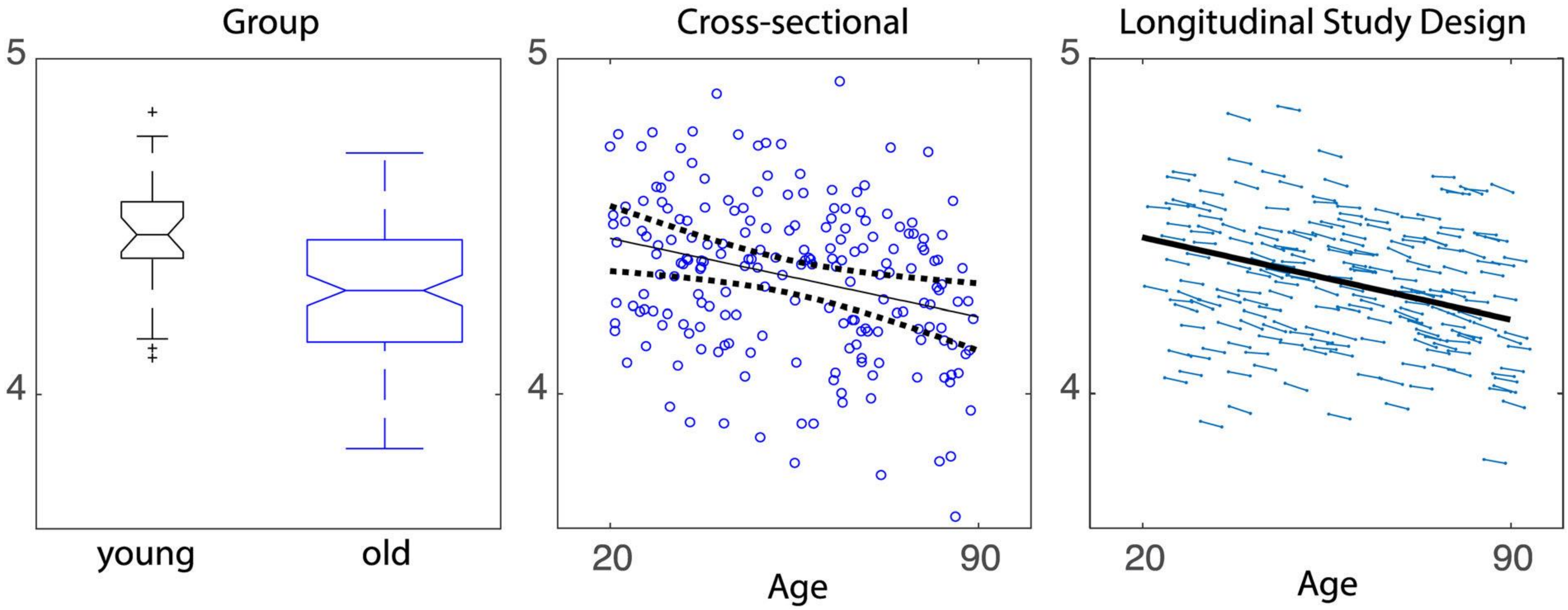
<u>Medical/Bio Research Topics Ⅱ: Week 10 (09.11.2023)</u>

Brain age estimation artificial intelligence models (2): model construction

(뇌나이 예측 인공지능 모델 개발 연습 (2): 예측 모델 구성)

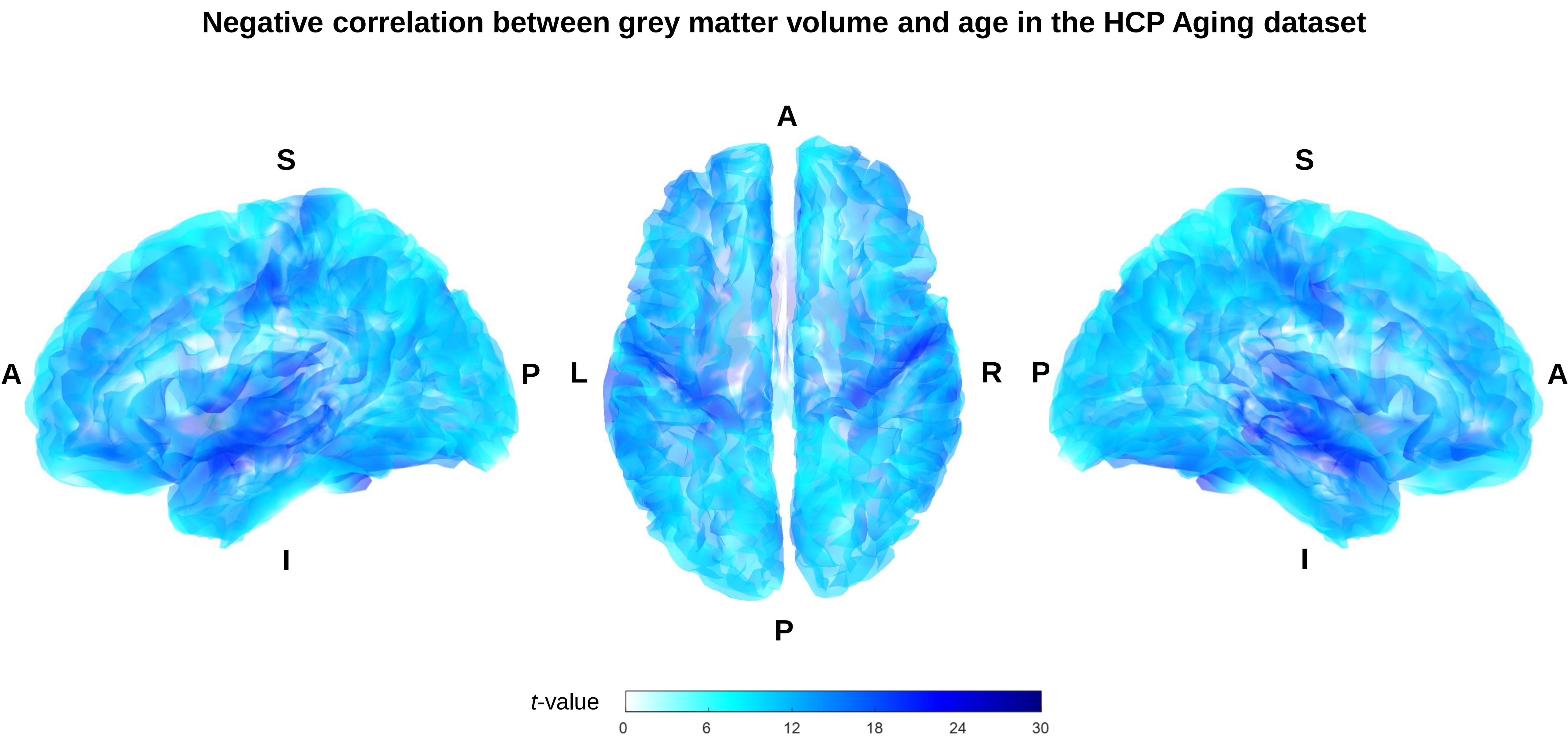
Brain Ageing on MRI

- Study designs for assessing age effects
  - Group
    - Age-related differences
  - Cross-sectional
    - Age-related differences
  - Longitudinal
    - Age-related changes
    - Enables to assess the rate of change with inter-individual variability removed

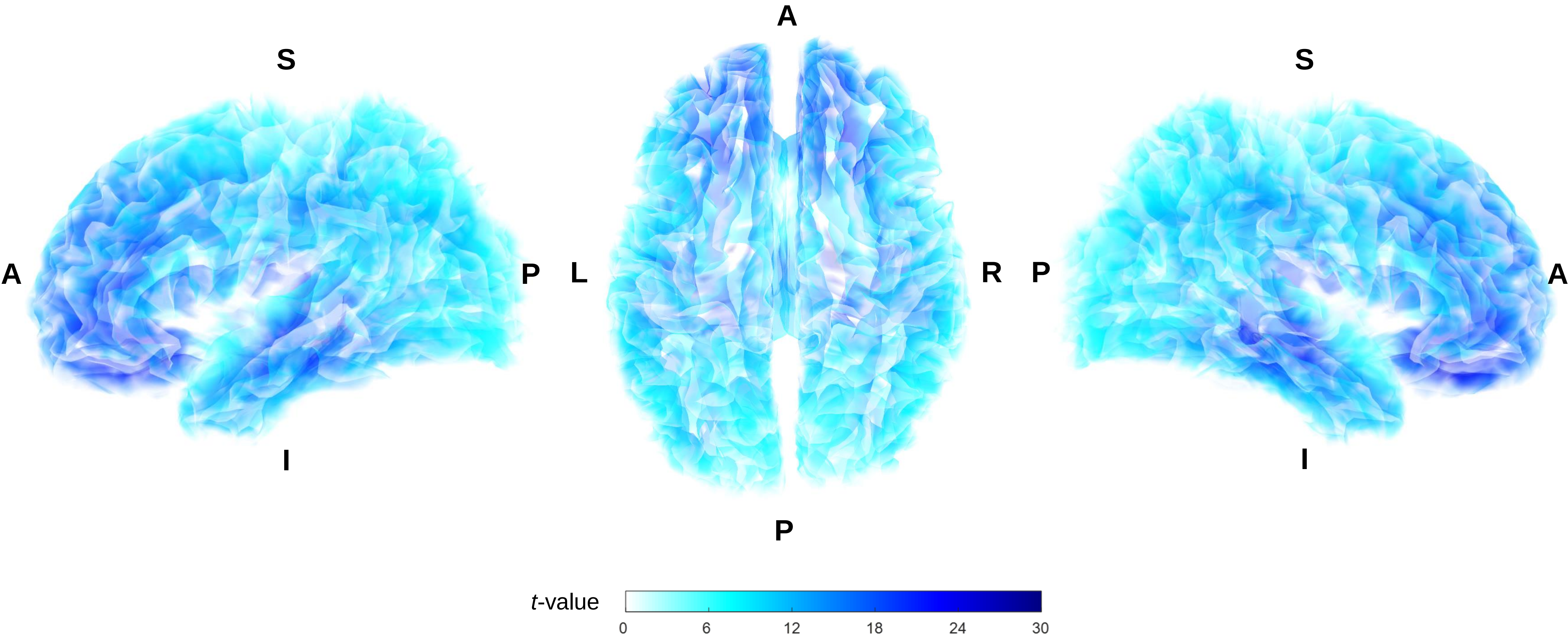


[MacDonald and Pike, 2021]

Study designs for assessing age effects


Negative correlation between grey matter volume and age in the HCP Aging dataset
S
A
P
I
A
L
R
P
S
P
A
I
t-value
0
6
12
18
24
30

Negative correlation between white matter volume and age in the HCP Aging dataset

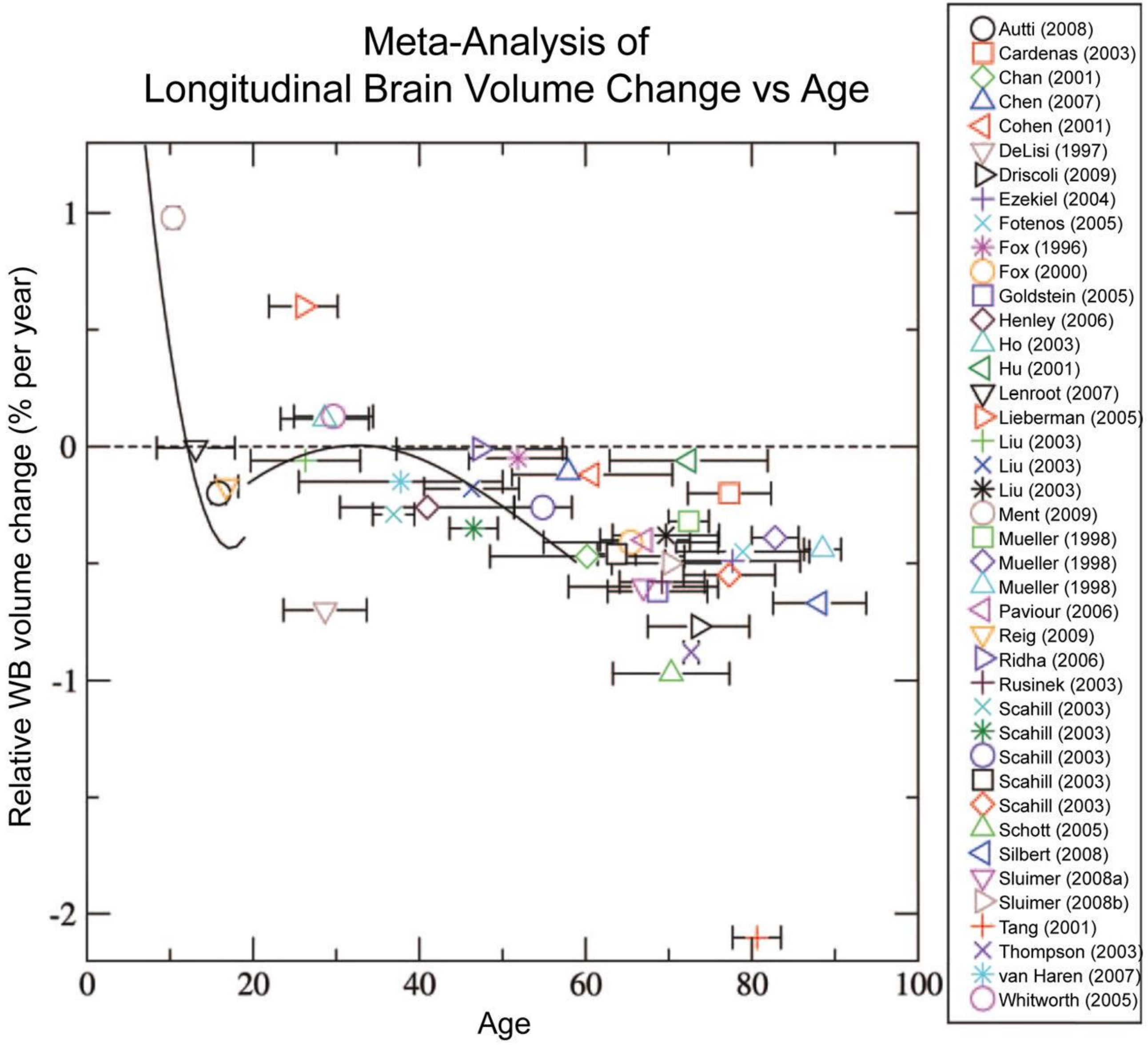


Region	Mean (95% Confidence Interval)	
	Cross-sectional Data	Longitudinal Data
Whole brain*	0.33 (0.25-0.41)	0.32 (0.10-0.54)
Temporal lobes*	0.35 (0.20-0.51)	0.68 (0.42-0.93)
Hippocampi*	0.35 (0.13-0.57)	0.82 (0.53-1.11)
Lateral ventricles, mm^3/y	521 (323-719)	650 (333-968)

