Medical/Bio Research Topics Ⅱ: Week 09 (02.11.2023)

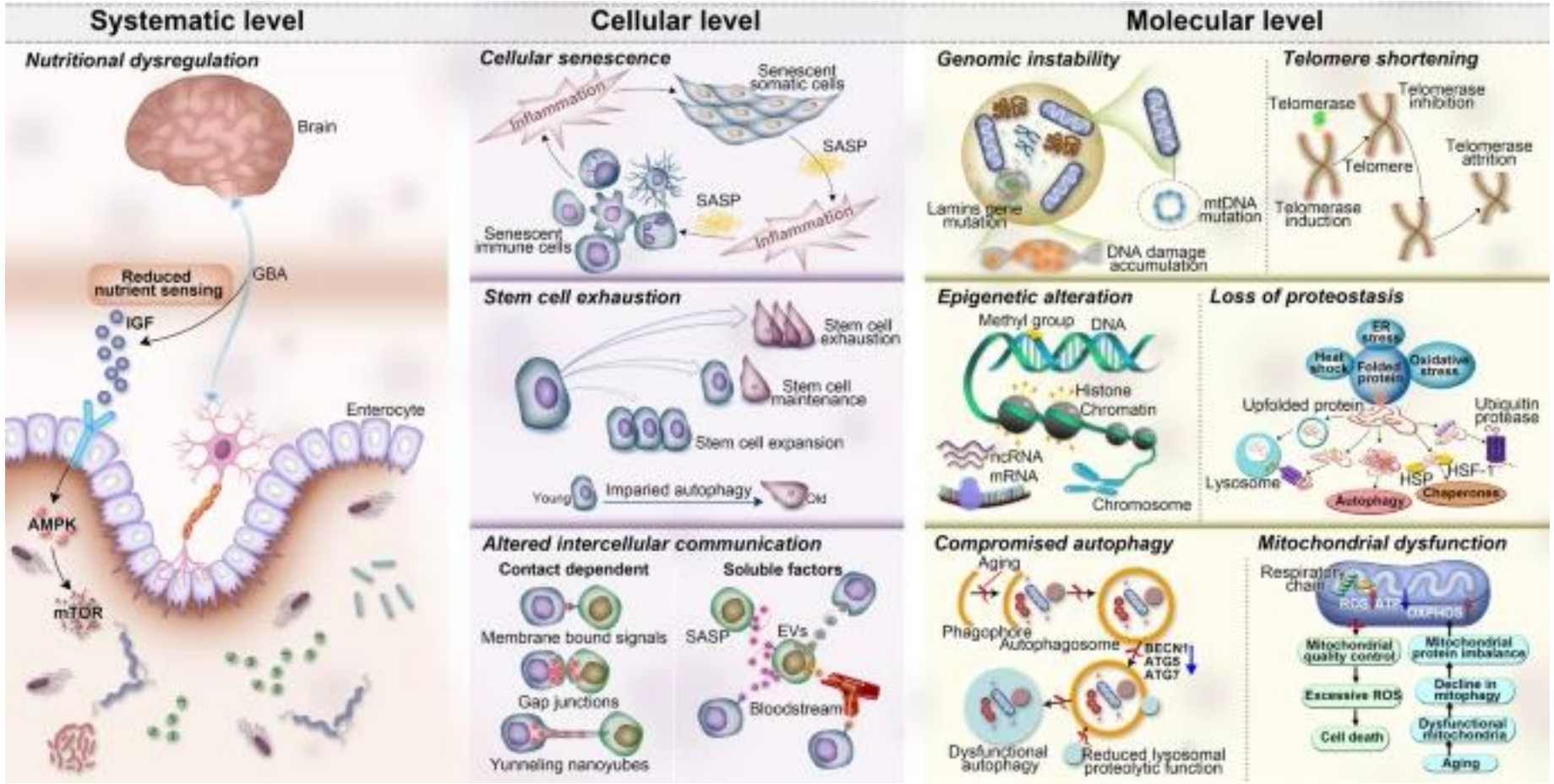
Brain age estimation artificial intelligence models (1): data and prediction problem

**(뇌나이 예측 인공지능 모델 개발 연습 (1):
데이터 및 예측 문제)**

Ageing

- Process of becoming older
  - Refers mainly to humans
- Biological basis [Jin, 2010]
  - Neither theory appears to be fully satisfactory
  - Two main theoretical categories:
    - Genetically programmed
      - Biological timetable, perhaps a continuation of childhood growth and development
    - Damage or error-related
      - Environmental assaults to living organisms

- Hallmarks of ageing [Guo et al., 2022]
  - Genomic instability
  - Telomere dysfunction
  - Epigenetic alterations
  - Loss of proteostasis
  - Compromise of autophagy
  - Mitochondrial dysfunction
  - Cellular senescence
  - Stem cell exhaustion
  - Altered intercellular communication
  - Deregulated nutrient sensing

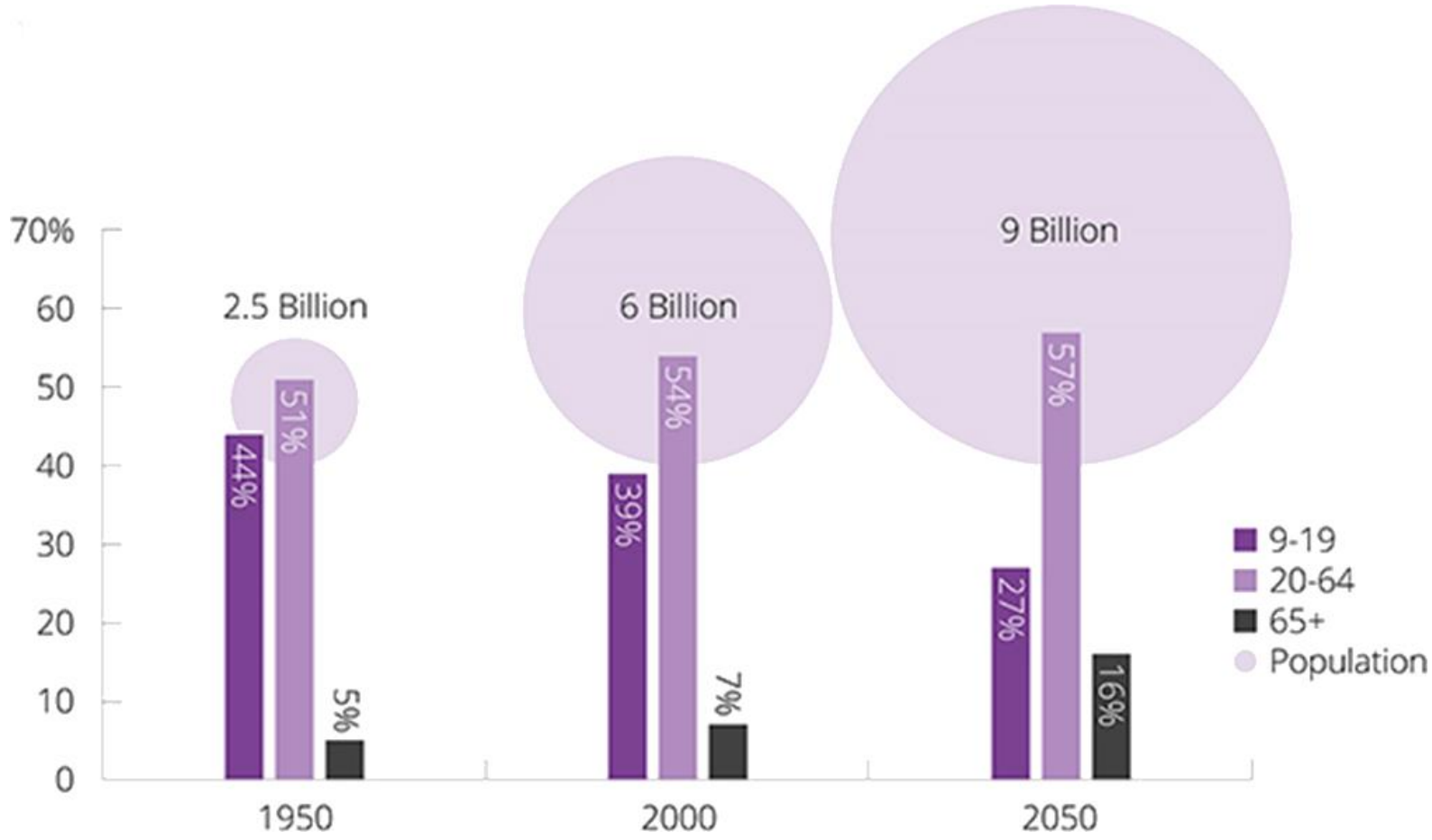


[Guo et al., 2022]

Ten hallmarks of ageing subdivided into three categories

- Changes by ageing
  - Accumulation of a wide variety of molecular and cellular damage over time
    → gradual decrease in physical and mental capacity
    → growing risk of disease
    → death
  - Neither linear nor consistent
  - Only loosely associated with an individual's age in years
  - Often associated with other life transitions such as retirement, relocation to more appropriate housing, and the death of friends and partners

- Population ageing
  - Shift in distribution of a country's population towards older ages
    - Increasing median age in a population
    - By 2030, 1 in 6 people in the world will be aged 60 years or over [https://www.who.int/news-room/fact-sheets/detail/ageing-and-health]
  - Because of declining fertility rates and rising life expectancy
  - Started in high-income countries and now extended to low- and middle-income countries



[http://study-aids.co.uk/dissertation-blog/population-ageing/]

Size and proportion of the global population as related to age over time

- Healthy ageing [https://www.who.int/news-room/fact-sheets/detail/ageing-and-health]
  - Proposed by the World Health Organization
  - Importance of an individual's healthspan (years lived in good health) in their lifespan (total years lived) with longevity
    - Related to non-random diversity seen in older age
  - Operationalises health as functional ability, which results from the interactions of intrinsic capacity and the environments
    - Although some of the variations in older people's health are genetic, most is due to their physical and social environments as well as their personal characteristics

Biological Age

- Two types of age
  - Chronological age
    - Number of years an individual has been alive
    - Unchangeable
  - Biological age
    - How old an individual's cells and tissues appear to be based on their current condition
    - May be reversed

CHRONOLOGICAL AGE

- The number of years that have passed since our birth.
- Cannot be influenced by lifestyle and eating habits.
- Has little relevance to how you feel and function.

