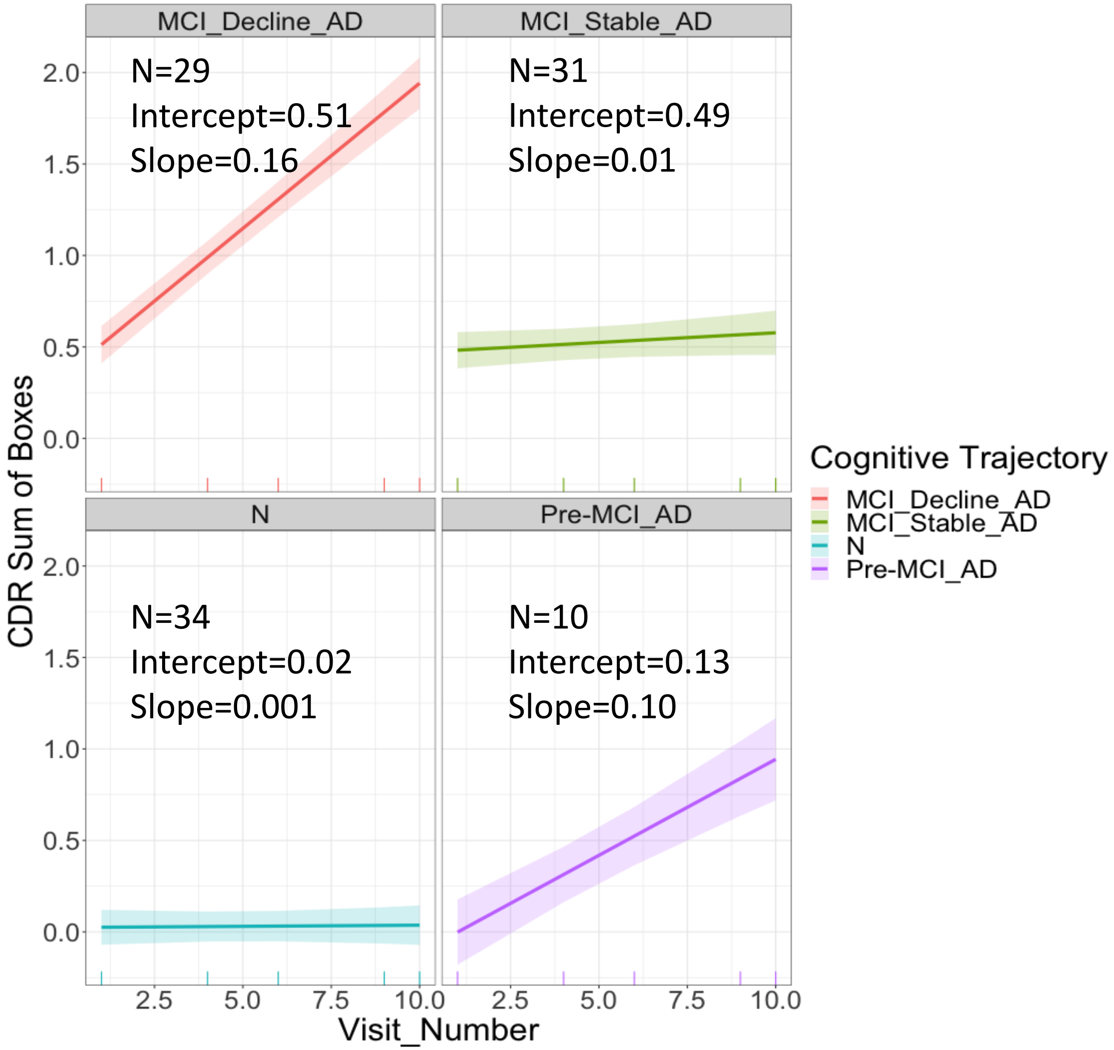

- We investigated the relationship of CDR Slopes with Neuro-Psych test slopes for *pre-MCI*
- Slopes were divided into high, med or low values
- Concordance of CDR slopes with *language* and *memory* slopes: Subjects with high CDR slopes also have high memory, high language slopes.
- CDR slopes not concordant with *MOCA* or *executive* slopes