[Schhill et al., 2003]

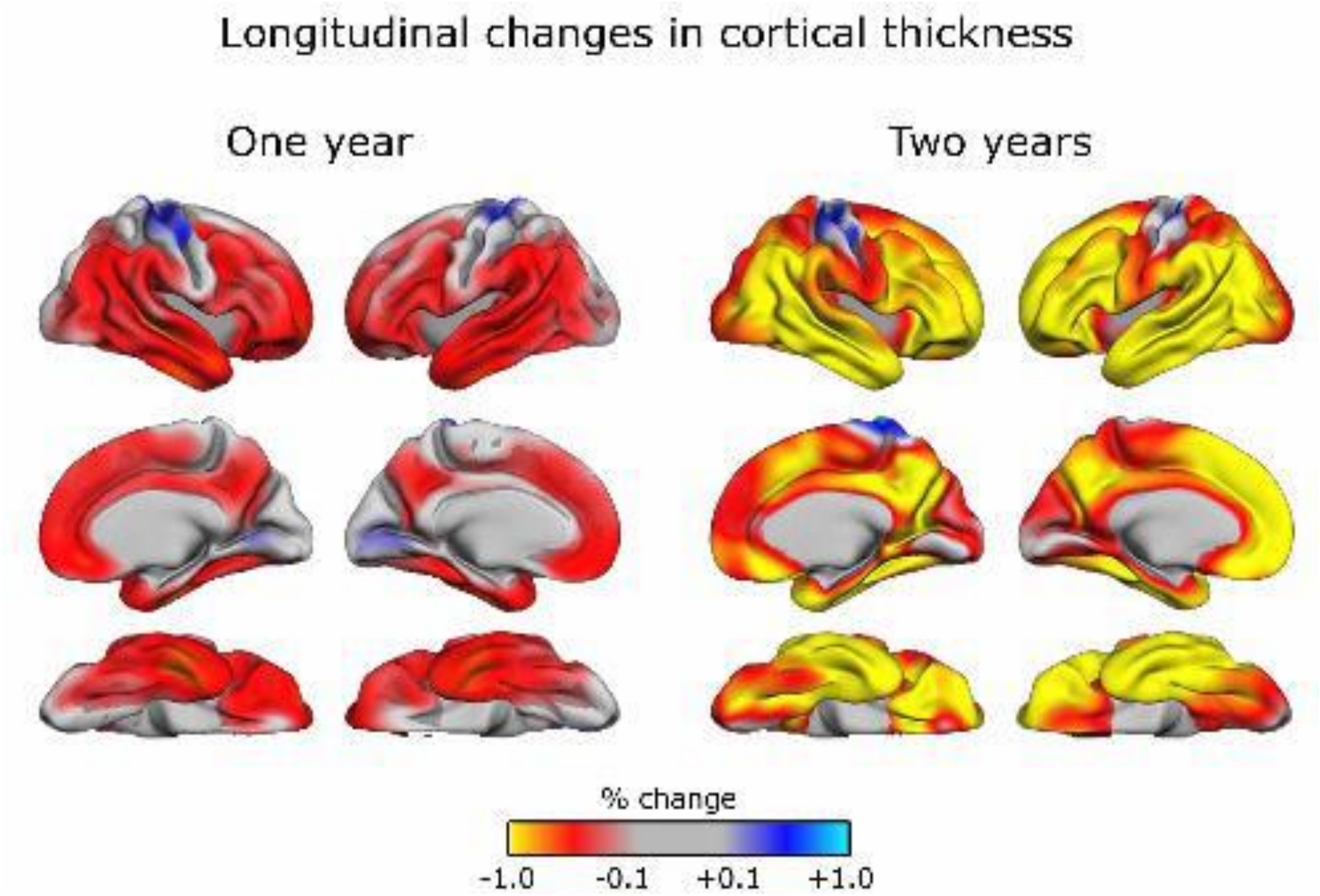
Annual rates of brain volume changes based on cross-sectional and longitudinal data

- Morphological changes
  - Shrinkage of brain volume
    - Continuous decline throughout the lifespan
      - With annual reductions of between 0.5% and 1.0% in most brain areas [Fjell and Walhovd, 2010]
    - For both grey matter and white matter
      - Slower rate of shrinkage for white matter [Ge et al., 2002]
    - Highly heterogeneous in the pattern of changes



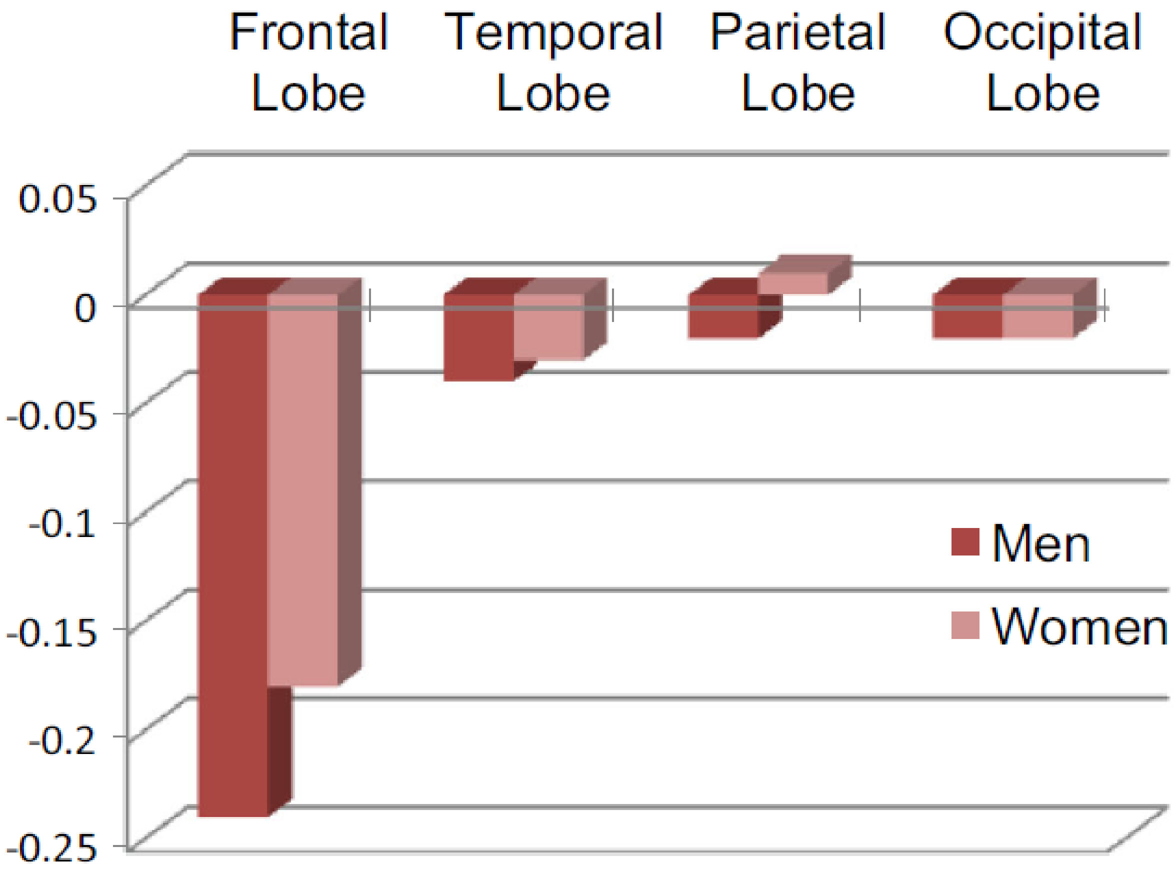
[Hedman et al., 2012]

Brain volume changes with age



[Fjell and Walhovd, 2010]

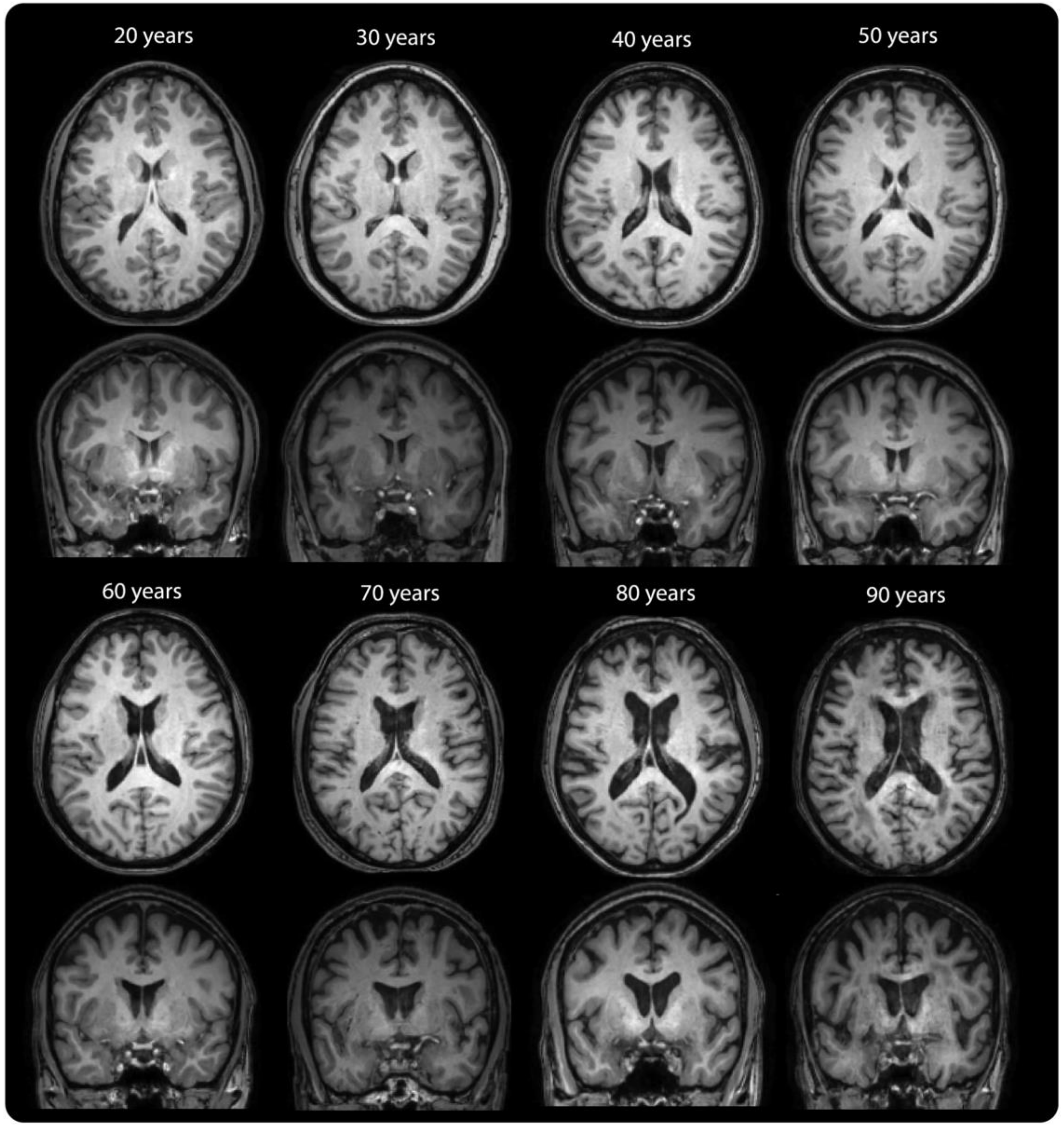
Percentage changes in cortical thickness over one and two years



[DeCarli et al., 2005]

