



# Contents

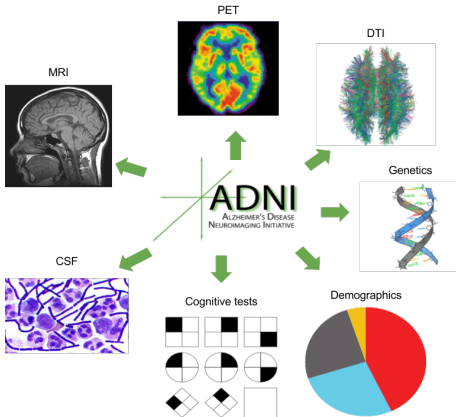
1. TADPOLE Overview
2. Modalities Available
3. Data Access

# TADPOLE Overview

The Alzheimer's Disease Prediction Of Longitudinal Evolution (TADPOLE) is a dataset created and formatted by EuroPOND consortium with funding from the European Union's Horizon 2020 research and innovation programme for use in a scientific challenge for senior high school students to develop machine learning algorithms to predict future evolution of individuals at risk of Alzheimer's Disease.



# TADPOLE Overview



TADPOLE is a longitudinal data derived from the ADNI dataset with

- preselected features
- fairly preprocessed
- well documented data descriptions at the tadpole website

Diagram showing TADPOLE biomarkers. Source of individual images: Wikimedia Commons

# Modalities Overview

- Diagnosis
- Demographics
- Cognitive Tests
- Genetic information
- MRI ROI
- PET ROI

# Diagnosis

Longitudinal

1730/1737 Patients (99.6%)

The entries specify both the current diagnosis and the baseline so, for example "MCI to Dementia" means that the current diagnosis is Dementia, while the diagnosis at the previous visit was MCI. EMCI,LMCI,SMC are only available at the baseline.

Code	Meaning
CN	Normal Aging/Cognitively Normal
EMCI/MCI	Early Mild Cognitive Impairment
LMCI	Late Mild Cognitive Impairment
SMC	Significant Memory Concern
AD	Alzheimer's disease

# Demographics and Genetic

## Demographics Static

1737/1737 Patients (100%)

- Age
- Sex
- Education
- Ethnicity
- Race
- Marital Status

## Genetic (ApoE4) Static

1725/1737 Patients (99.3%)

- The alipoprotein E4 variant (APOE E4) is a gene that is the largest known risk factor for AD. Subjects with APOE E4 have a risk 10 to 30 times higher of developing AD compared to non-carriers (i.e. subjects without the gene).

Two of the biggest two known risk factors associated with Alzheimer's Disease are Age and APOE4

# Cognitive Tests

Longitudinal

1737/1737 Patients (100%)

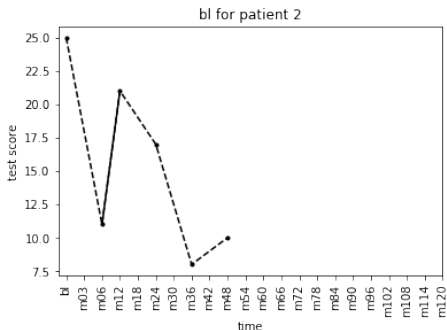
Measure cognitive decline in a direct and quantifiable manner.

- CDR Sum of Boxes 1737/1737 Patients (100%)
  - ▶ Clinical Dementia Rating Scale Sum of Boxes
- ADAS11, ADAS13 1734/1737 Patients (99.8%)
  - ▶ Alzheimer's Disease Assessment Scale
- MMSE 1737/1737 Patients (100%)
  - ▶ Mini-Mental State Examination
- RAVLT 1735/1737 Patients (99.9%)
  - ▶ Rey Auditory Verbal Learning Test