# How do slopes of CDR SOB trajectory relate to slopes of Neuro-Psych in *MCI*?




- We investigated the relationship of CDR Slopes with Neuro-Psych test slopes for all *MCI*
- Slopes were divided into high, med or low values
- Higher concordance of CDR with MOCA, memory, executive or language slopes

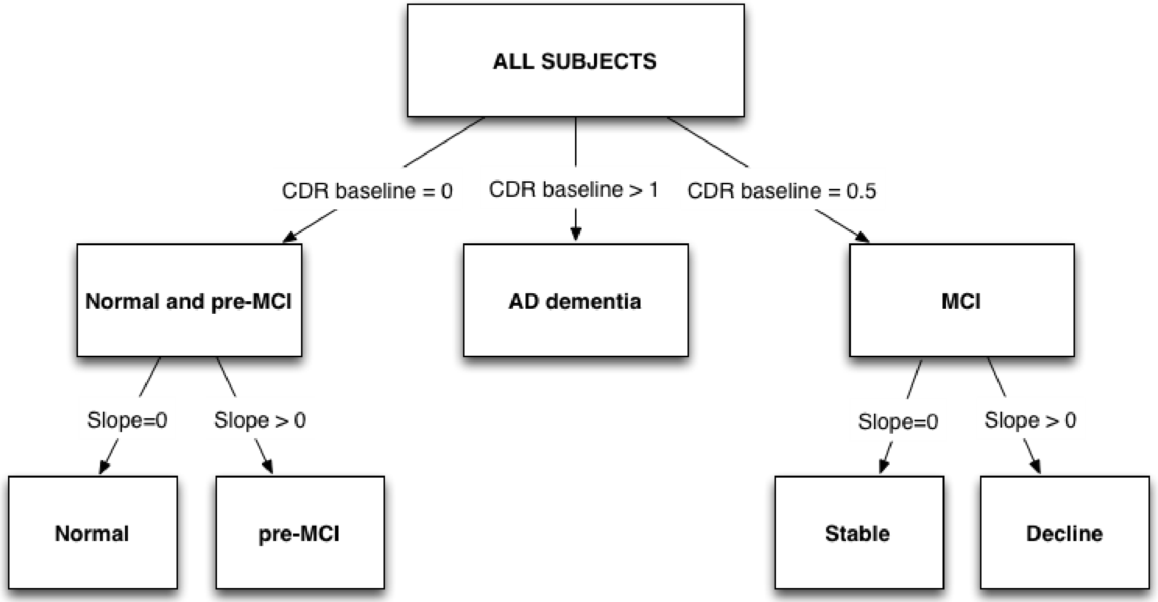
# Do cognitive-categories have distinct patterns of CDR Sum of Boxes trajectories?




- Mixed-effect model with subject specific intercepts and slopes were fitted to the data
- Different cognitive trajectory have distinct slopes and intercepts

# Can we use machine learning to predict categories?

Manually generated decision tree




- Decision Tree
  - Technique in supervised learning algorithm for classification
  - We can generate a tree for small number of variables (see tree on left) but need machine-learning for complex data
  - Several tree-based methods: bagging, boosting and random forest
  - We used boosting method as it is resistant to overfitting
  - Boosting builds ensemble of trees from bootstrapped datasets and improves model by fitting residuals