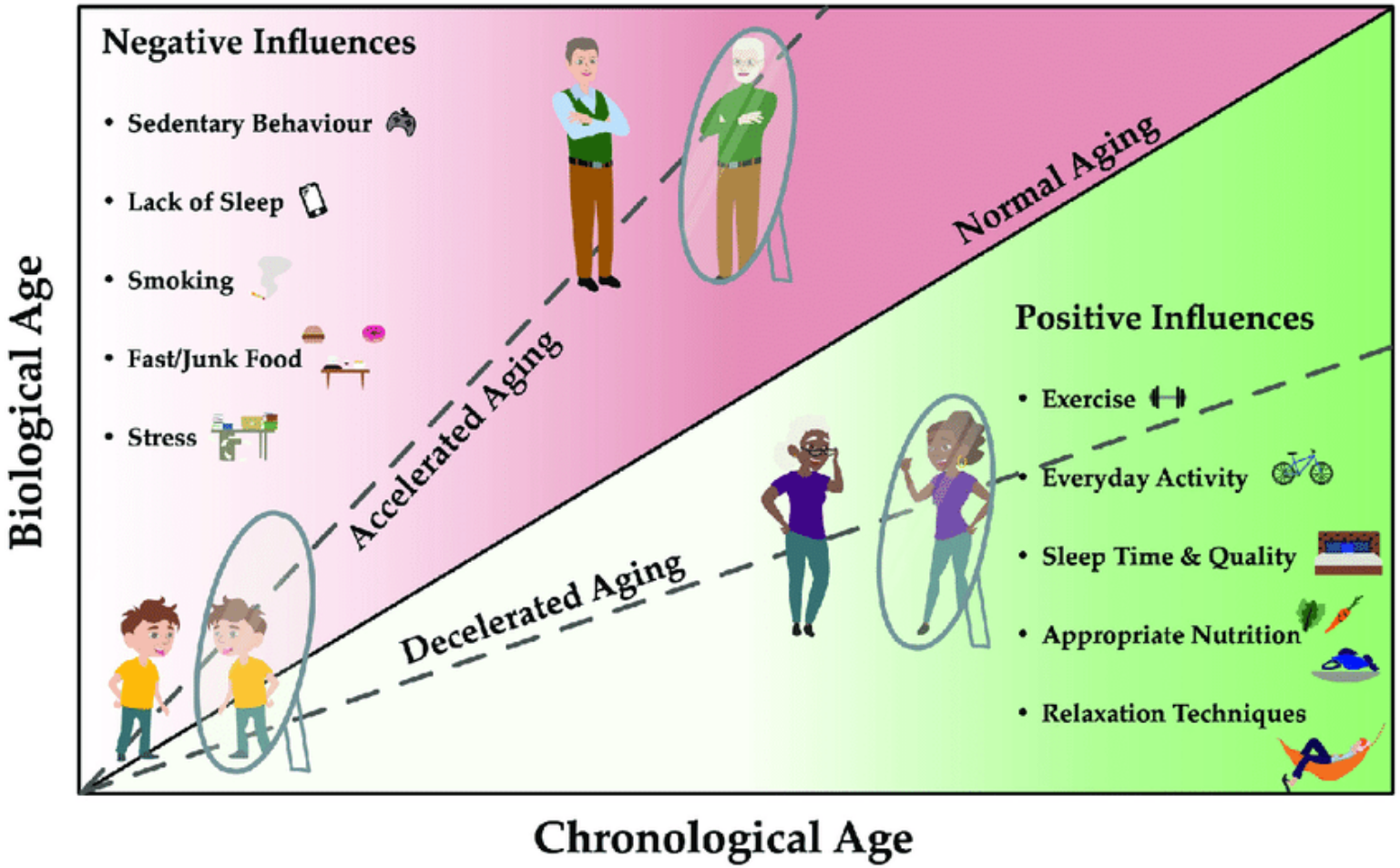
BIOLOGICAL AGE

- How old our cells really are, therefore, our real age
- Can be reversed by attending to your health
- The most important component to the aging process

[https://reyouvenate.com/truage/]

Chronological vs. biological age

- Departure of biological age from chronological age
  - May exhibit greatly different susceptibilities to age-related diseases and death for individuals of the same chronological age
  - Likely reflective of differences in underlying biological ageing processes
- Biological age as a biomarker of ageing
  - Individual-level measure of ageing that captures inter-individual differences in the timing of disease onset, functional decline, and death over the life course
  - Ccrucial to enable evaluation of interventions aimed at promoting healthier aging, by providing a measurable outcome



[Haupt et al., 2022]

Factors that have the potential to both delay and accelerate ageing processes

- How biological age is determined
  - Often by assessing an individual's genetic material
    - Telomeres (repeats of a hexametric DNA sequence capping the end of chromosomes)
      - Shorter telomeres → older biological age
    - DNA methylation (DNAm, hypo- and hyper-methylation changes at many regions across the genome)

TruAge™ vs. Everyone Else

What we do that Zymo and Elysium don't

While Zymo (MyDNAge) and Elysium (Index) market their biological age test to consumers all over the world, there are a few things they don't do.

	TruAge™	MyDNAge	Index
Test over 900,000 Loci	✓	✗	✗
Custom Algorithm	✓	✗	✗
Track your lab and health information in one place	✓	✗	✗
Validated Collection Method	✓	✗	✗
Additional Testing	TruTelomere™	✗	✗
Report	50+ Pages of Interpreted Data and Clinical Recommendations	1 Page Report	3 Page Report
Time it Takes to Report Results	4-6 Weeks	6-8 Weeks	6+ Weeks

– A combination of numerous potential markers that specifically measures all important aspects of ageing processes may be the key to a valid composite biomarker of ageing [Hartmann et al., 2021]
  - Routine laboratory
  - Epigenetic
  - Non-epigenetic
  - Physical capability
  - Organ function
  - Cellular senescence

Potential biomarkers	Material	Age linked processes[#]	e-score	rc-score*	c-score
Lymphocytes/WBC [CDC] [PA]	blood/EDTA	Inflammation autoimmune disorders	-	202	2240
Insulin	blood/serum	Diabetic state	--	148	1143
Glucose/glucose fastened [PA]	blood/glucose monovette	Diabetic state	-	111	1175
C-reactive protein (CRP/hsCRP) [IA] [PA]	blood/plasma	Inflammation, cancer, cardiovascular disease	-	71	1146
Cholesterol	blood/plasma	Cardiovascular disease	-	67	896
Albumin [PA]	blood/plasma	Kidney and liver dysfunction	-	65	1062
IL6 [IA]	blood/plasma	Inflammation	-	58	979
Tumor necrosis factor alpha (TNFα) [IA]	blood/serum	Inflammation, cancer	--	51	751
Hemoglobin [CDC]	blood/EDTA	Anemia, other hematopoietic disorders	-	39	471
Insulin-like growth factor 1 (IGF-1)	blood/serum	Metabolic disease	--	29	263
LDL-cholesterol	blood/plasma	Cardiovascular disease	-	24	280
Triglycerides	blood/plasma	Cardiovascular disease	-	23	498
HDL-cholesterol	blood/plasma	Cardiovascular disease	-	23	349
Creatinine [PA]	blood/plasma	Kidney dysfunction	-	19	479
Monocytes	blood/EDTA	Inflammation	-	16	378
Glycated hemoglobin (Hba1c)	blood/EDTA	Diabetic state	-	13	220
Cystatin C	blood/plasma	Kidney dysfunction	-	12	142
N-terminal prohormone of brain natriuretic peptide (NT-proBNP)	blood/EDTA	Heart failure	-	10	119
Alkaline phosphatase [PA]	blood/plasma	Liver damage, bone disorder	-	9	252
Hematocrit/RBC [CDC]	blood/EDTA	Anemia	-	8	159
D-dimer	blood/citrate monovette	Hypercoagulable state	-	8	91
IL8 [IA]	blood/plasma	Inflammation	--	7	164
Plasminogen activator inhibitor-1 (PAI1)	blood/EDTA	Prothrombotic state in cancer and other acute phases	--	6	72
Bilirubin	blood/plasma	Liver dysfunction	-	5	46
Urea	blood/plasma	Renal dysfunction	-	3	137
IL15	blood/plasma	Inflammation	--	3	55
Mean corpuscular volume/MCV [CDC] [PA]	blood/EDTA	Anemia, other hematopoietic disorders	-	2	42
Mean corpuscular hemoglobin concentration/MCHC [CDC]	blood/EDTA	Anemia, other hematopoietic disorders	-	2	32
CD4/CD8 ratio	blood/EDTA	Immune deficiency, autoimmunity	--	1	103
C-peptide (preferable to insulin)	blood/serum	Diabetic state	-	1	32
IL1-β [IA]	blood/plasma	inflammation	--	1	5

* *rows are sorted by rc-score.*
[#] *frequently mentioned general or disease-linked processes.*