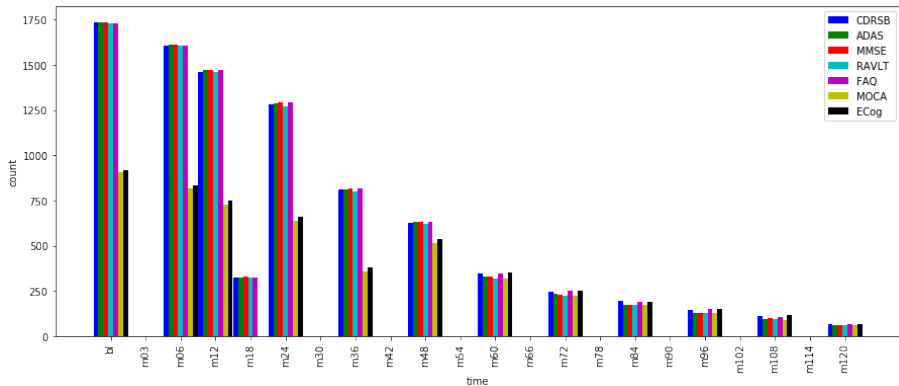


# Cognitive Tests

- MOCA 1202/1737 Patients (69.2%)
  - ▶ Montreal Cognitive Assessment
- FAQ 1732/1737 Patients (100%)
  - ▶ Functional Activities Questionnaire
- Ecog 1219/1737 Patients (70.2%)
  - ▶ Everyday Cognition



# Cognitive Tests

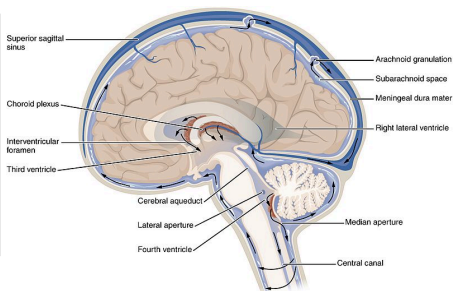
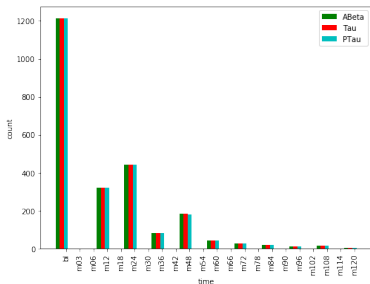


Longitudinal

1255/1737 (72.3%)

Abnormal levels of certain proteins in the cerebrospinal fluid (CSF) are some of the earliest signs of Alzheimer's disease and can indicate abnormalities many years before symptom onset.

This dataset contains amyloid-beta level, tau level, and phosphorylated tau level.



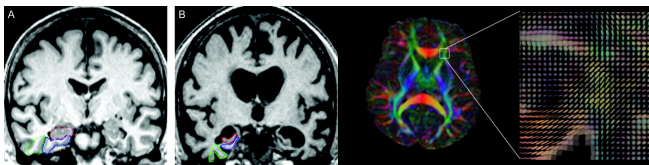
## ■ MRI

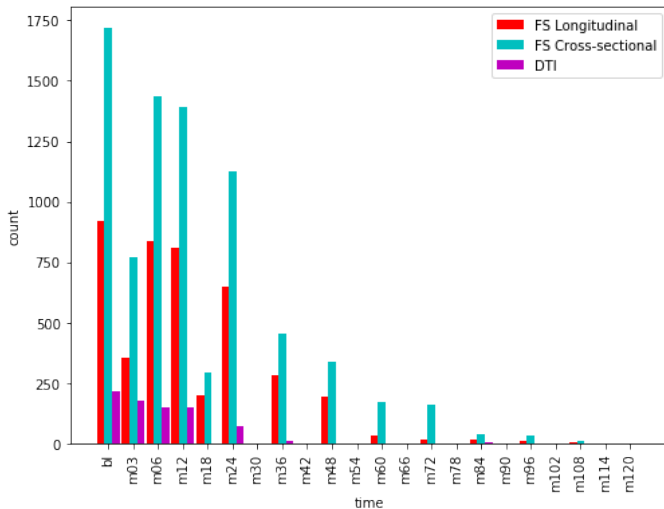
TADPOLE datasets include three main types of structural MRI markers of atrophy: 1. ROI volumes 2. ROI cortical thicknesses 3. ROI surface areas