# Performance of machine learning tool

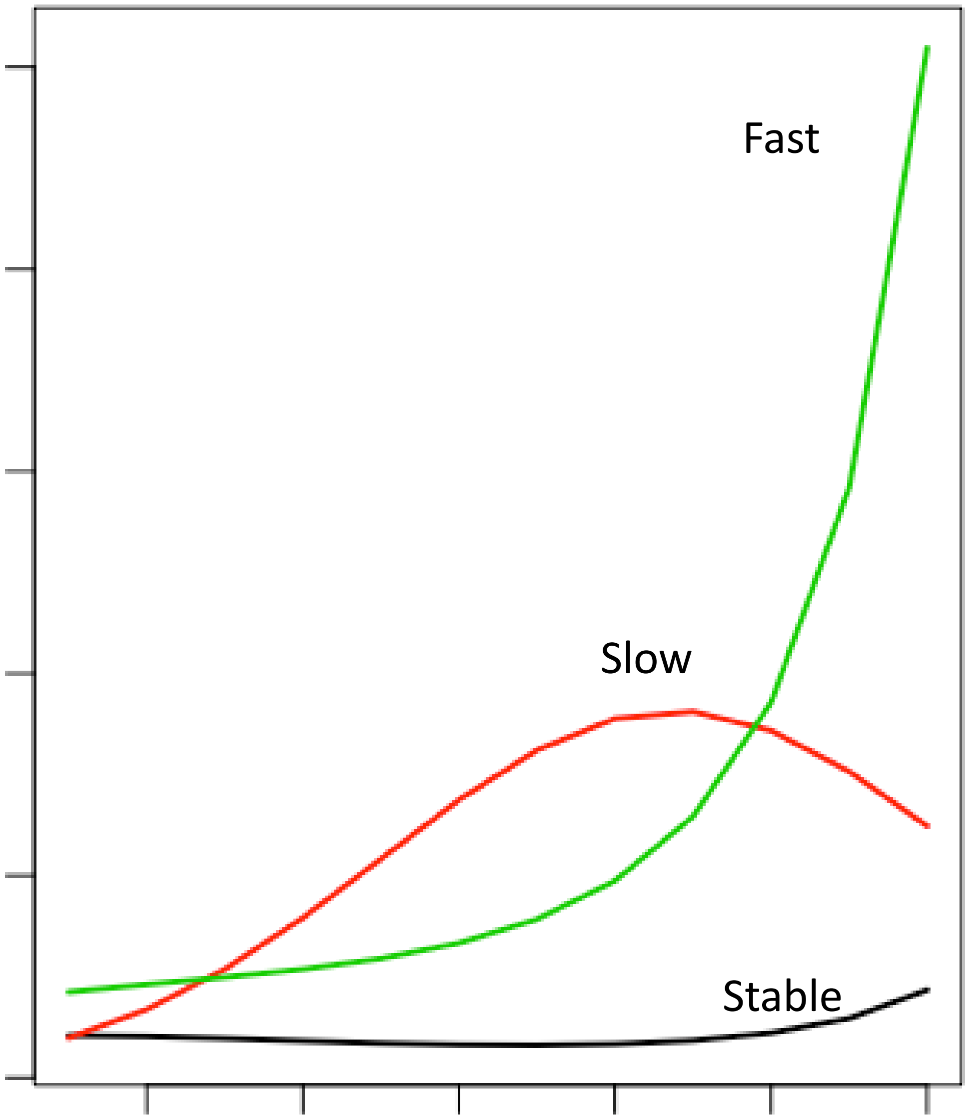
- We used CDR slope and intercept as input variables
- Used 1/3 dataset as hold-out for testing
- Achieved > 80% accuracy against manual labels
- Adding additional neuro-psych test  variables did not improve model performance

- Classification Error examples:
  - Pre-MCI AD -> N
  - Pre-MCI AD -> MCI_Decline_AD
  - MCI_Decline_AD -> Dem_AD
    - Example