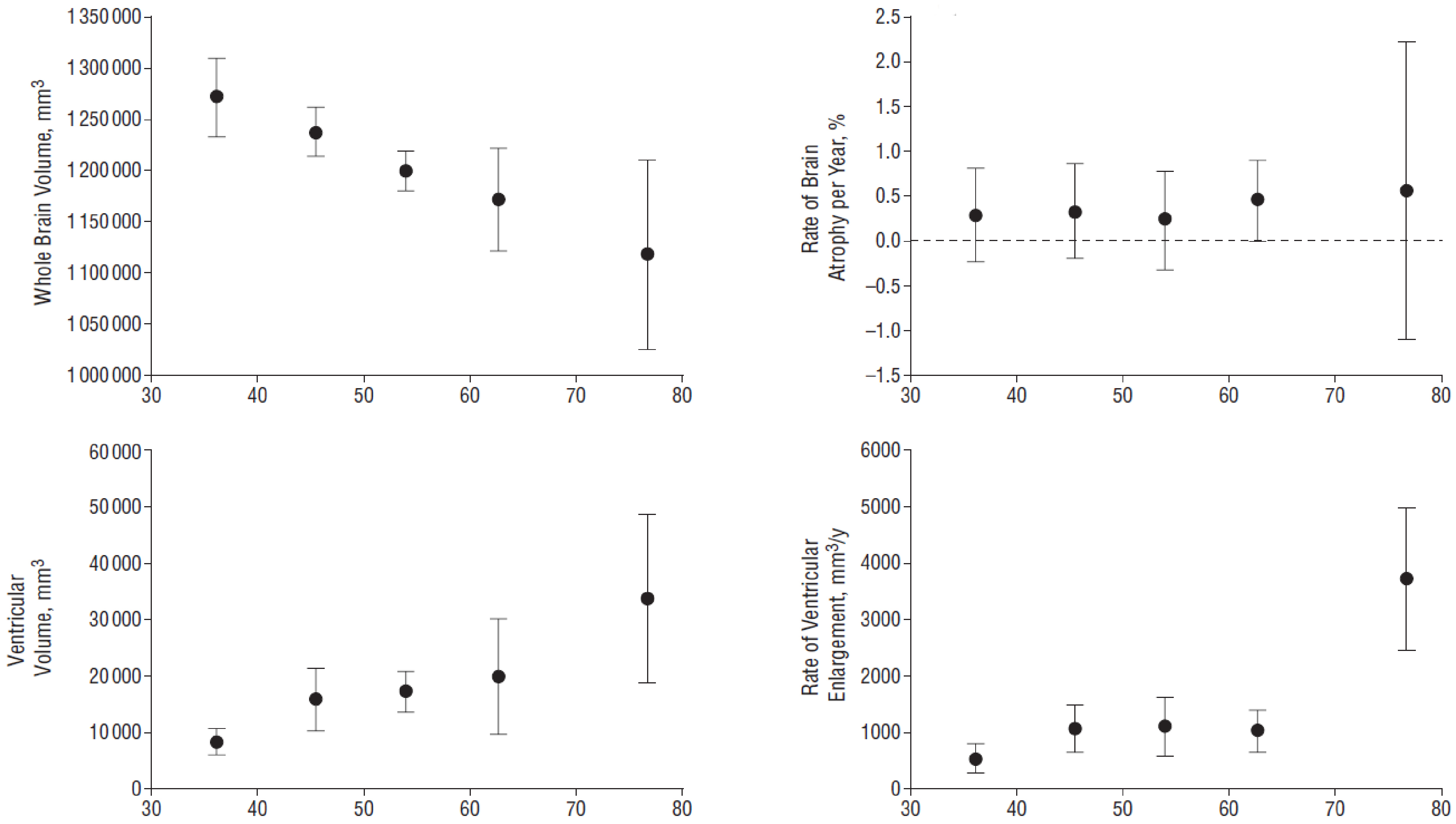
Regional differences in yearly brain volume changes

- Enlargement of ventricular size
  - Increase in cerebrospinal fluid volume
  - With most marked changes occurring after 70 years of age [Schhill et al., 2003]



[MacDonald and Pike, 2021]

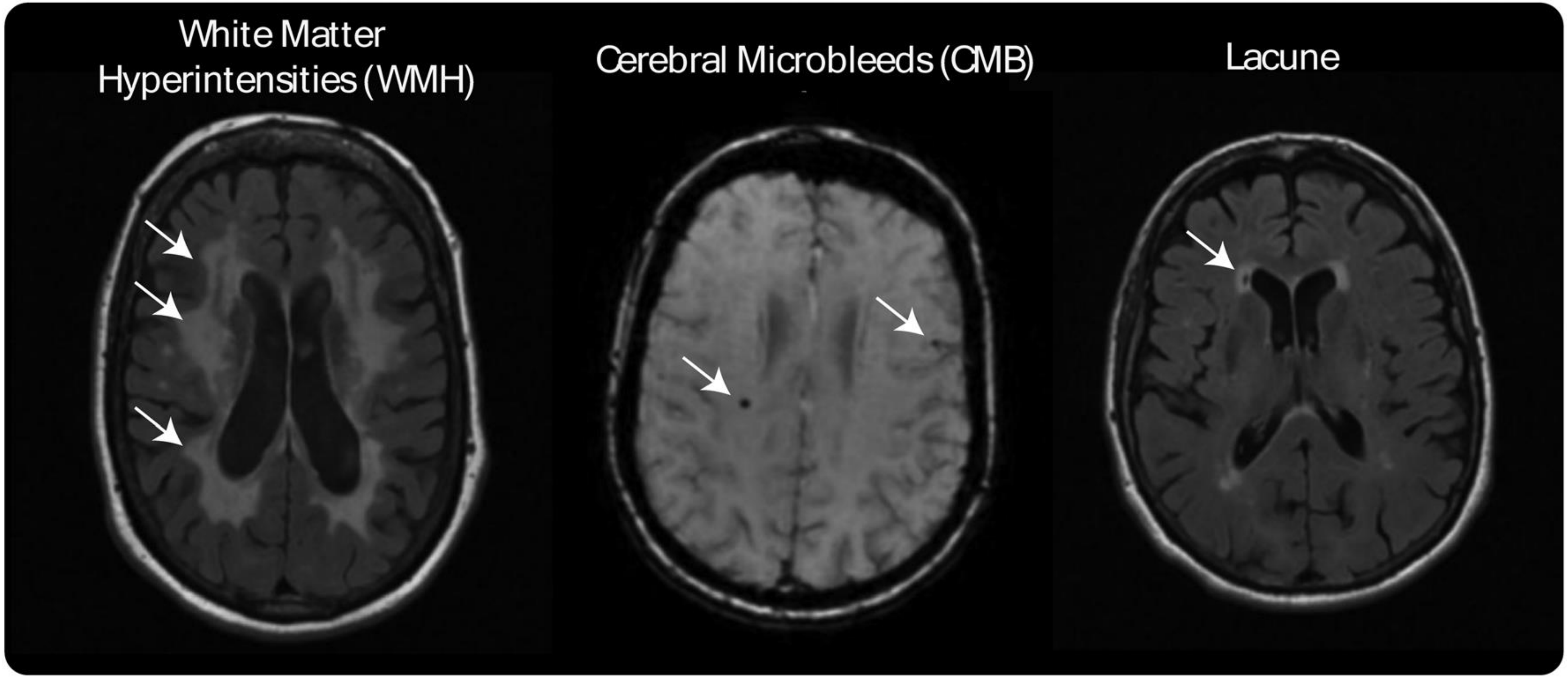
Ventricular size changes with age



[Schhill et al., 2003]

Cross-sectional volume and longitudinal rate of volume changes in the whole brain and ventricles

- Accrual of silent lesions
  - White matter hyperintensities
    - Focal white matter spots that are hyperintense on T2-weighted MRI
  - Cerebral micorbleeds
    - Small hemorrhages caused by rupture of small vessels in basal ganglia or subcortical white matter
  - Lacunar infarcts
    - Small noncortical infarcts caused by occlusion of a single penetrating branch of a large cerebral artery



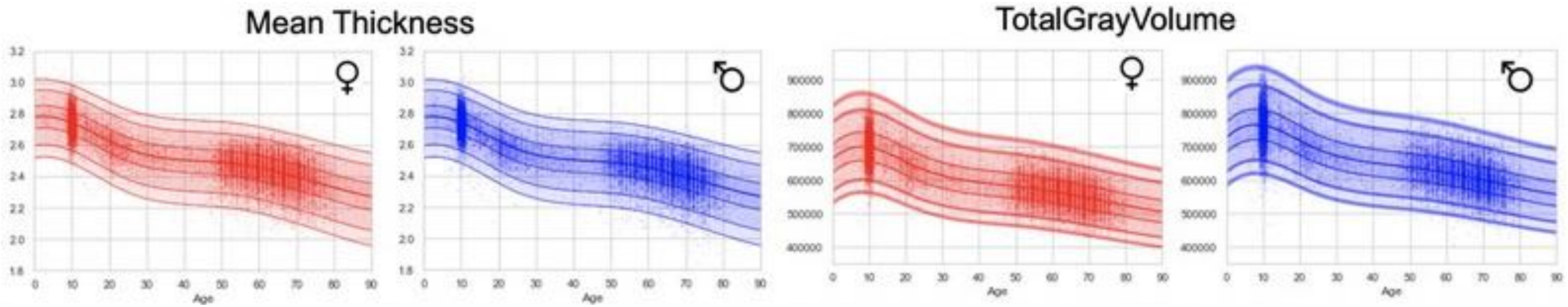
[[MacDonald and Pike, 2021]

Examples of silent lesions

Normative Model

- Reference model for population variation [Rutherford et al., 2022]
  - Enables to quantify individual variation against centiles of variation in a reference population
  - Shifts focus away from group-level (e.g., case-control) inferences to the level of an individual
    - The ability to study individual deviations is essential for understanding inter-individual variability and its relation to the onset and progression of clinical conditions

- Framework for mapping population-level trajectories of the relationships between health-related variables while simultaneously preserving individual-level information [Rutherford et al., 2023]
  - Health-related variables may involve:
    - Demographics (i.e. age and gender)
    - Simple (i.e. height and weight) or complex (i.e. brain structure and function, genetics) biological measures
    - Environmental factors (i.e. urbanicity, pollution)
    - Self-report measures (i.e. social satisfaction, emotional experiences)
    - Behavioural tests (i.e. cognitive ability, spatial reasoning)



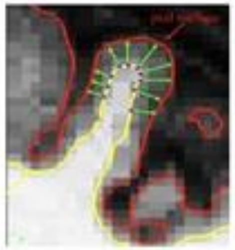
[Rutherford et al., 2022]

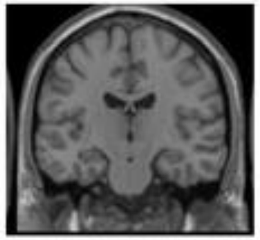
Prediction of cortical thickness and subcortical volume from age

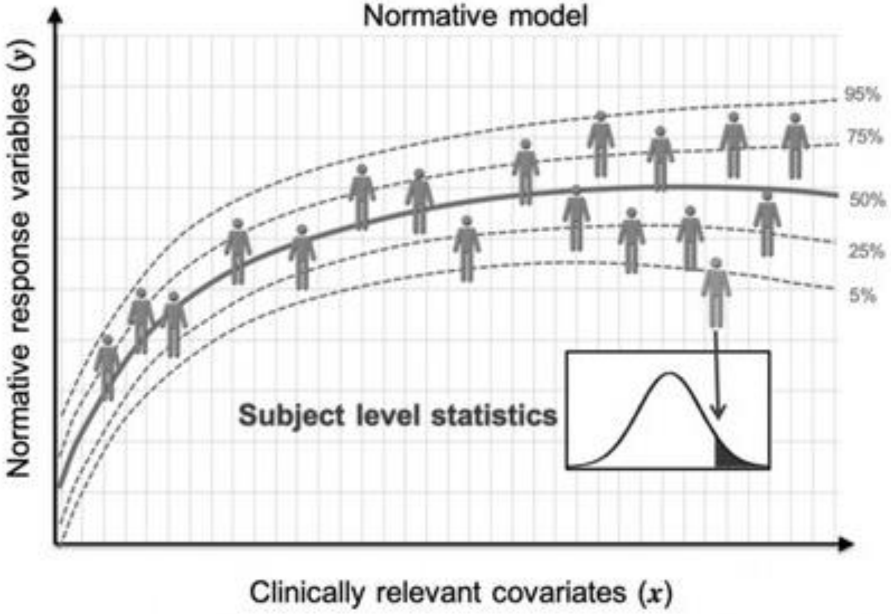
- Deviation score [Rutherford et al., 2023]
  - Output of a normative model
  - Represents where an individual is in comparison to the population the model was estimated on
    - Positive deviation score: greater cortical thickness or subcortical volume than average
    - Negative deviation score: less cortical thickness or subcortical volume than average
  - Advantageous compared to using raw features in regression and classification tasks