[IA] = inflammaging
[PA] = Phenotypic Age
[CDC] = complete blood count

[Hartmann et al., 2021]

Routine laboratory biomarkers of ageing

Potential biomarkers	Material	Methods	Age linked processes[#]	e-score	rc-score*	c-score
Telomere length (TL):			Morbidity, mortality, cell stress		191	932
Average TL	DNA	Q-PCR, TRF, TCA		--		**
TL structure	DNA	Q-FISH, Flow-FISH		---		**
Shortest TL	DNA	STELA, TeSLA		---		**
DNA damage	DNA	Various methods	Morbidity, mortality	--	174	713
Reactive oxygen species (ROS)	Tissue mitochondria	Various methods	Morbidity, cell stress, DNA/protein damage	---	168	712
Mitochondrial dysfunction	living cells, mitochondrial DNA	Various methods	Morbidity, mortality, neurodegenerative diseases	---	86	289
EVs (extracellular vesicles)	blood/plasma, liquor, cell culture supernatant	Immuno-histochemistry Western Blot, FACS	Cellular senescence, cancer	---	65	194
Autophagy	cells, cell extract	Electron microscopy immunoblotting flow cytometry	Morbidity, cancer, Parkinson's and Alzheimer's disease	---	46	207
Transforming growth factor beta (TGF-β)	blood/serum	ELISA	Inflammation, fibrosis, cellular senescence, cancer	--	45	315
Telomerase activity	cell extract, DNA	PCR-ELIDA, TRAP	Morbidity, mortality, tumor progression	---	41	169
Gut microbiome	fecal specimen	Next generation sequencing	Morbidity, mortality	--	29	101
α-Klotho	blood/plasma tissue	Immuno-histochemistry ELISA	Morbidity, mortality, renal function	--	20	107
Adiponectin	blood/plasma blood/EDTA	ELISA	Morbidity, mortality, frailty, metabolic syndrome, liver cirrhosis, diabetes type 2	-	14	217
Sirtuin 1 (SIRT1)	blood/serum	ELISA immuno-histochemistry PCR	Morbidity, mortality, inflammation, cancer	--	12	112
Growth differentiation factor 15 (GDF15)	blood/plasma	Proteomics immunoassays	Morbidity, organ damage (liver, heart, kidney)	--	12	63
Sirtuin 6 (SIRT6)	blood/serum	ELISA immuno-histochemistry PCR	Morbidity, mortality, diabetic risk, arthritis	--	4	50
Growth differentiation factor 11 (GDF11)	blood/plasma	Proteomics immunoassays	Morbidity	--	3	22
CXCL1	blood/plasma	Immunoassays, ELISA	Immune response, inflammation, cancer, Alzheimer's disease	--	0	15
Skin microbiome	skin swab	Next generation sequencing	Morbidity, mortality	--	0	4

** rows are sorted by rc-score.*
*** included in the c-score of TL.*
[#] frequently mentioned general or disease-linked processes.

[Hartmann et al., 2021]

Research laboratory biomarkers based on non-epigenetic measurements

Potential biomarkers	Material	Methods	Prediction	e-score	rc-score	c-score*
DNA methylation and aging clocks:					n.a.	2158
Horvath's clock	DNA (broad spectrum of tissues)	DNA methylation analysis	Chronological age	--	n.a.	214
Hannum's clock	DNA (blood)		Chronological age	--	n.a.	190
DNAm GrimAge	DNA (blood)		Biological age	--	n.a.	31
DNAm PhenoAge	DNA (blood)		Biological age	--	n.a.	26
Weidner clock	DNA (blood)		Chronological age	--	n.a.	8
EpiTOC	DNA (blood)		Biological age	---	n.a.	2
miRNA (microRNA)	RNA (blood/plasma PBMCs)	Next generation sequencing microarrays	Morbidity, mortality	---	198	635
Non-coding RNA expression profiles	RNA	RNA sequencing	Chronological age	---	167	602
exRNA (extracellular RNA)	blood/plasma	Next generation sequencing	Morbidity, mortality	---	25	119
Histone modifications:					36	73
H4K20 methylation		DNA methylation analysis mass spectrometry, HPLC, ChIP Immunohisto-chemistry	Cell stress	---	n.a.	n.a.
H4K16 acetylation				---	n.a.	n.a.
H3K4 methylation	protein extract			---	n.a.	n.a.
H3K9 methylation	from tissue DNA			---	n.a.	n.a.
H3K27 methylation				---	n.a.	n.a.
Chromatin remodeling	DNA	Chromatin remodeling assays	Chronological age	---	13	26

** rows are sorted by c-score.*
n.a.: not assigned due to high variation of terminology in literature.

[Hartmann et al., 2021]

Research laboratory biomarkers based on epigenetic measurements

Potential biomarkers	Method	Age linked processes[#]	Domain	e-score	rc-score*	c-score
Physical capability						
Grip strength	Physical exam	Mortality, morbidity	Strength	--	11	229
Walking speed	Physical exam	Mortality, morbidity	Locomotor function	--	3	106
Standing balance	Physical exam	Mortality, morbidity	Balance	--	1	26
Timed up and go test	Physical exam	Mortality, morbidity	Locomotor function	--	0	11
Organ function						
Atherosclerotic lesions	IMT, ultrasound	Mortality, CAD	Cardiovascular system	--	158	680
Muscle mass	MRI	Mortality, cardiovascular risk	Body composition	--	81	495
Systolic blood pressure	Auscultatory method	Mortality, cardiovascular risk	Cardiovascular system	--	65	844
Cognitive function	Various	Mortality, morbidity	Brain function	---	56	581
Body mass index	Calculated	Mortality CAD	Body composition	--	24	1280
Bone density	Bone density test	Mortality, morbidity	Body composition	--	17	84
Lung function	Spirometry	Mortality, morbidity	Respiratory system	--	16	84
Waist circumference	Tape measure	Mortality, cardiovascular risk	Body composition	--	3	202
General well being						
Health assessments	Questionnaire	Mortality, morbidity	General	--	n.a.	n.a.

** rows are sorted by c-score.*
[#] frequently mentioned general or disease-linked.
n.a.: not assigned due to high variation of terminology in literature.

[Hartmann et al., 2021]

Non-blood physical capability and organ function biomarkers

Potential biomarker		Material and Methods	e-score	rc-score*	c-score
SASP				442	2646
	Cytokines (e.g., IL-6, IL-7, IL-15)	ELISA from Serum or EDTA plasma samples proteomics	--° --° --° --	n.a.	n.a.
	Chemokines (e.g., IL-8, CCL3, CCL4)			n.a.	n.a.
	Growth factors (e.g., GDF-15, activin A)			n.a.	n.a.
Cell cycle arrest	p53	qPCR from blood samples/staining of cultured cells/flow cytometry NGS/microarray	-- -- --	66	561
	p16			27	422
	p21			21	435
SA-βGal		Microscopy/flow cytometry	---	9	359
SAHF	Histone fragments (H3K9Me2, HP1γ)	DAPI/heterochromatin staining	---°	3	19
Lamin B1		Immunohistochemistr Western Blot	---	0	12
Cell morphology (e.g., progerin)	Cell shape	Microscopy of cultured cells	---	n.a.	n.a.

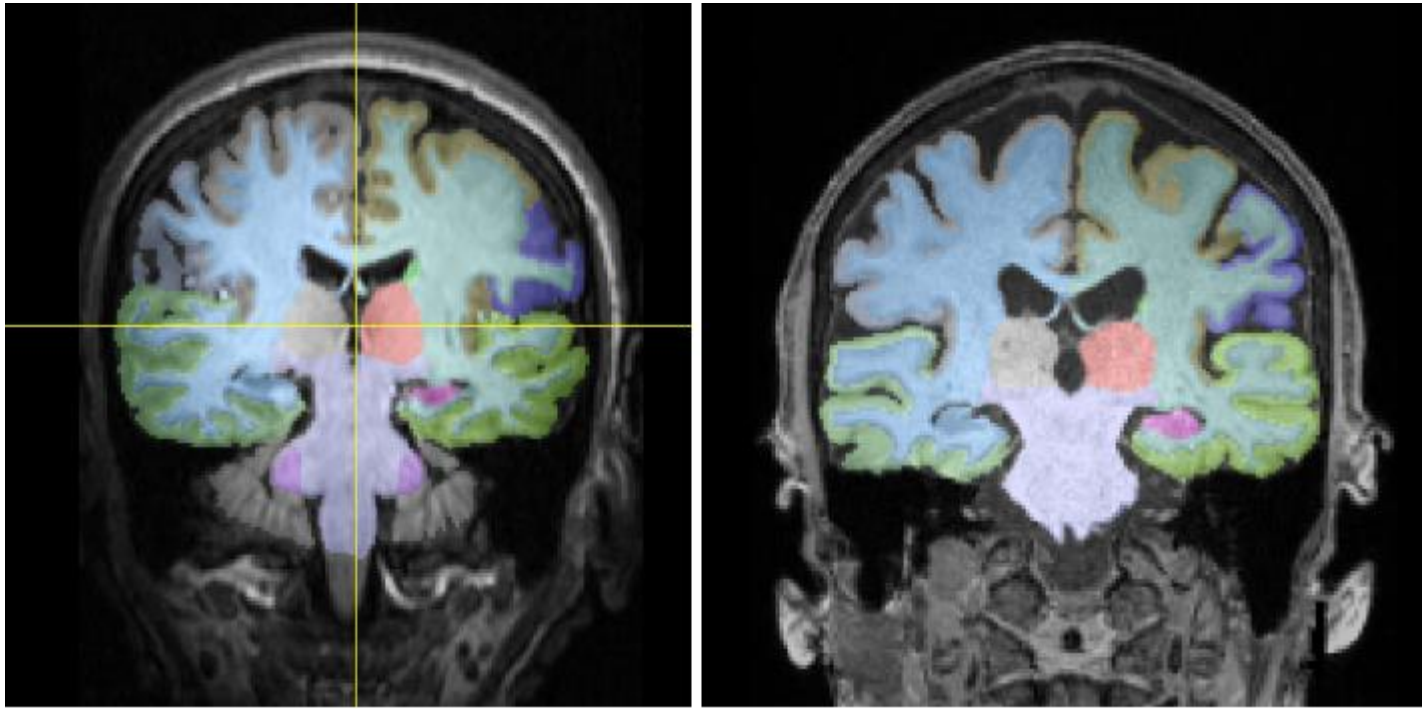
** rows are sorted by rc-score.*
*° on average (detailed in **Supplementary Table 1, 5**).*
n.a.: not assigned due to high variation of terminology in literature.