9400 **AD** : **MCI_O** : **MCI_AD** : **AD** : **MCI_AD** : **MCI_AD** : **AD**

# Study heterogeneity of amnestic MCI with data-driven methods

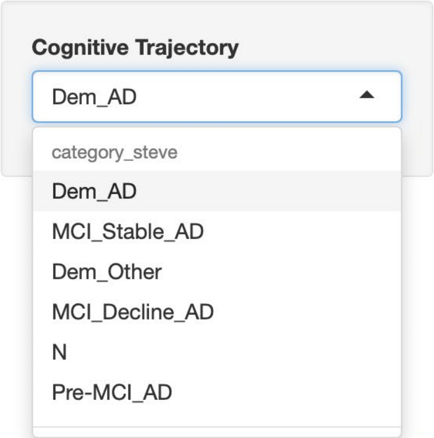
- Next we explored whether data-driven-methods can identify clusters within the *amnestic MCI group*
- Used a finite mixture models for clustering trajectories
- CV errors are used to set # of latent classes in group-based trajectory analysis



# Classify cognitive trajectory of all MADRC subjects

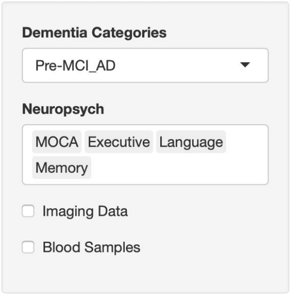
| Category | N |
|---|---|
| N | 341 |
| Pre-MCI_AD | 114 |
| MCI_Stable_AD | 242 |
| MCI_Decline_AD | 200 |
| Dem_AD | 160 |
| **Total** | **1057** |

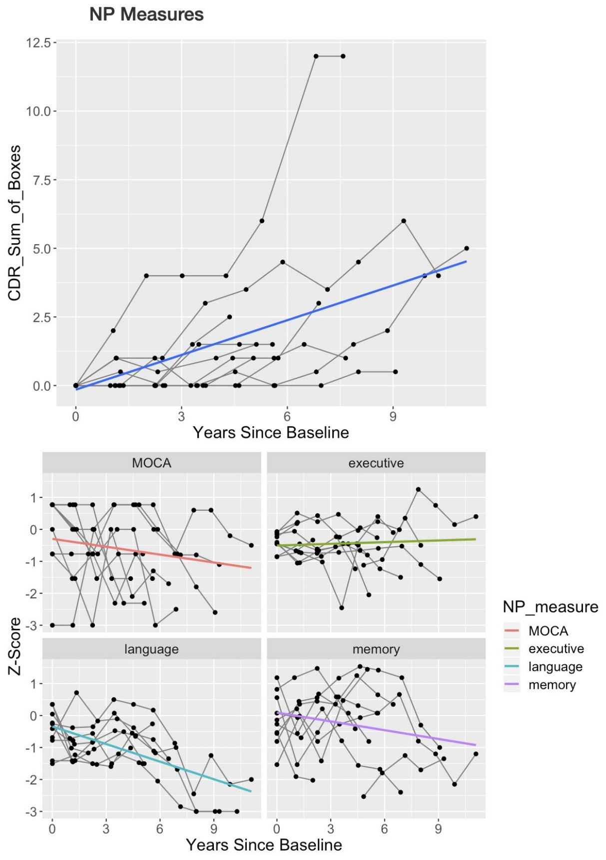
# Interactive web-based tool




**Cognitive Dx Chains**

9413 MCI_AD : MCI_AD : DO : AD : AD : AD
10249 AD : AD : AD : AD
10543 AD : AD : AD : AD : AD : AD
9420 MCI_AD : MCI_AD : AD : AD : AD : AD : AD : AD : AD : AD
9485 AD : AD : AD
9370 MCI_AD : AD : AD : AD : AD : AD : AD : AD : AD
9412 AD : AD : AD : AD : AD
9442 AD : AD : AD : AD : AD : AD
9378 AD : AD : AD : AD : AD : AD : AD : AD : AD : AD : AD
9466 AD : AD : AD : AD : AD
9665 AD : AD : AD : AD : AD : AD
9407 AD : AD : AD : AD : AD : AD : AD : AD : AD : AD
10243 AD : AD : AD : AD : AD : AD : AD
10464 AD : AD : AD
10516 AD : AD : AD
10524 AD : AD : AD
10550 AD : AD : AD : AD : AD : AD
10562 AD : AD : AD : AD
10602 AD : AD : AD
9318 DO : DO : AD