Input

Raw Cortical Thickness







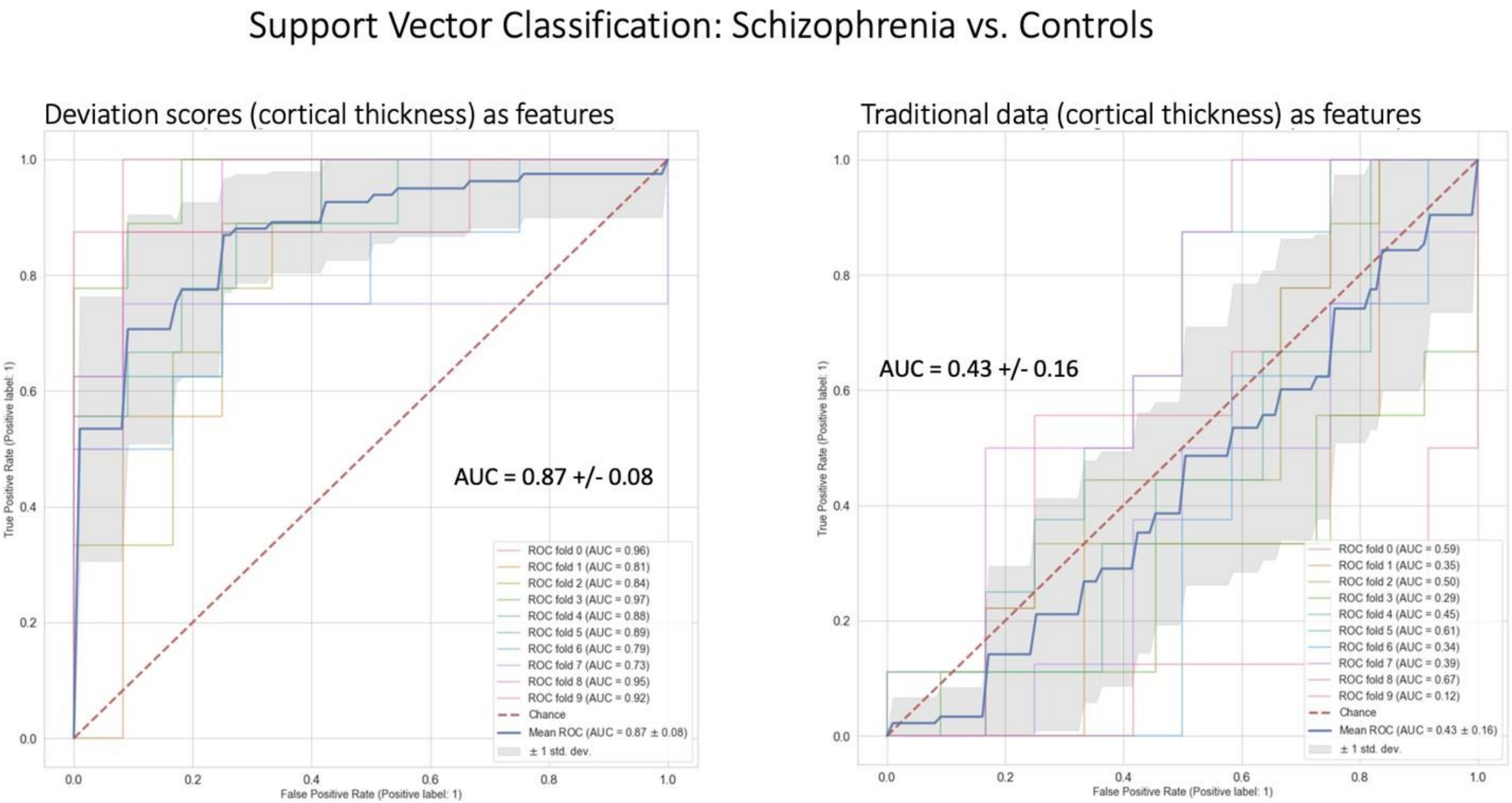
Output

$$z_{nd} = \frac{y_{nd} - \hat{y}_{nd}}{\sqrt{\sigma_d^2 + \left(\sigma_*^2\right)_d}}$$

Cortical Thickness
Deviation (Z) Score

[Rutherford et al., 2023]

Normative model-derived deviation scores that represent individual-level devations



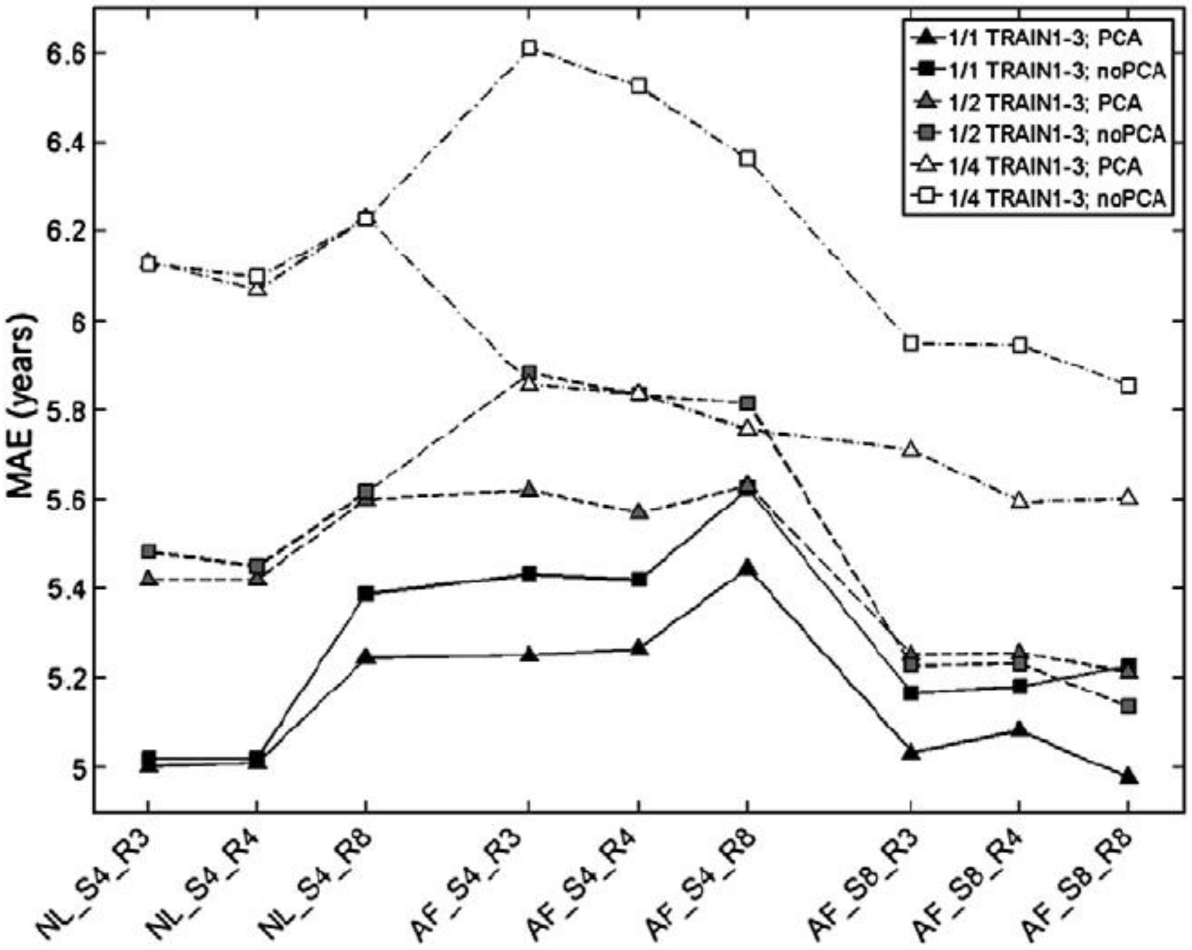
[Rutherford et al., 2023]

Comparison of classification accuracy between using deviation scores and using raw features

- Brain age estimation model as a normative model
  - Describes population-level trajectories of the relationship between brain structure and age
  - Prediction of age from brain structural features
    - Age ~ brain structural features
  - Deviation score
    - Brain age gap = brain age – chronological age

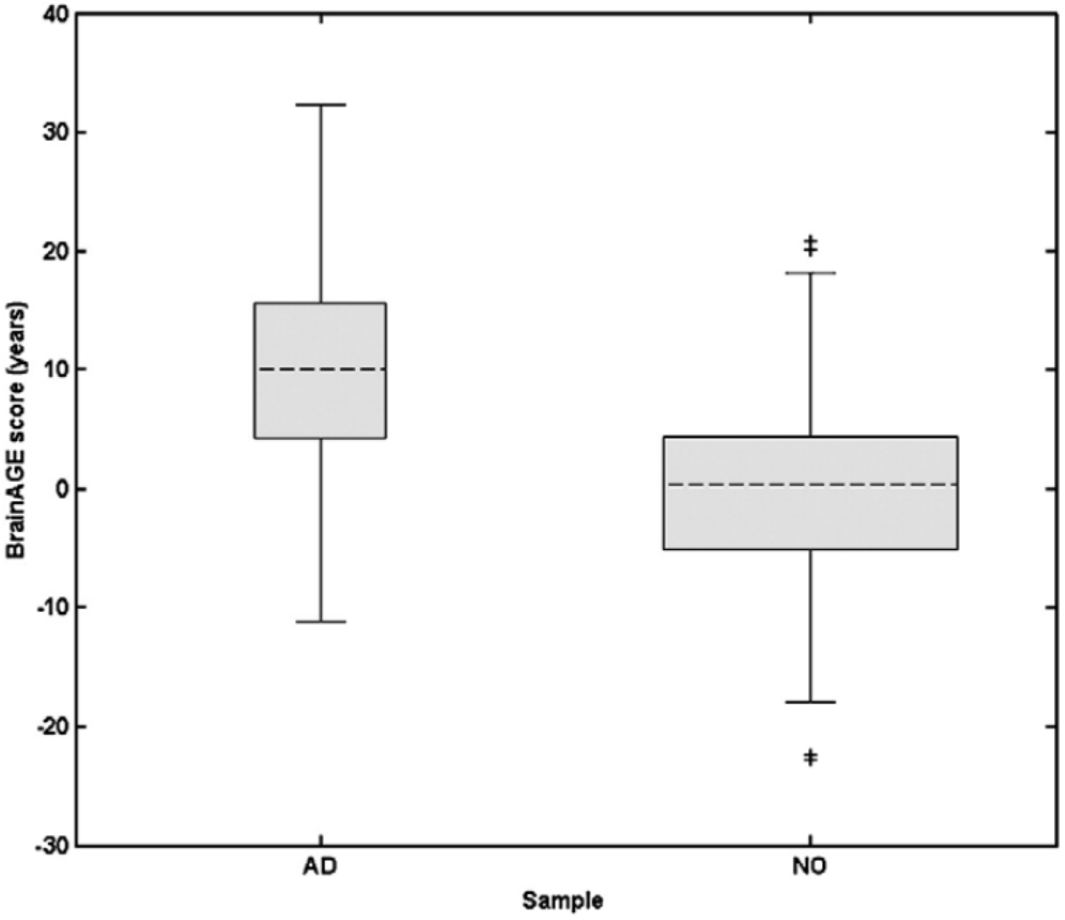
Machine Learning for Brain Age Estimation

- Relevance vector regression
  - Franke et al., 2010
    - Input:
      - Voxel-wise values of grey matter probability → principal component analysis
    - Datasets:
      - Training: n = 410 (20-86 years)
      - Test: n = 137 (19-83 years)
      - External test: n = 108 (20-59 years)
    - Performance:
      - Test: mean absolute error (MAE) = 4.61 years
      - External test: MAE = 5.44 years



[Franke et al., 2010]

Influences of various parameters on the performance of brain age estimation

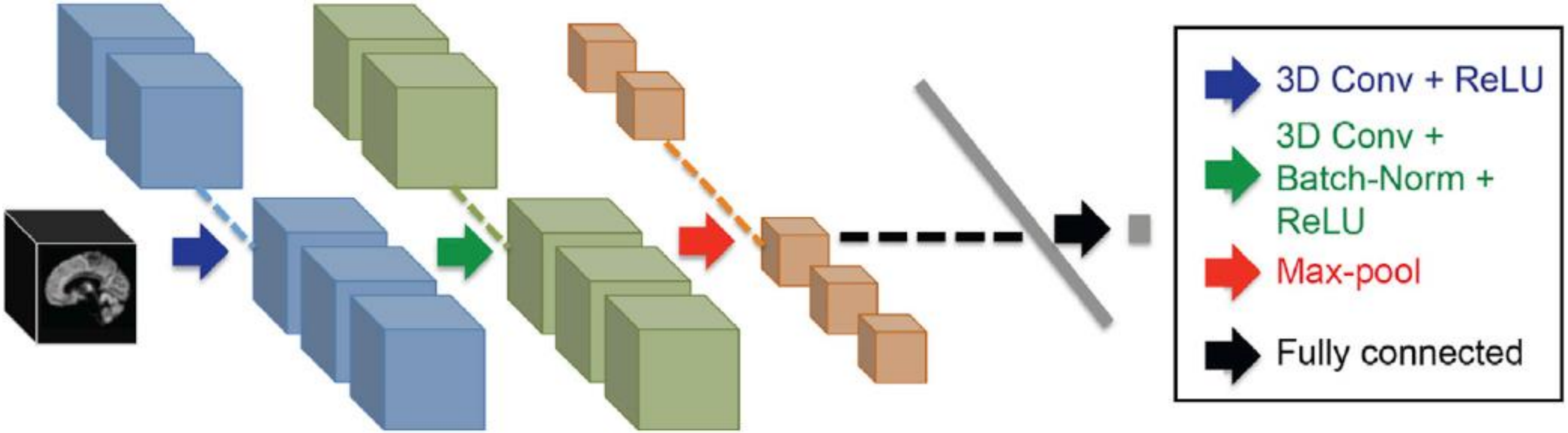


AD, Alzheimer's disease

[Franke et al., 2010]