[Hartmann et al., 2021]

Biomarkers associated with cellular senescence

Brain Age

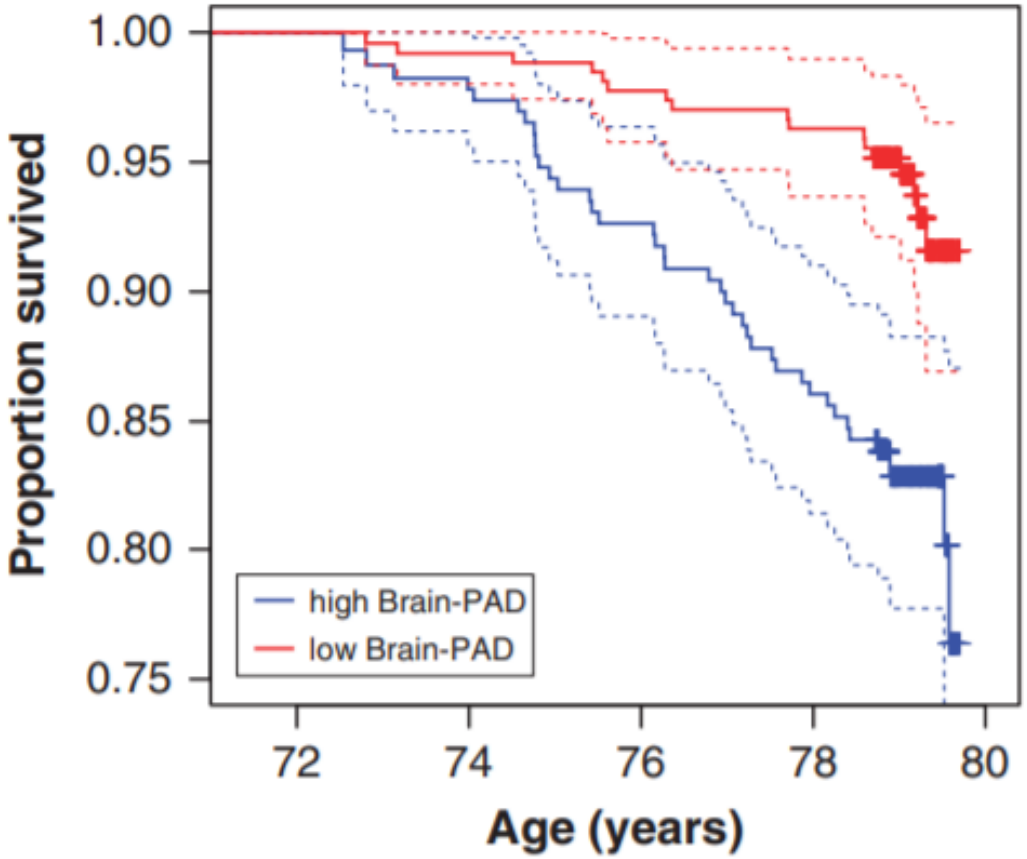
- Biological age usually estimated with information derived from brain MRI data
- Sums up the progression of ageing processes in the brain
  - Reflects relatively advanced or delayed brain maturation, while all individuals' brains undergo the general progression such as the thinning of grey matter



[https://www.brainkey.ai/blog/brain-age-how-we-calculate-it-and-what-it-means]

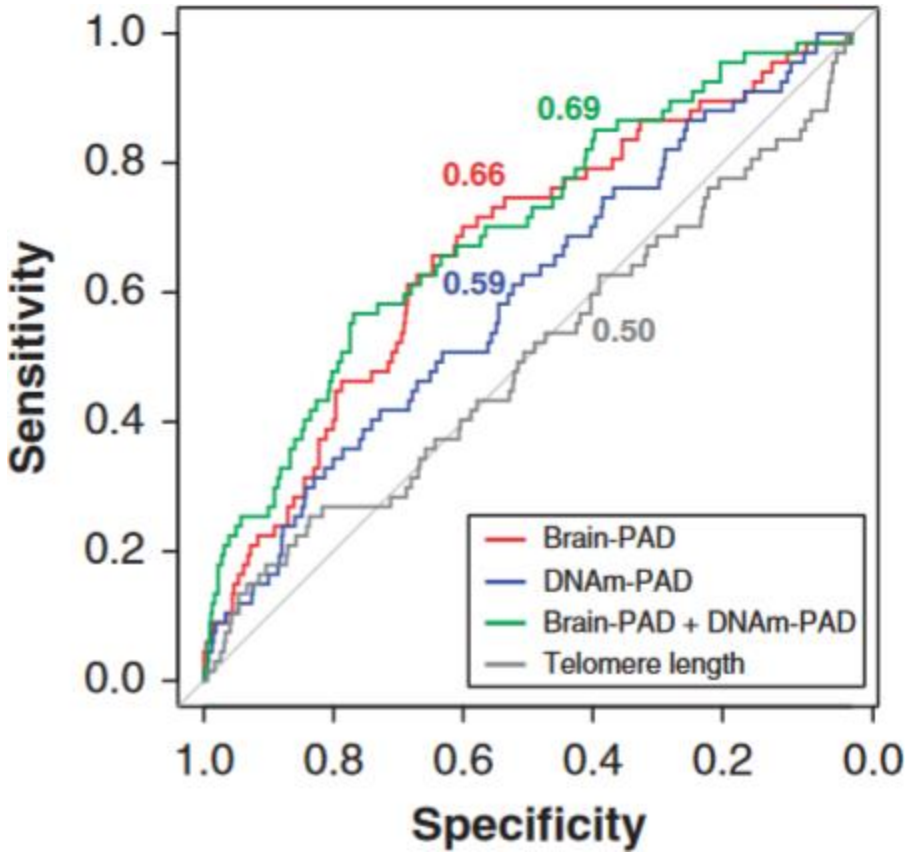
Representative brain images for young (22 years) and old (83 years) individuals

- Brain age gap (BAG, also called brain-predicted age difference, delta, etc.)
  - Difference between brain age and chronological age
  - Indicates whether an individual's brain appears to have aged more or less than the population average for their actual chronological age
    - Brain age gap > 0: accelerated ageing
      - Related to functional impairment [Liem et al., 2017], brain diseases such as Alzheimer's disease [Yin et al., 2023], Parkinson's disease [Eickhoff et al., 2021], schizophrenia [Nenadic et al., 2017], stroke [Egorova et al., 2019], epilepsy [Sone et al., 2021], and diabetes mellitus [Franke et al., 2013], and mortality [Cole et al., 2017]
    - Brain age gap < 0: resilience to ageing



[Cole et al., 2017]

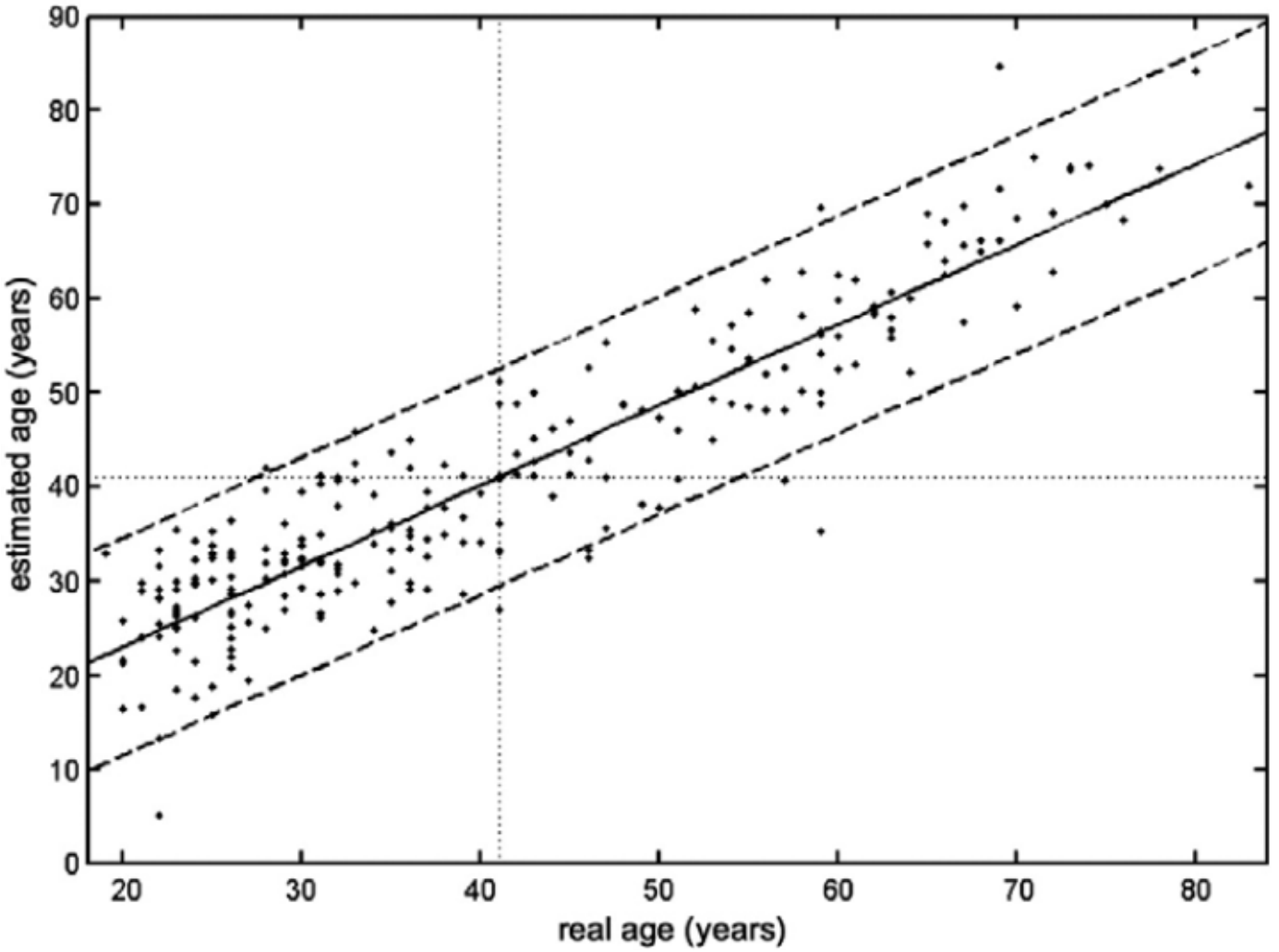
Larger proportion of survival in individuals with lower BAG than those with higher BAG