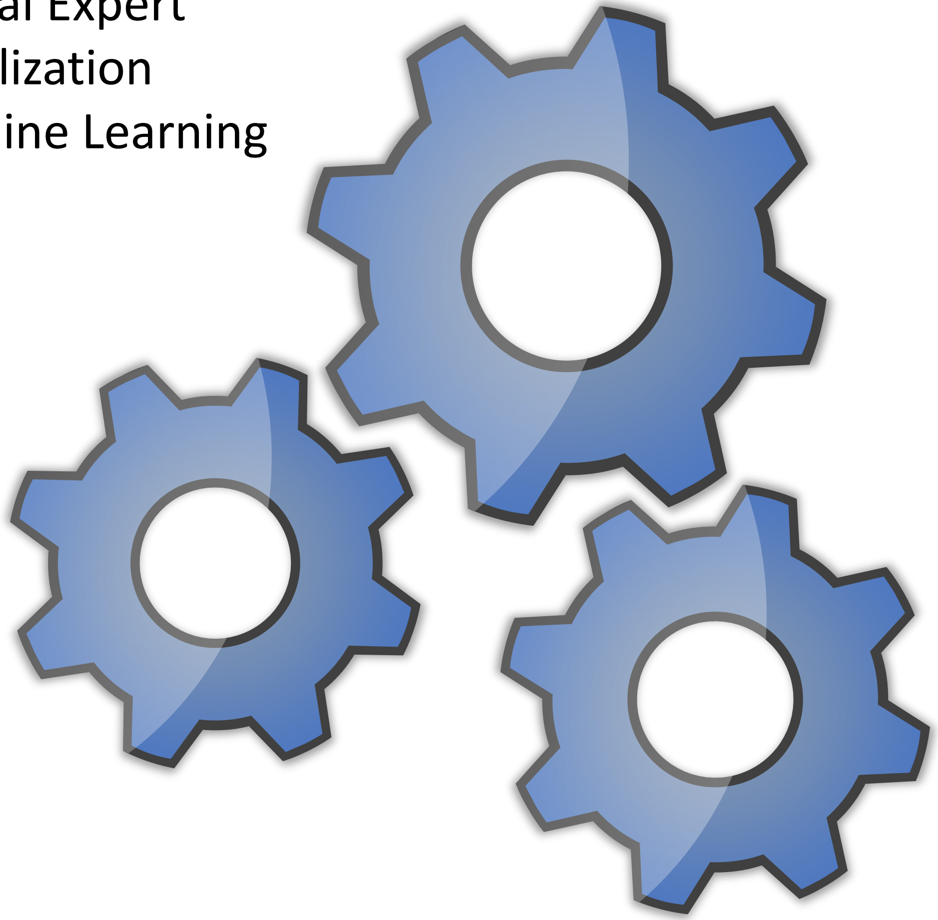



# Conclusions & Future Directions

- Developed machine-learning model to categorize cognitive trajectory
- Built visualization tools to illustrate trajectories
- In future too will be deployed to center researchers
  - Allow researchers to query databases for cognitive trajectories of interest
  - Add demographics, APOE genotype, longitudinal blood-draw and imaging information to visualization

# Multi-disciplinary Approach

Study Researcher
Clinical Expert
Visualization
Machine Learning



Challenge: Assess cognitive Dx trajectory using tabular format in spreadsheets

Visualization provides a manageable representation of the same data leading to more efficient labeling

Computational model further increases scalability

Model is further refined by input from clinical experts

Multidisciplinary framework can be applied to other problems

# Acknowledgements


- Bradley T Hyman, MD, PhD
- Steven E Arnold, MD, PhD
- Derek Oakley, MD, PhD
- Colin Magdamo
- Emily Merrill