Comparison of brain age gap

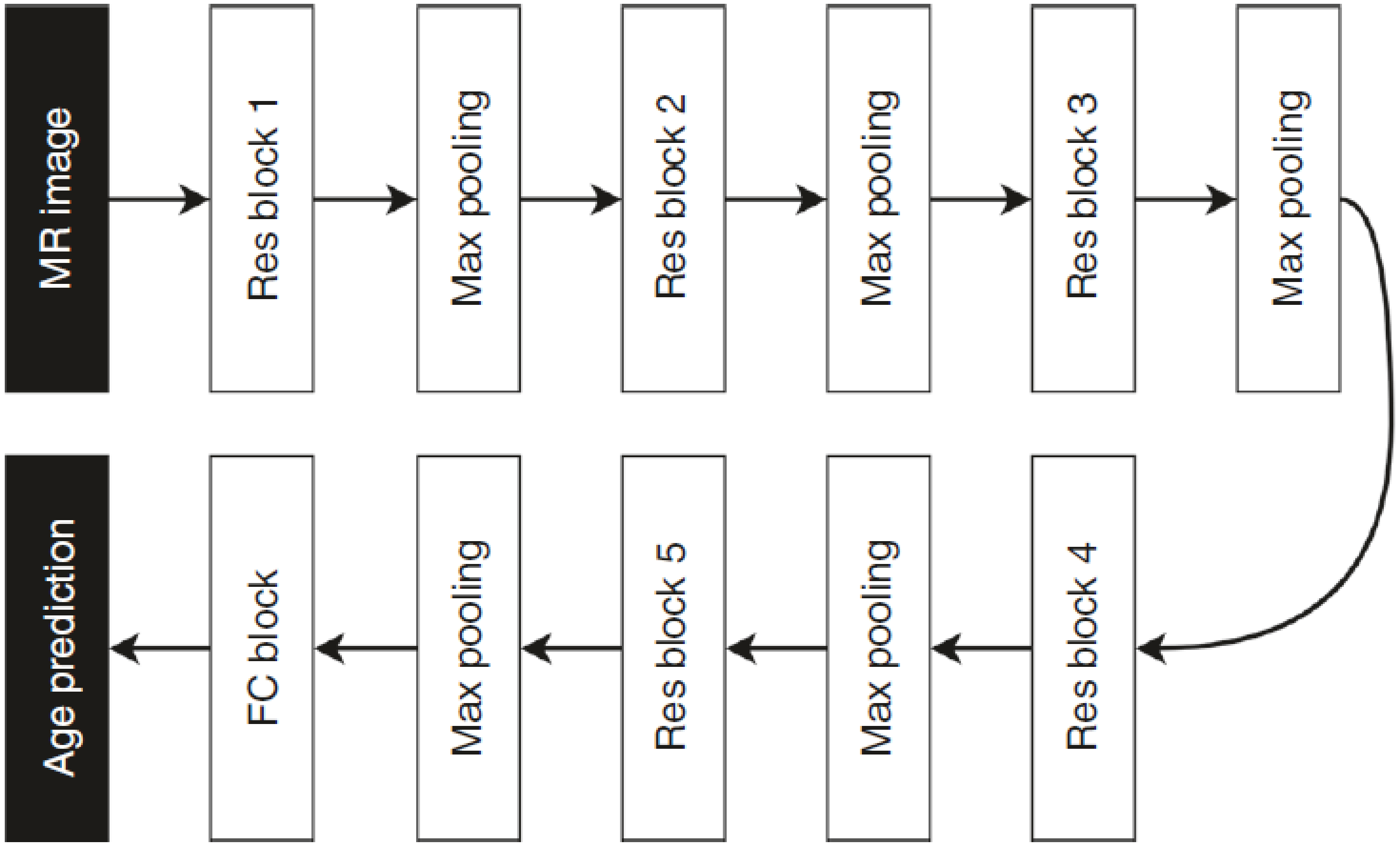
- 3D convolutional neural networks (CNN)
  - Cole et al., 2017
    - Input:
      - T1-weighted brain image (T1)
      - Grey matter probability image (GM)
      - White matter probability image (WM)
    - Datasets (18-90 years):
      - Training: *n* = 1,601
      - Validation: *n* = 200
      - Test: *n* = 200
    - Performance:
      - Test: MAE = 4.16 (GM), 4.34 (GM + WM), 4.65 (T1), 5.14 (WM) years



[Cole et al., 2017]

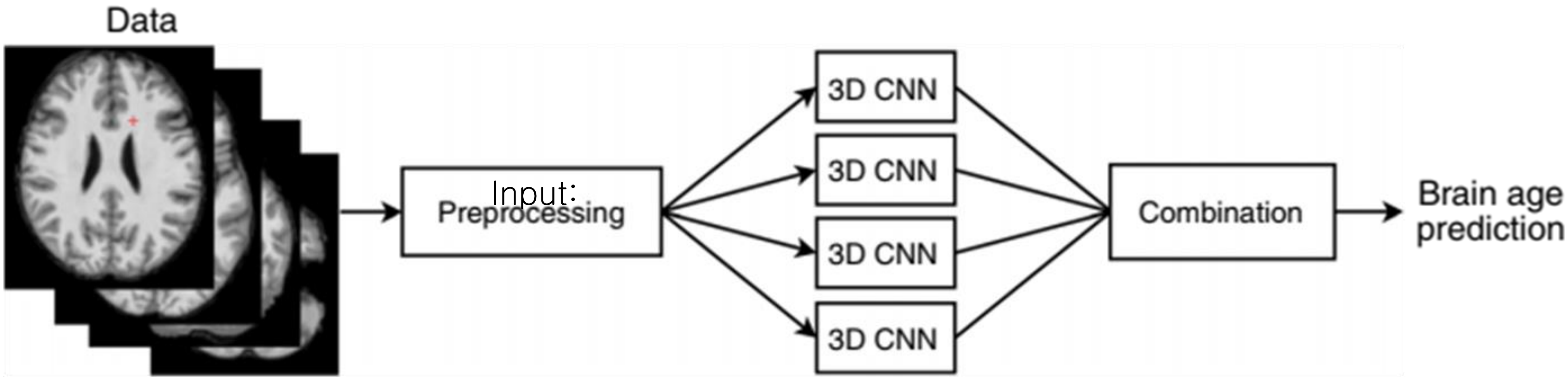
3D CNN

- Jonsson et al., 2019
  - Input:
    - T1-weighted brain image (T1)
    - Jacobian map (JM)
    - Grey matter probability image (GM)
    - White matter probability image (WM)
    - Individuals' sex and MRI scanner type
  - Datasets (18-75 years):
    - Training: n = 809 (1,171 images)
    - Validation: n = 202 (298 images)
    - Test: n = 253 (346 images)
  - Performance
    - Validation: MAE = 3.581 (T1, JM, GM, and WM composition), 3.996 (T1), 4.676 (WM), 4.766 (GM), 4.801 (JM) years
    - Test: MAE = 3.388 (T1, JM, GM, and WM composition), 4.006 (T1), 4.189 (WM), 4.641 (GM), 4.804 (JM) years



[Jonsson et al., 2019]

3D CNN that employed residual blocks



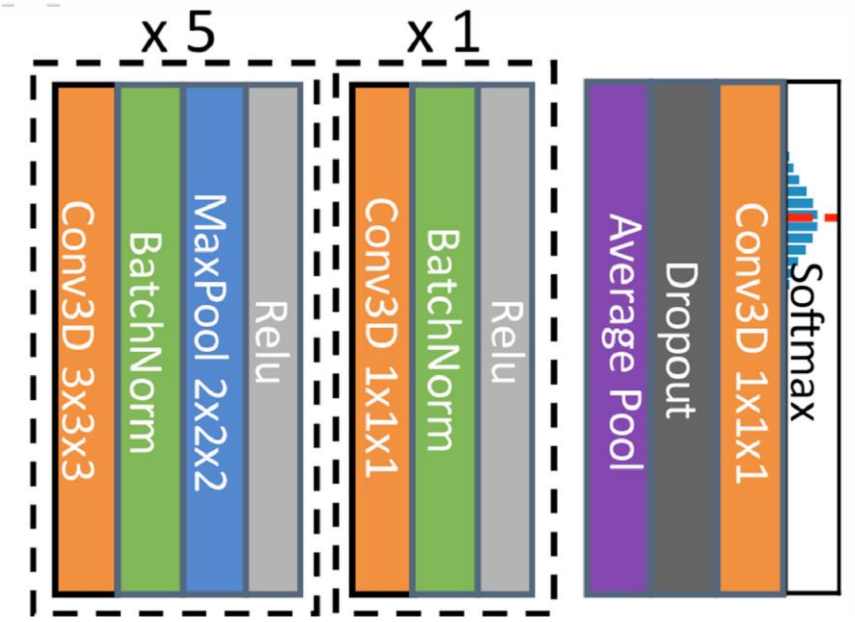
[Jonsson et al., 2019]

Combination of predictions from multiple CNNs by training a linear regression blender

– Peng et al., 2021

- Based on approaches that achieved the first place in the Predictive Analytic Challenge (PAC) 2019
- Input:
  – Linearly registered T1-weighted brain image (T1Lin)
  – Nonlinearly registered T1-weighted brain image (T1Nonlin)
  – Grey matter probability image (GM)
  – White matter probability image (WM)
- Datasets (44-80 years):
  – Training: n = 12,949
  – Validation: n = 518
  – Test: n = 1,036
- Data augmentation
  – Randomly shifted by 0, 1, or 2 voxels along every axis
  – Mirrored with a probability of 50% about the sagittal plane

- Ensemble strategy
  - 5 (identical network structure but randomly-initialised parameters) models trained on each of the 4 input data types
- Performance
  - Simple fully convolutional network (SFCN) with data augmentation and regularization
    - » Train: MAE = 1.36 years (T1Lin)
    - » Validation: MAE = 2.18 years (T1Lin)
    - » Test: MAE = 2.14 years (T1Lin)
  - Model ensemble (n = 2,590 for training)
    - » Test: MAE = 2.58 (T1Lin + T1Nonlin + GM + WM ensemble), 2.62 (T1Nonlin ensemble), 2.71 (T1Lin ensemble), 2.72 (GM ensemble), 2.78 (WM ensemble) years
- https://github.com/ha-ha-ha-han/UKBiobank_deep_pretrain/

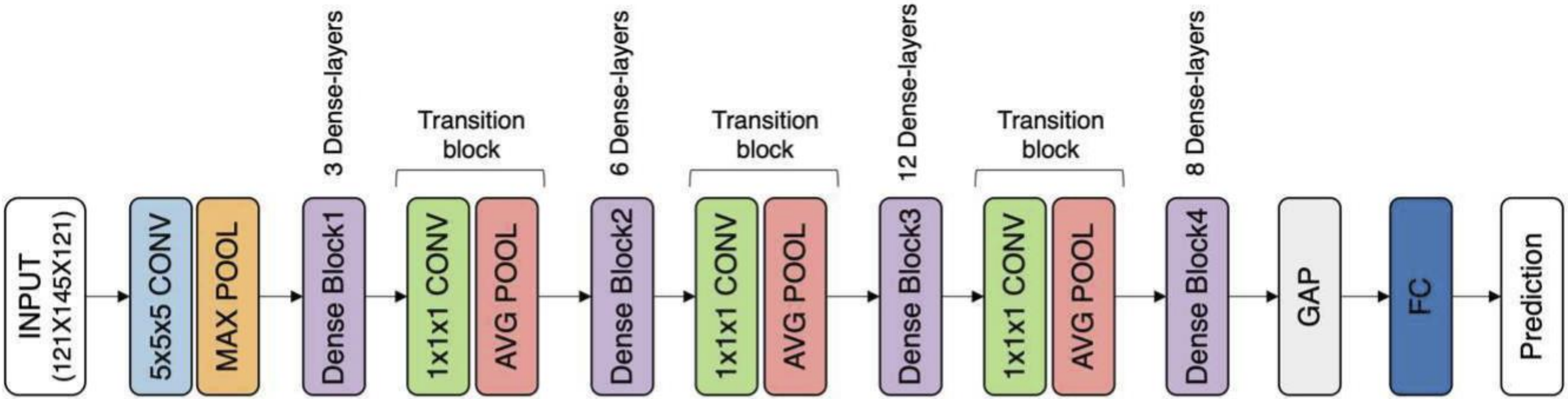


[Peng et al., 2021]