[Cole et al., 2017]

Receiver operator characteristic (ROC) curves for four survival models

- Estimation of brain age
  - Modelling chronological age as a function of various information from brain MRI data by using a supervised-learning algorithm
  - Input data
    - Brain MRI data of one or more modalities, for example, T1-weighted structural MRI data for individuals
  - Supervised-learning algorithm
    - Regression
    - Conventional machine learning
    - Deep learning



[Franke et al., 2010]

Estimation of brain age

Human Connectome Project (HCP)

- Launched in 2009 as a Blueprint Grand Challenge by the National Institutes of Health in the US
- Maps the neural pathways that underlie human brain function
  - To identify functionally distinct subdivisions of the human brain
  - To understand the relationships among the subdivisions
- Acquires and shares data about the structural and functional connectivity of the human brain

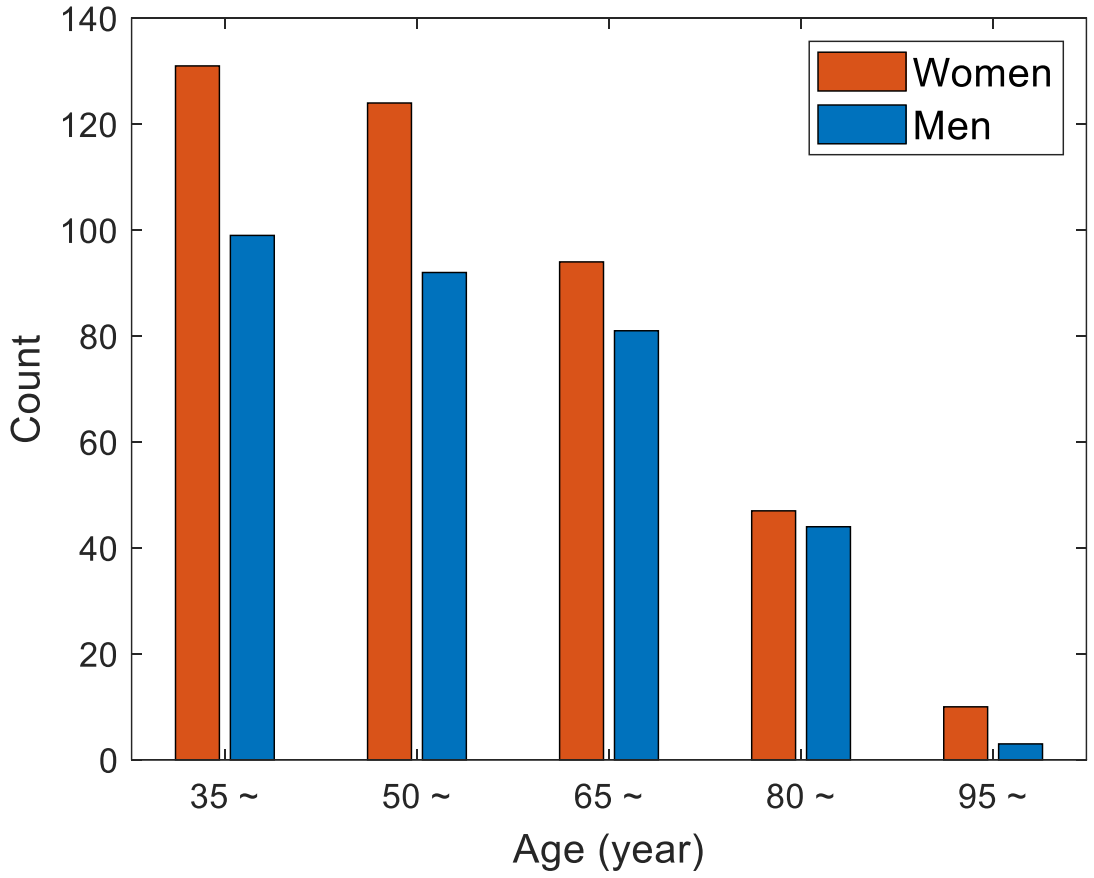
- HCP original study
  - HCP Young Adult (HCP-YA)
    - Age: 22-35 years
  - 1200 subjects data release (S1200)
    - Released on 1 March 2017
    - 1,206 healthy young adult participants collected in 2012-2015
    - Imaging and behavioural data
      - 3T MRI data for 1,113 participants
      - 3T HCP protocol (MRI and behaviour) retest data for 46 participants
      - Multimodal 7T MRI data for 184 participants
  - S1200 extensively processed functional MRI data
    - Released on 21 July 2017

- HCP lifespan studies
  - Lifespan developing human connectome project
    - Age: 20-44 weeks post-conception
  - Lifespan baby connectome project
    - Age: 0-5 years
  - HCP Development (HCP-D)
    - Age: 5-21 years
  - HCP Aging (HCP-A)
    - Age: 36-100 years

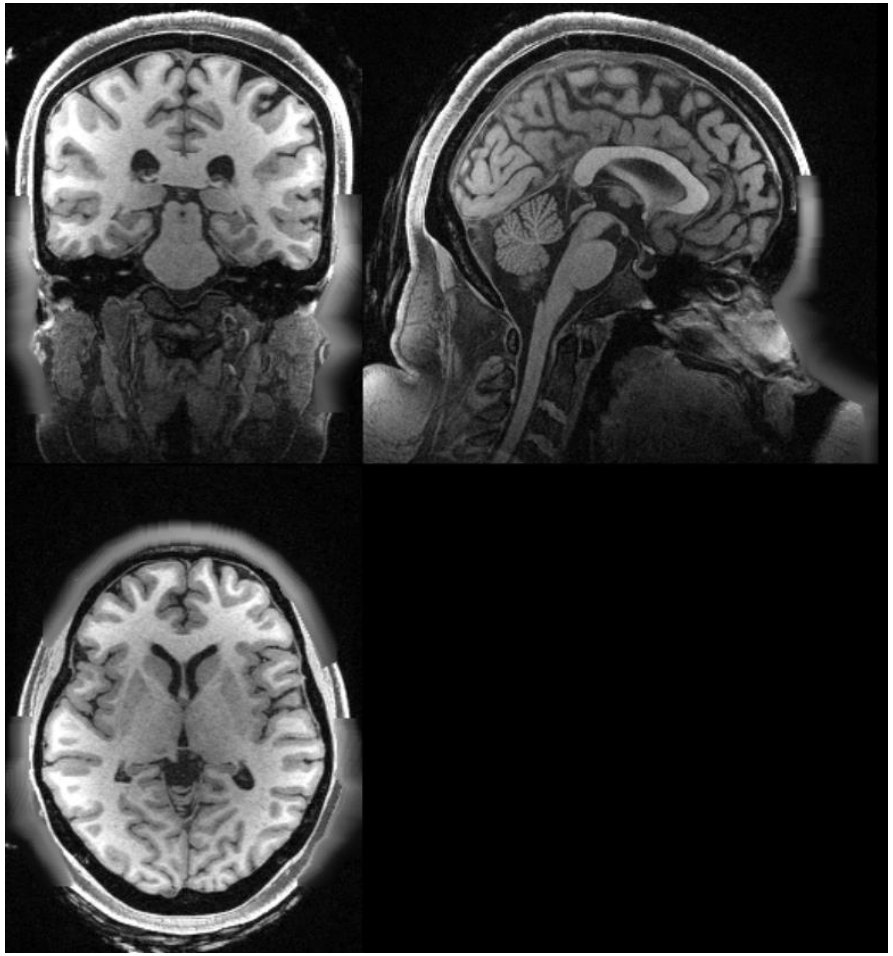
- Lifespan 2.0 release
  - Released on 26 February 2021
  - HCP-A & HCP-D
    - HCP-A: 725 healthy participants
    - HCP-D: 652 healthy participants
    - 22+ TB of data per project
  - Imaging and behavioural data at cross-sectional visit 1 (V1)
    - Unprocessed V1 imaging data for all included modalities
    - Preprocessed structural and functional imaging data
    - Non-imaging demographic and behavioral assessment data