- ▶ FreeSurfer (cross-sectional) 1735/1737 Patients (99.9%)
- ▶ FreeSurfer (longitudinal) 1105/1737 Patients (63.6%)

## ■ Diffusion Tensor Imaging (DTI) 249/1737 Patients (14.3%)

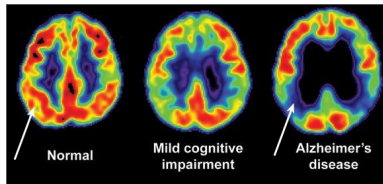
DTI can measure the degeneration of white matter (connections between neurons) in the brain. This is done by analyzing the diffusion of water molecules along the neuron fibre connections.

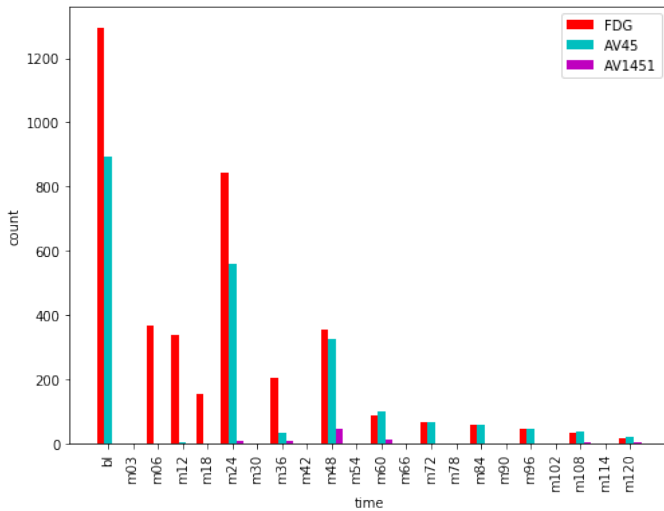




# PET

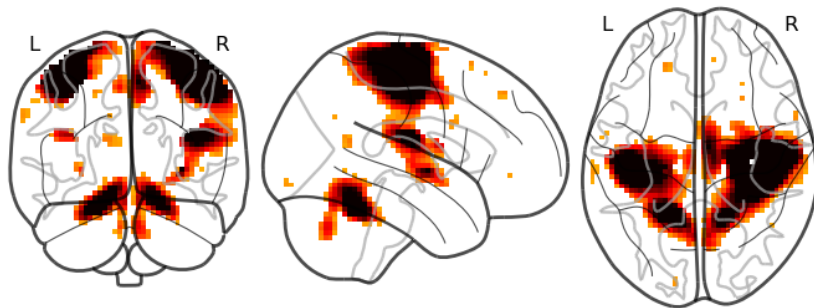
- FDG PET ROI Averages 1401/1737 Patients (80.7%)  
measure cell metabolism, where cells affected by AD show reduced metabolism
- AV45 PET ROI averages 1094/1737 Patients (63.0%)  
measures amyloid-beta load in the brain, where amyloid-beta is a protein that mis-folds (i.e. its 3D structure is not properly constructed), which then leads to AD
- AV1451 PET ROI 85/1737 Patients (4.9%)  
measures tau load in the brain, where tau is another protein which, when abnormal, damages neurons and thus leads to AD





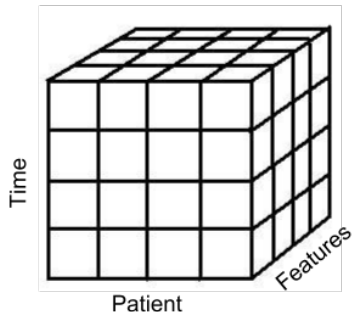
# Brain Visualization

All brain imaging data, MRI, DTI, and PET, has been sectioned into regions of interest ROIs, and such can be visualized with the python nilearn package.





# Data Access



- Linked on Slack
- Data and Name Files for all Modalities
  - ▶ `finished_(modaility)_dat.npy`
  - ▶ `(modaility)tags.npy`
- Dimensions: (patient, time, features)
- Size: (1737 x 22 x feature length)

- TADPOLE Website Data Information
- TADPOLE Challenge Paper

# Questions