SFCN

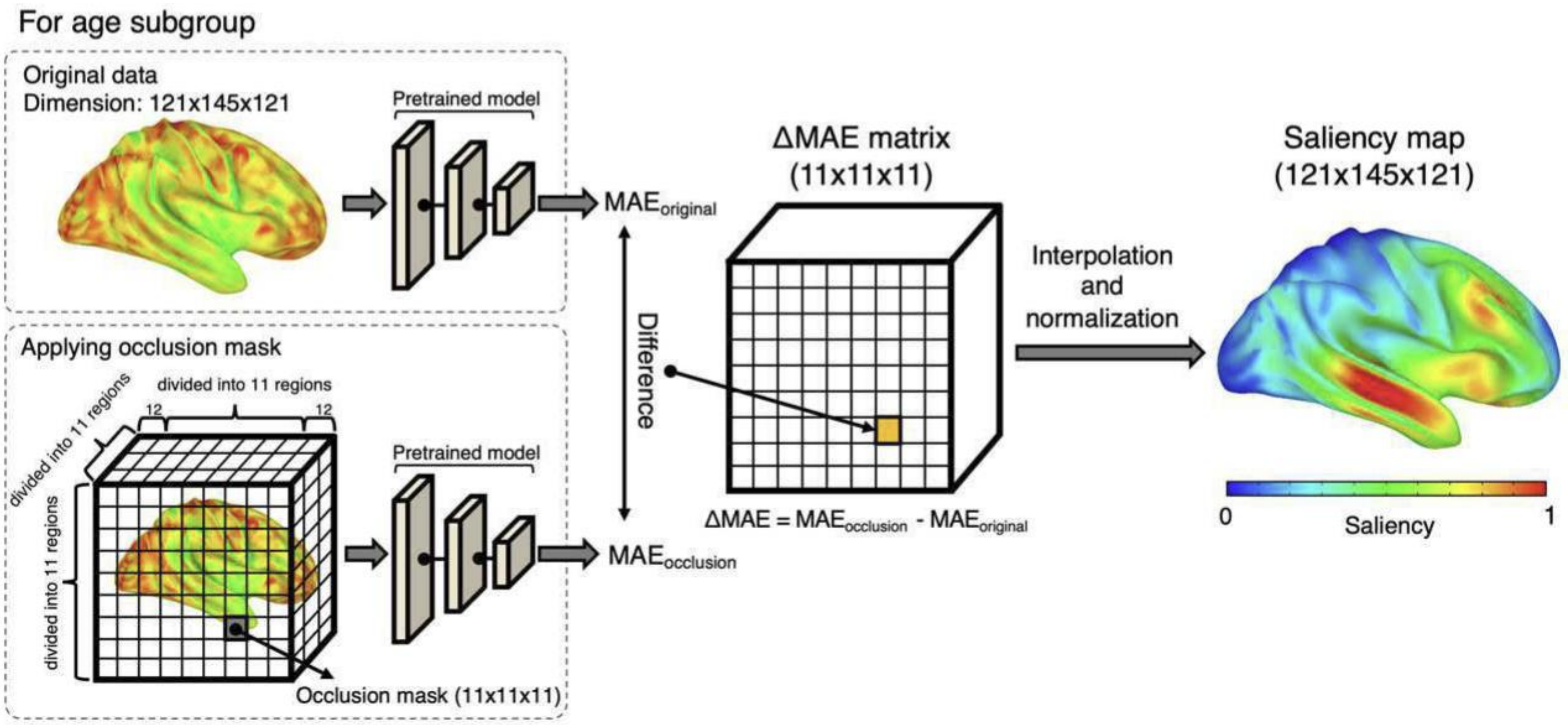
– Lee et al., 2022

- Input:
  - T1-weighted image
- Datasets: n = 1,805 (20-98 years)
  - 5-fold cross validation
    - » Training: 60%
    - » Validation: 20%
    - » Test: 20%
- Performance
  - Test: MAE = 4.2055 years
- Interpretability
  - Through occlusion sensitivity analysis with occlusion masks of 11^3 mm^3
- https://github.com/Neurology-AI-Program/Brain_age_prediction.git



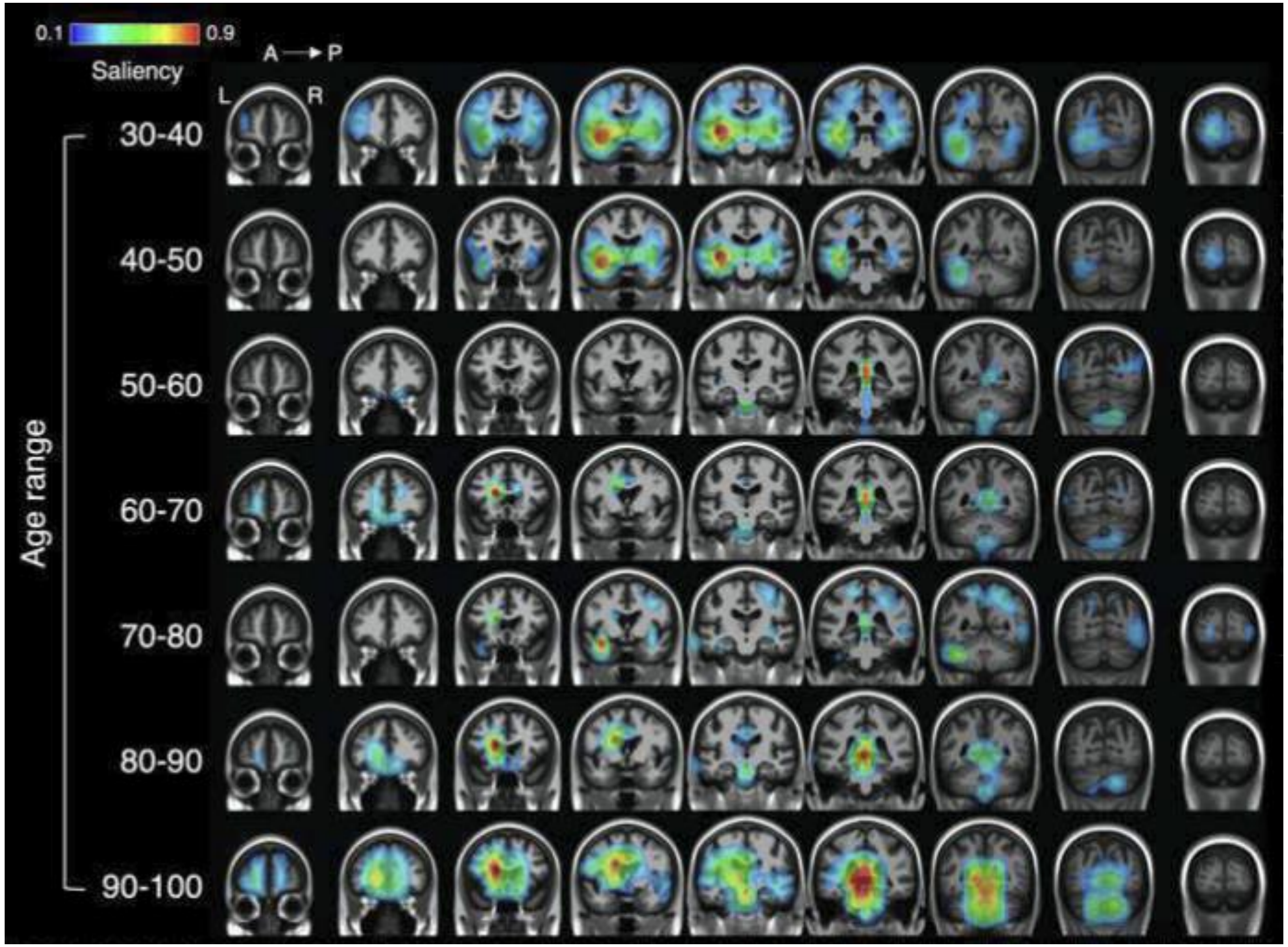
[Lee et al., 2022]

3D CNN that employed dense blocks



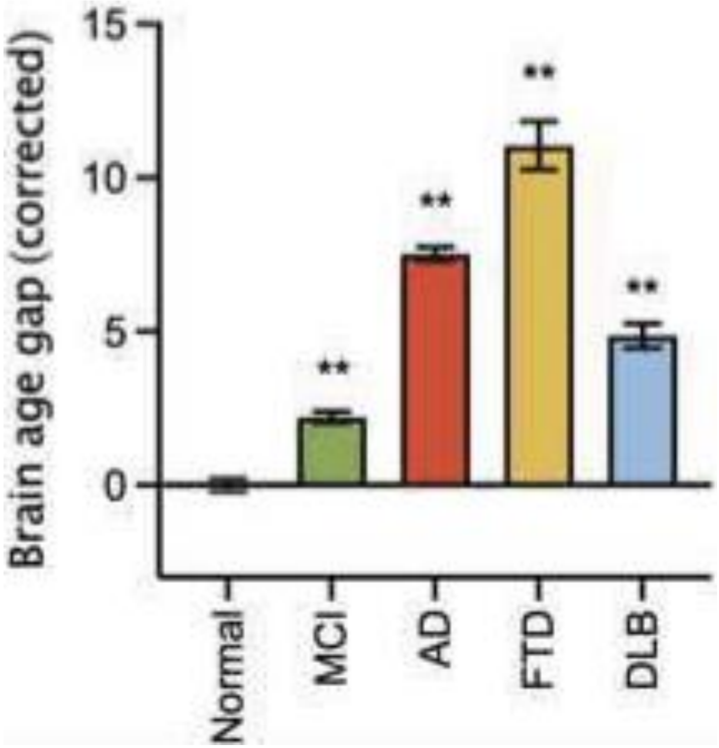
[Lee et al., 2022]

Occlusion sensitivity analysis



[Lee et al., 2022]

Saliency maps across age range groups

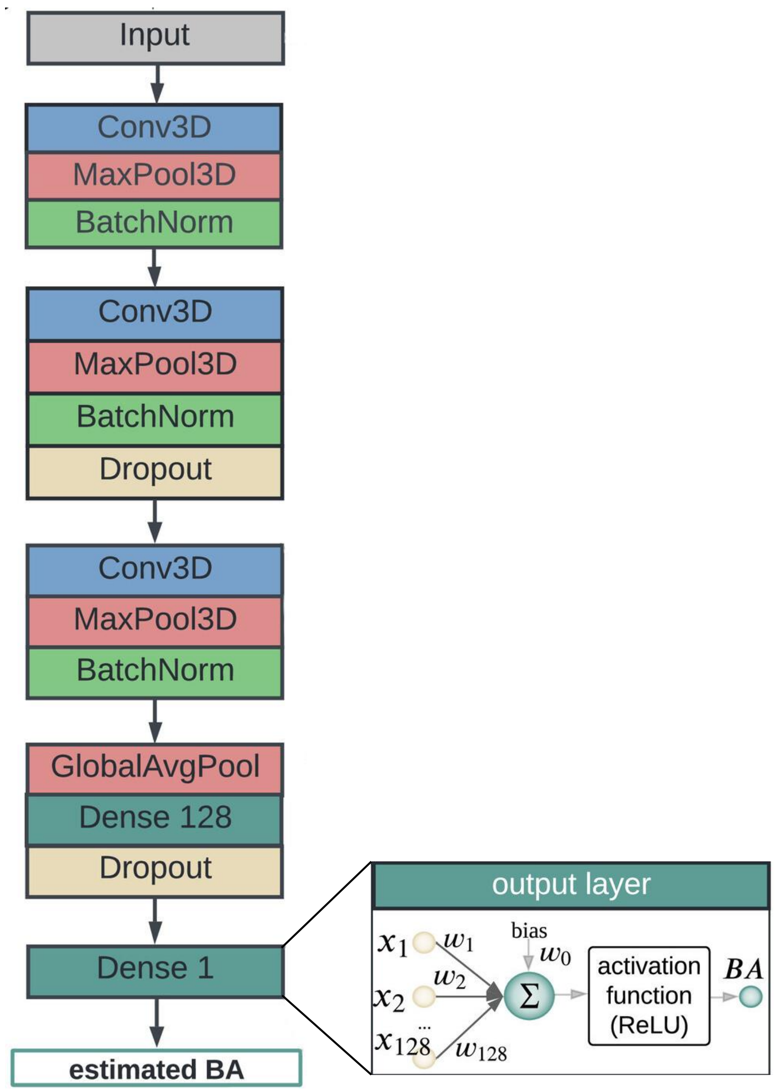


MCI, mild cognitive impairment
AD, Alzheimer's disease
FTD, frontotemporal dementia
DLB, dementia with Lewy bodies

[Lee et al., 2022]