Dataset

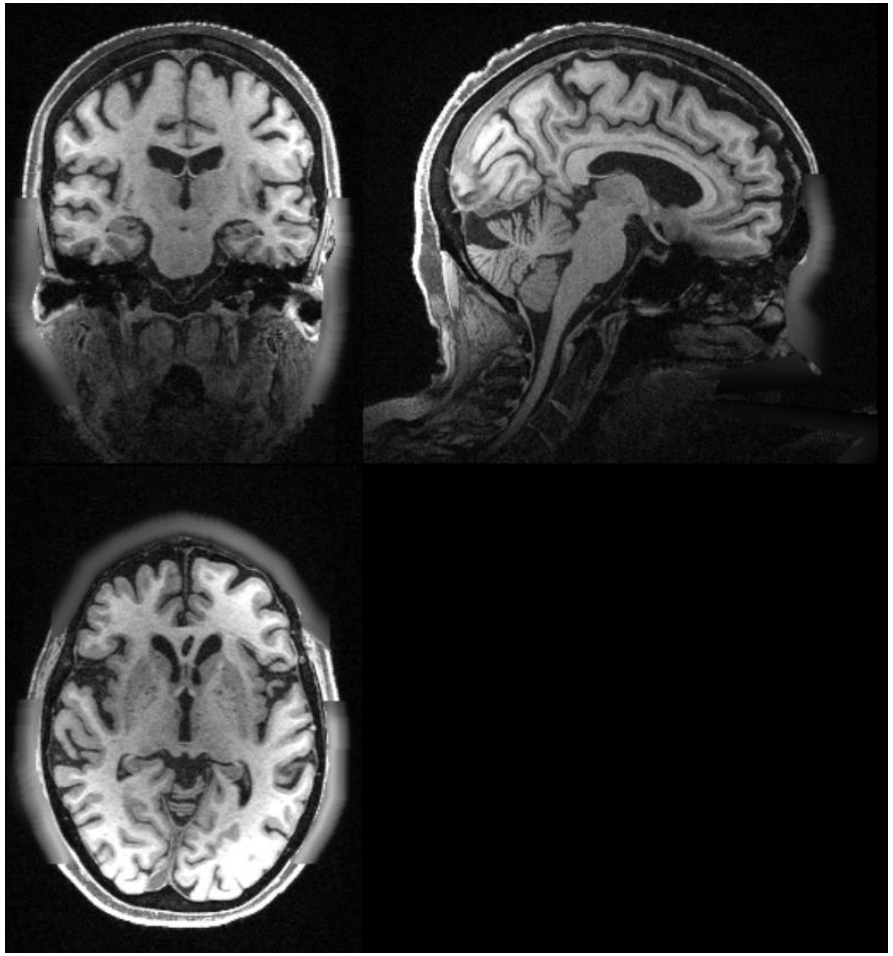
- Data from HCP-A (*n* = 725)
  - T1-weighted MRI scans
  - Demographic information including chronological age and sex



Distribution of age and sex for 725 individuals



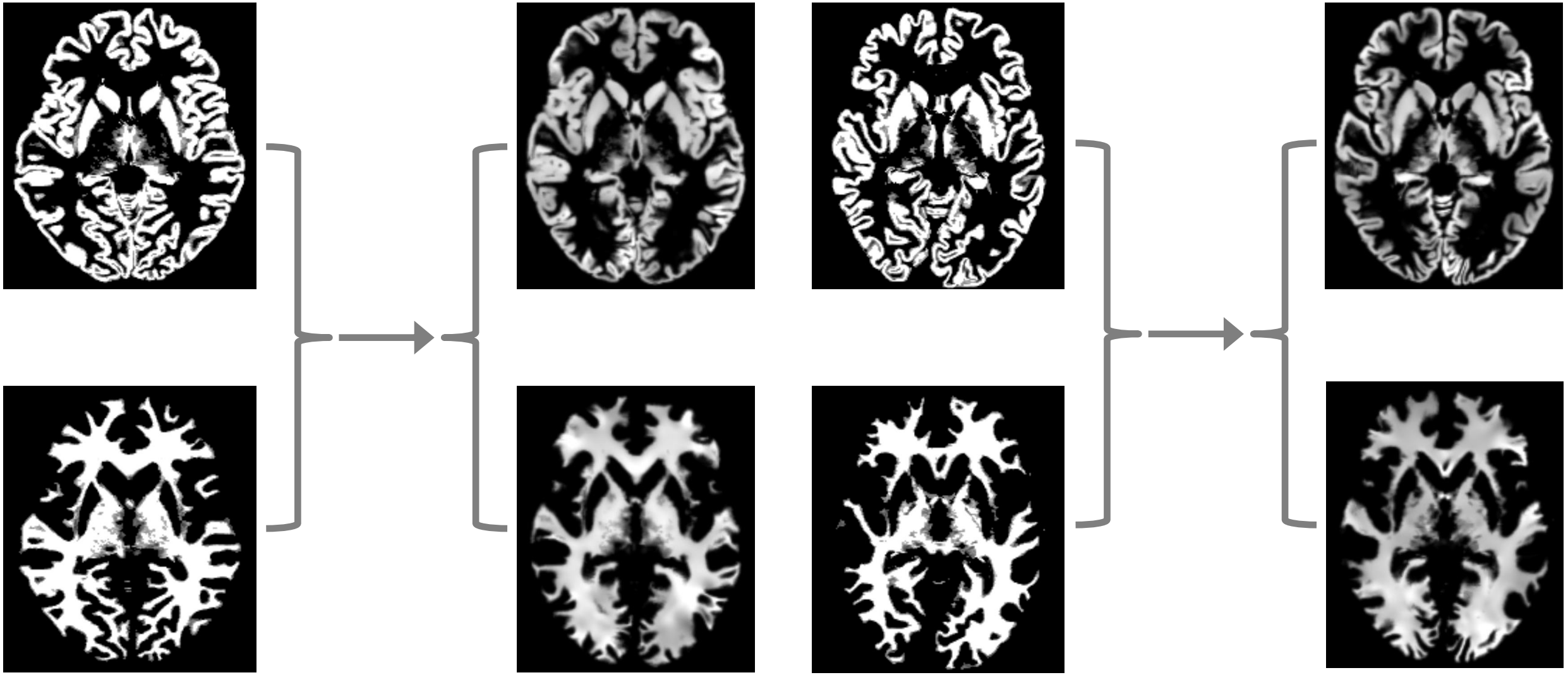
36-year-old woman



100-year-old woman

T1-weighted MRI scan

- Preprocessing
  - Correction for intensity non-uniformity (bias field)
  - Segmentation into grey matter, white matter, and cerebrospinal fluid
  - Normalisation into the MNI standard brain space