Comparison of brain age gap

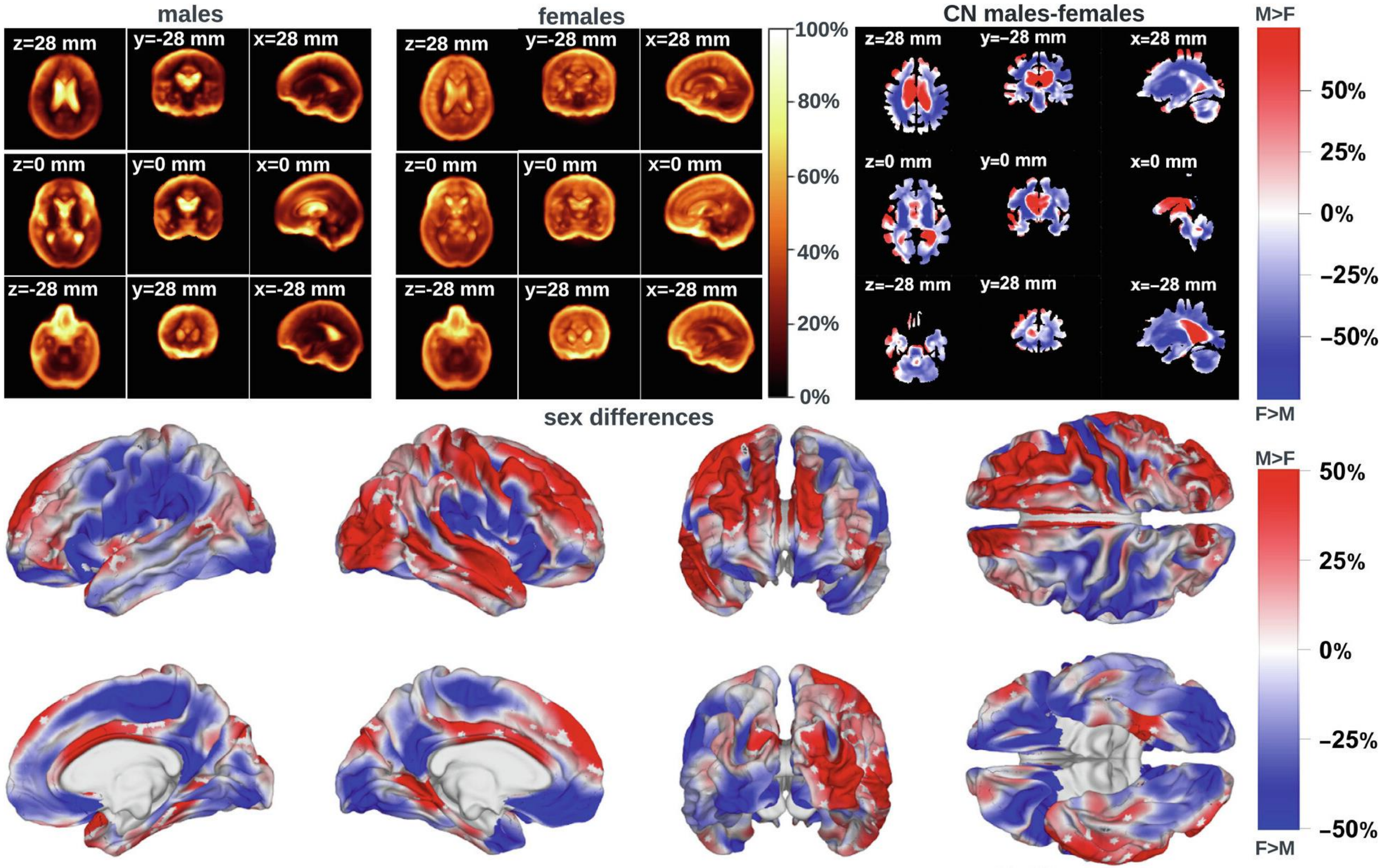
- Yin et al., 2023
  - Input:
    - T1-weighted image (brain.mgz from FreeSurfer)
  - Datasets:
    - Training: *n* = 4,681 (22-95 years)
    - Test: *n* = 1,170 (22-95 years)
    - External test: *n* = 650 (18-88 years)
  - Performance
    - Test: MAE = 2.41 (males), 2.23 (females) years
    - External test: MAE = 3.01 (males), 4.71 (females) years



[Yin et al., 2023]

3D CNN

- Interpretability
  - Through occlusion sensitivity analysis with occlusion masks of 1 mm^3
  - Reveals typical neuroanatomic patterns of aging
    - » Ventricular enlargement
    - » Atrophy of frontal, temporal, and hippocampal cortices
    - » Cortical thinning
- https://github.com/irimia-laboratory/USC_BA_estimator



[Yin et al., 2023]

Comparison of brain saliency maps between sexes

Females

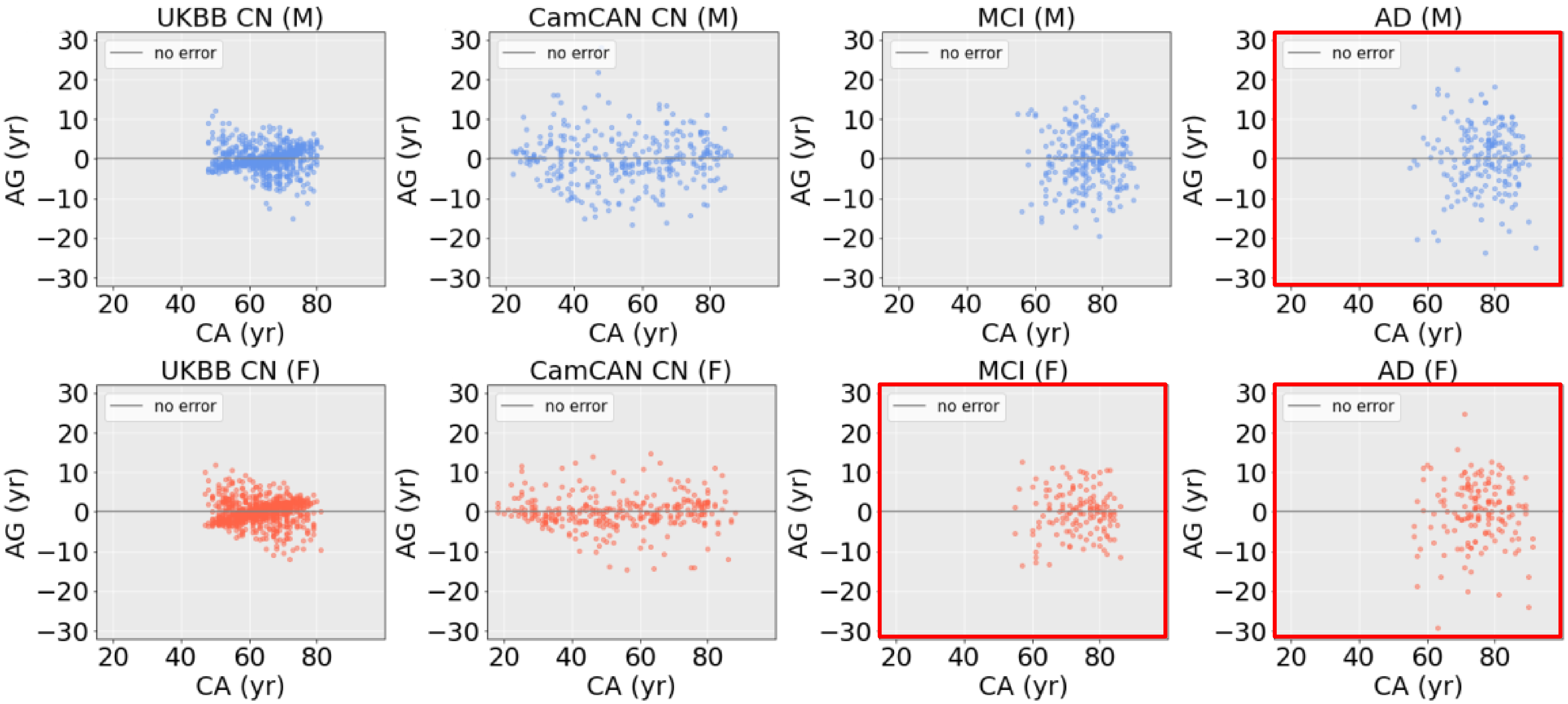
metric	dataset	status	3D-CNN	SFCN	$\Delta[\%]$
MAE	UKBB	CN	2.27	**2.14**	6.07
	CamCAN	CN	**4.71**	8.17	**-29.87**
	ADNI	MCI	**5.26**	7.50	**-42.35**
	ADNI	AD	**6.48**	8.65	**-25.08**
R^2	UKBB	CN	**0.85**	0.84	**1.19**
	CamCAN	CN	**0.95**	0.67	**41.79**
	ADNI	MCI	**0.44**	0.05	**780.00**
	ADNI	AD	**0.21**	0.05	**320.00**

Males

metric	dataset	status	3D-CNN	SFCN	$\Delta[\%]$
MAE	UKBB	CN	2.31	**2.14**	7.94
	CamCAN	CN	**3.01**	9.90	**-69.59**
	ADNI	MCI	**4.33**	7.72	**-43.91**
	ADNI	AD	**5.98**	8.24	**-27.42**
R^2	UKBB	CN	0.83	**0.84**	**1.19**
	CamCAN	CN	**0.90**	0.66	**36.36**
	ADNI	MCI	**0.31**	0.15	**106.67**
	ADNI	AD	**0.17**	0.12	**41.67**

[Yin et al., 2023]

3D CNN vs. SFCN



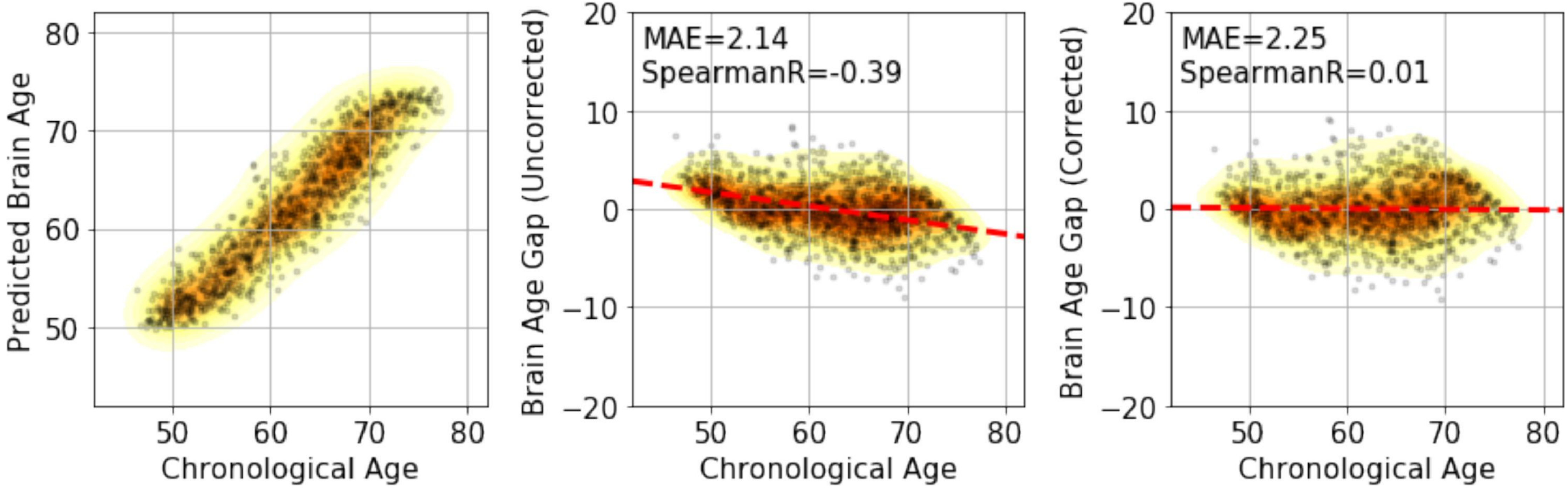
[Yin et al., 2023]

Brain age gap across testing cohorts

Bias in Brain Age Estimation

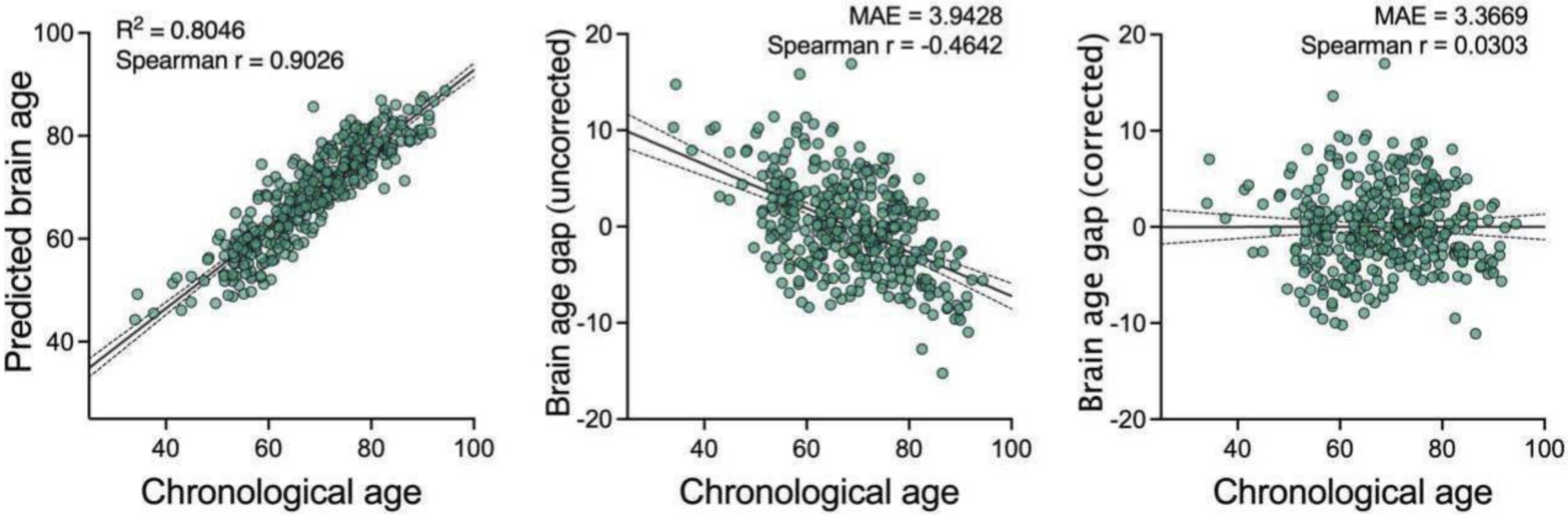
- Tendency to be biased towards the mean age of the total cohort
  - Overestimated brain age in younger individuals, but underestimated brain age in older individuals
- Induces the correlation between chronological age and brain age gap
  - Impacts the relationship between brain age gap and other variables of interest when they are also related to age

- Explained by the concept of 'regression to the mean' (RTM) in statistics
  - For values observed with random error
  - Neither data-dependent nor specific to particular methods including deep learning
- Needs to be adjusted by regressing chronological age on brain age or brain age gap to provide corrected brain age gap
  - (brain age) ~ a × (chronological age) + b [Liang et al., 2019]
  - (brain age gap) ~ a × (chronological age) + b [Le et al., 2018]



[Peng et al., 2021]

Correction of brain age gap

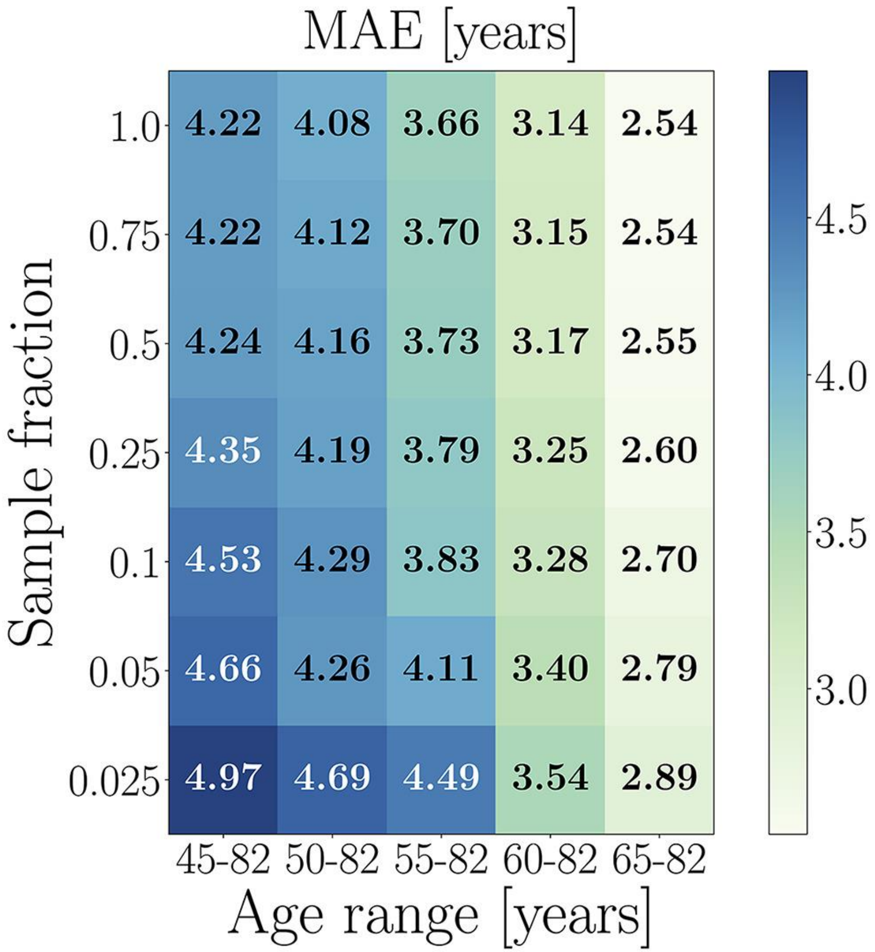


[Lee et al., 2022]

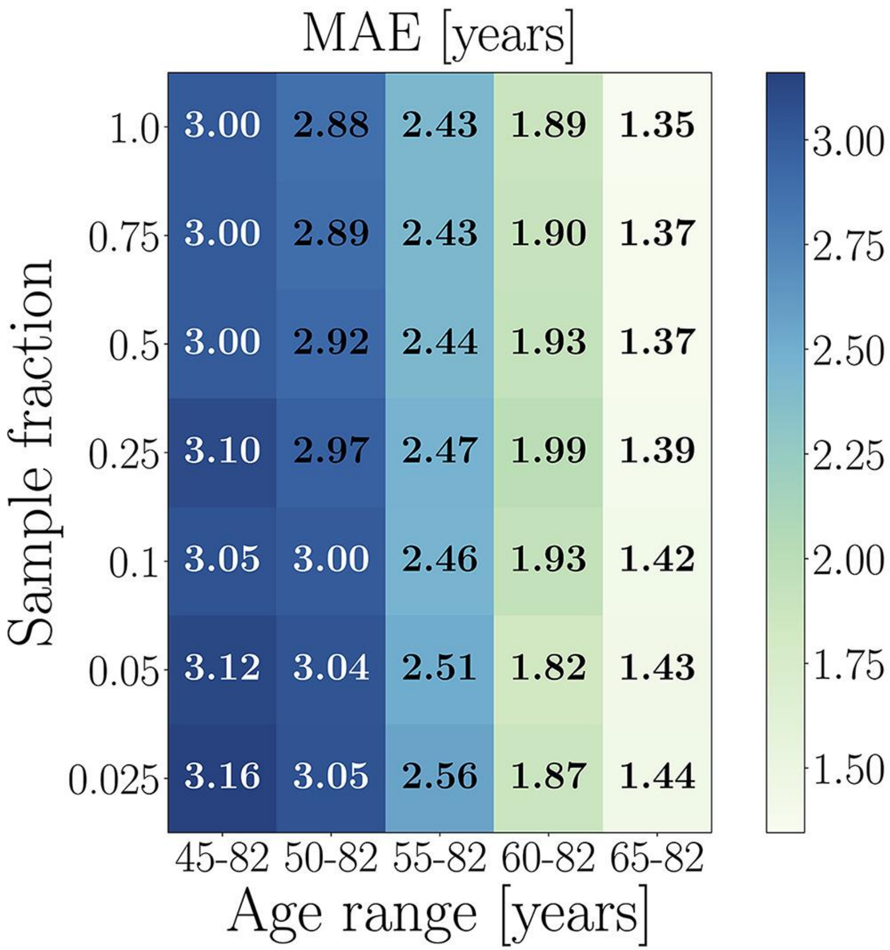
Correction of brain age gap

Effects of Age Range and Sample Size on Brain Age Estimation

- Better performance in samples with a narrower age range
  - Due to smaller error when predictions are closer to the mean age of the total cohort
- Better performance for larger sample sizes across different age ranges



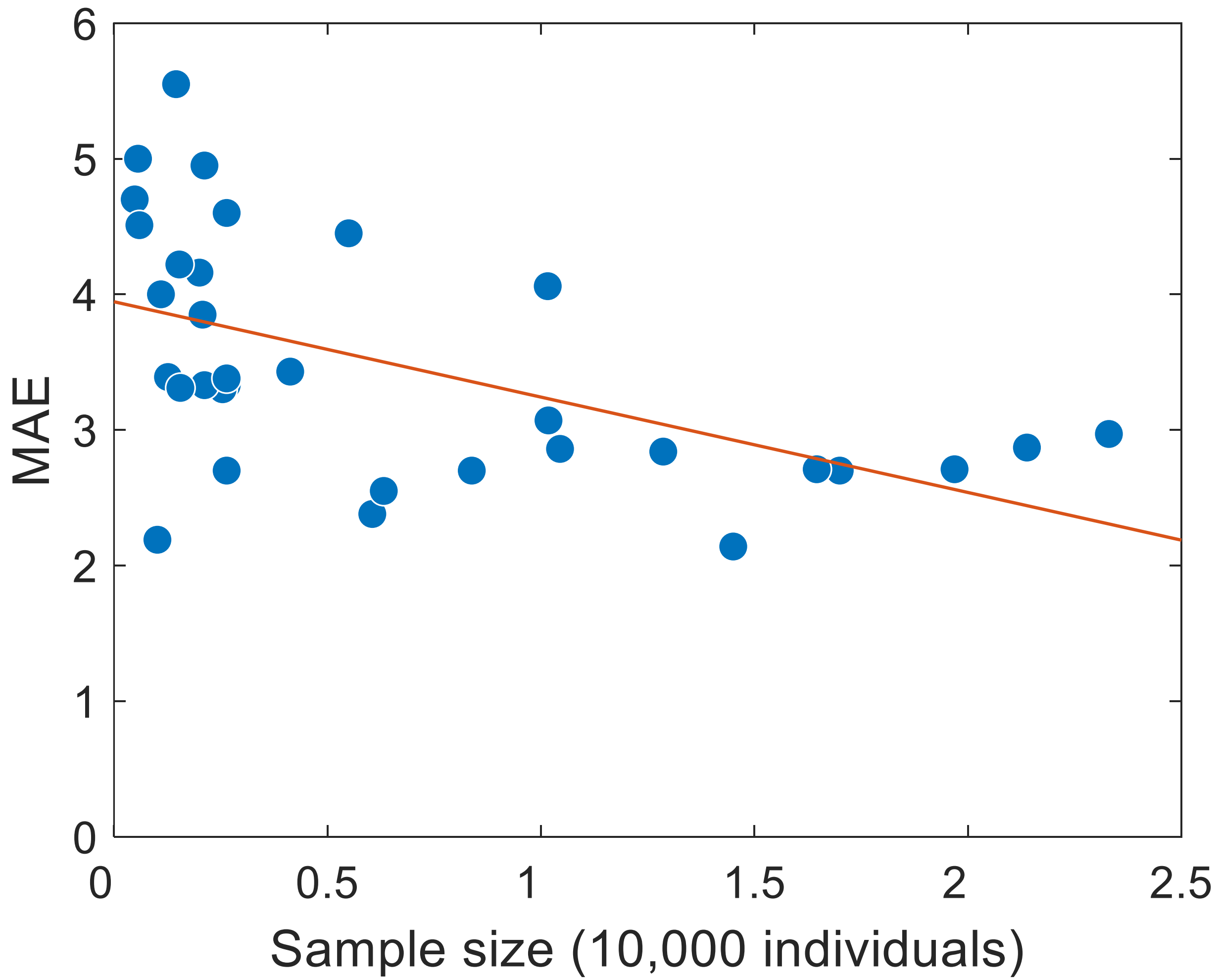
Without bias adjustment



With bias adjustment

[de Lange et al., 2022]

Comparison of performance for different age ranges and sample sizes

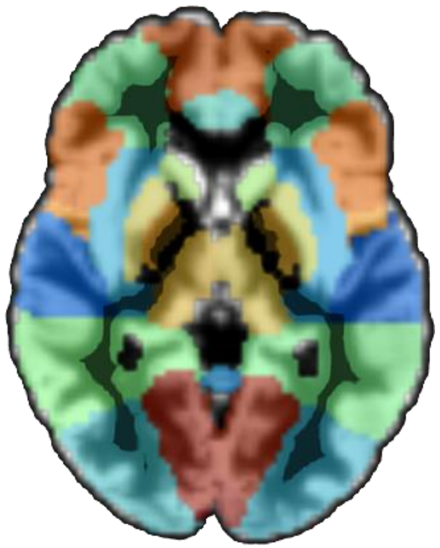


Relationship between sample size and performance

Hands on Machine Learning Modelling for Brain Age Estimation

- Support vector regressor
  - Features [https://github.com/hauin/MedicalBioResearchTopics2/blob/main/10_20231109/X_train.txt]

	Response	Confounding		Predictor		
	Age	TIV	Sex	Region 1 GM volume	…	Region 60 GM volume
Subject 1	-	-	-	-	-	-
Subject 2	-	-	-	-	-	-
Subject 3	-	-	-	-	-	-
⋮	-	-	-	-	-	-



– Performance in 5-fold cross validation

Input	Training	Test
GM (60 features)	6.5±0.3 years	7.2±0.2 years
WM (48 features)	6.3±0.2 years	5.8±0.2 years
GM + WM (108 features)	5.4±0.3 years	5.8±0.2 years
GM (60 features) and WM (48 features) combination	5.5±0.3 years	5.6±0.3 years

- 3D CNN
  - `Regressor` in MONAI

    [https://github.com/hauin/MedicalBioResearchTopics2/blob/main/10_20231109/Age_Regressor.ipynb]

```
==========================================================================================
Layer (type:depth-idx)                                  Output Shape              Param #
==========================================================================================
Regressor                                               [5, 1]                    --
├─Sequential: 1-1                                       [5, 128, 4, 4, 4]         --
│    └─ResidualUnit: 2-1                                [5, 16, 32, 32, 32]       --
│    │    └─Conv3d: 3-1                                 [5, 16, 32, 32, 32]       448
│    │    └─Sequential: 3-2                             [5, 16, 32, 32, 32]       7,378
│    └─ResidualUnit: 2-2                                [5, 32, 16, 16, 16]       --
│    │    └─Conv3d: 3-3                                 [5, 32, 16, 16, 16]       13,856
│    │    └─Sequential: 3-4                             [5, 32, 16, 16, 16]       41,538
│    └─ResidualUnit: 2-3                                [5, 64, 8, 8, 8]          --
│    │    └─Conv3d: 3-5                                 [5, 64, 8, 8, 8]          55,360
│    │    └─Sequential: 3-6                             [5, 64, 8, 8, 8]          166,018
│    └─ResidualUnit: 2-4                                [5, 128, 4, 4, 4]         --
│    │    └─Conv3d: 3-7                                 [5, 128, 4, 4, 4]         221,312
│    │    └─Sequential: 3-8                             [5, 128, 4, 4, 4]         663,810
├─Sequential: 1-2                                       [5, 1]                    --
│    └─Flatten: 2-5                                     [5, 8192]                 --
│    └─Linear: 2-6                                      [5, 1]                    8,193
├─Reshape: 1-3                                          [5, 1]                    --
==========================================================================================
```

– Performance

Input	Validation	Test
T1	5.3 years	5.5 years
GM	4.9 years	6.0 years
WM	4.7 years	4.7 years