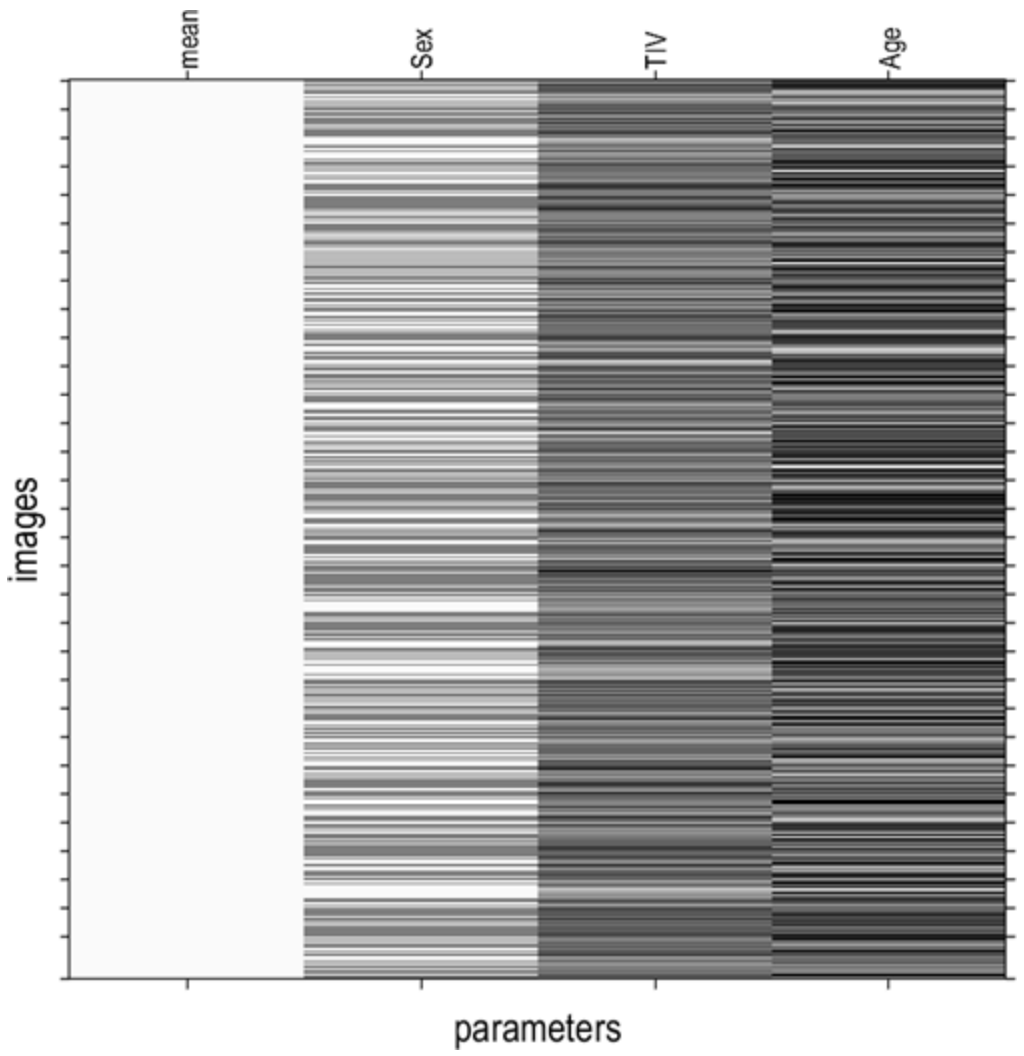


36-year-old woman

100-year-old woman

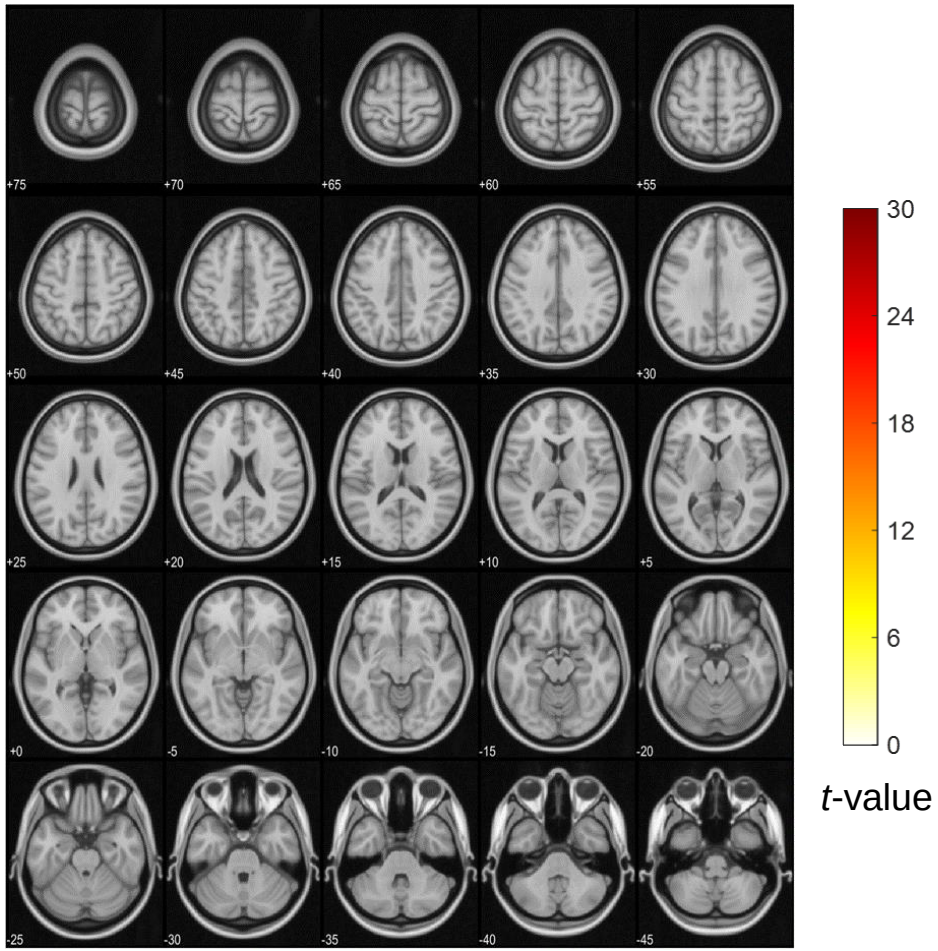
Segmentation and normalisation

- Statistical inferences on processed maps: age-related changes in brain structure
  - Grey matter volume ~ sex + total intracranial volume (TIV) + age
  - White matter volume ~ sex + TIV + age

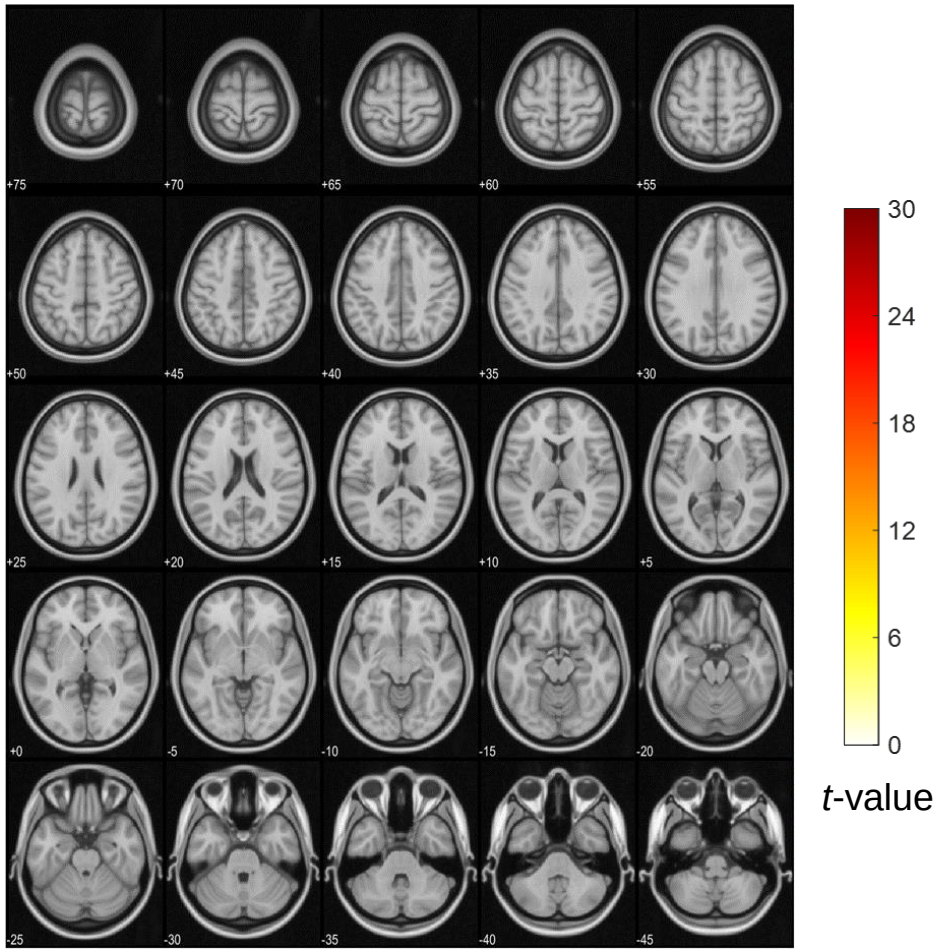


Design matrix for age correlation

Positive correlation with age



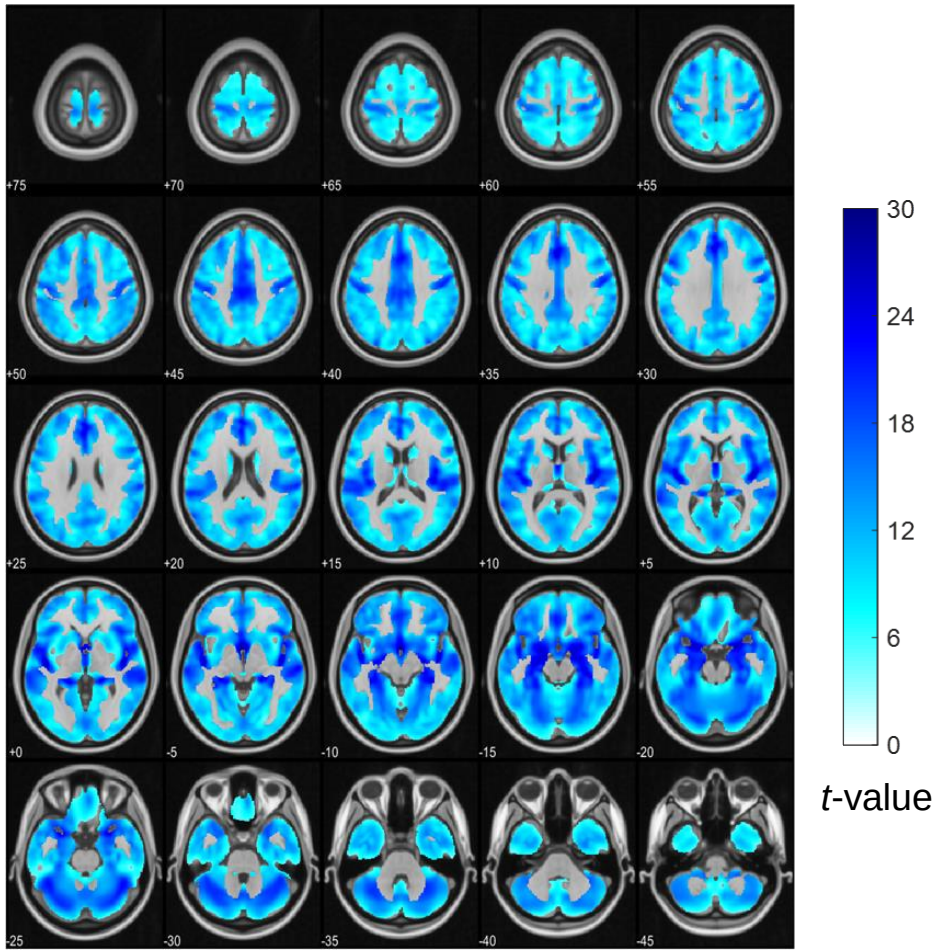
Grey matter volume



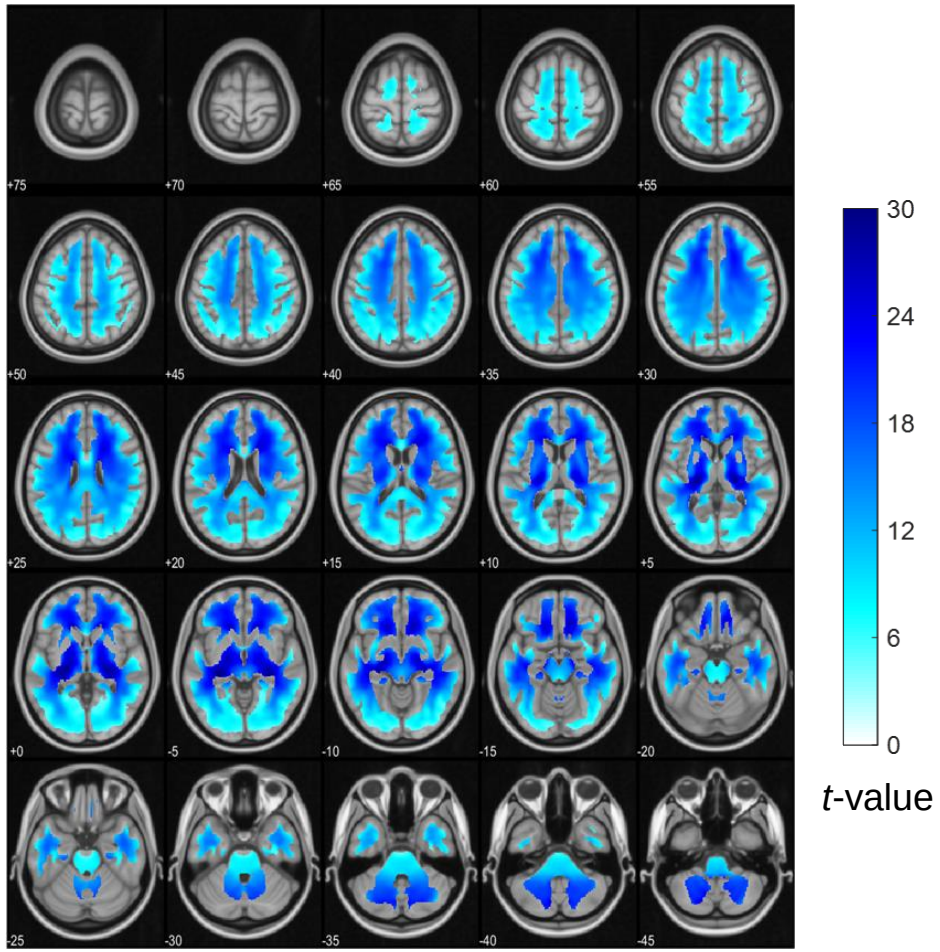
White matter volume

Thresholded at FDR corrected $p = 0.05$ at the cluster level and FWE corrected $p = 0.05$ at the voxel level

Negative correlation with age



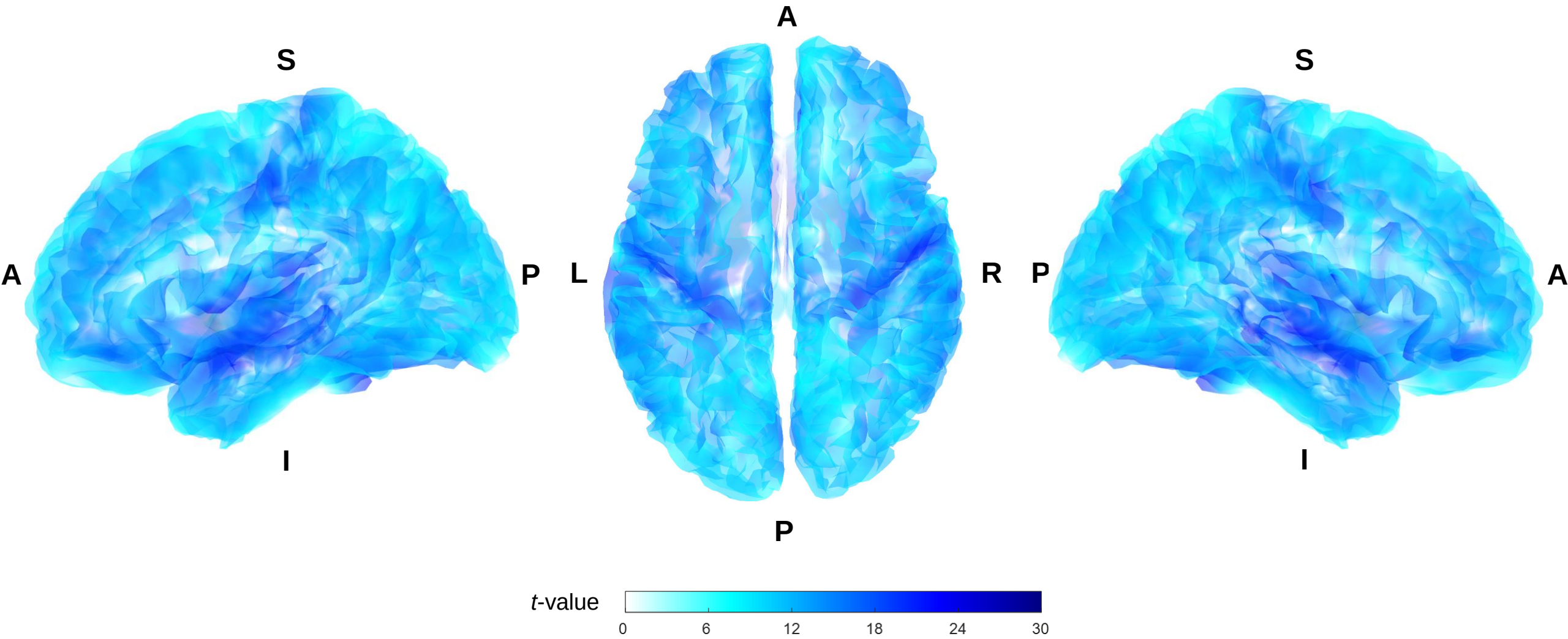
Grey matter volume



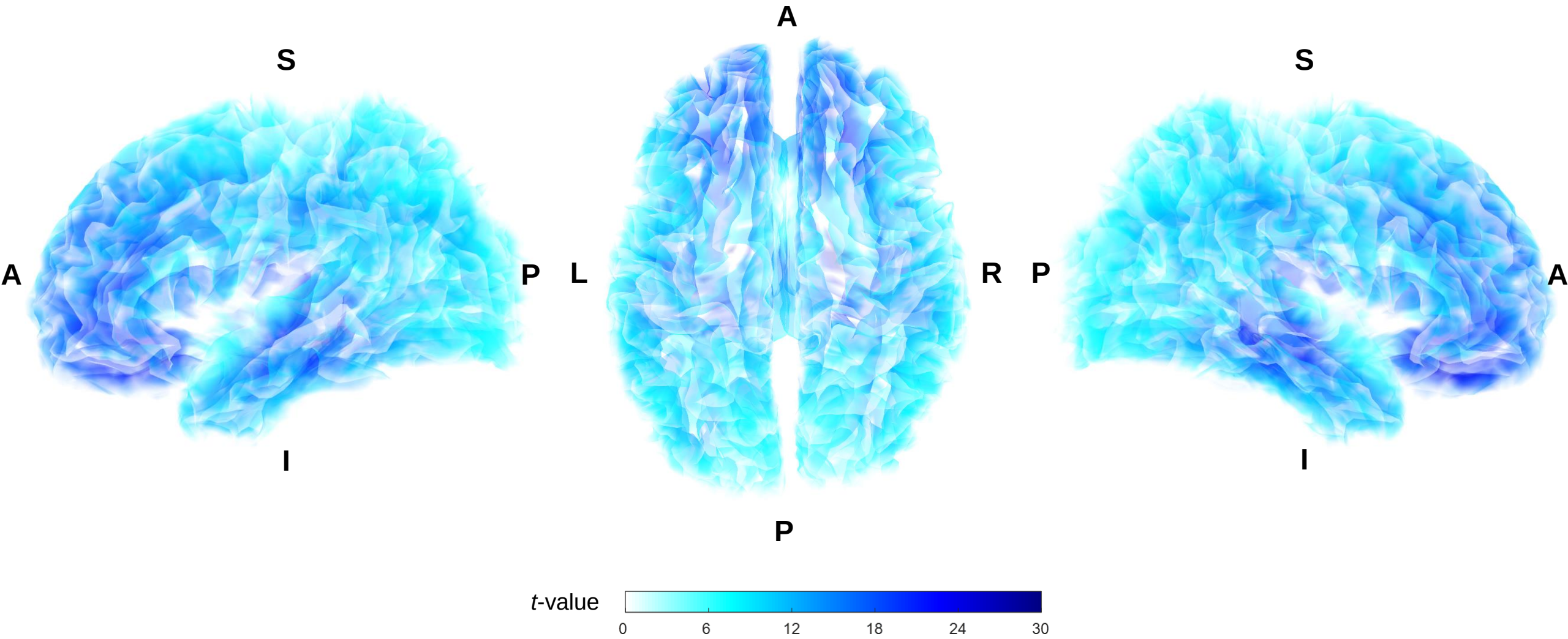
White matter volume

Thresholded at FDR corrected $p = 0.05$ at the cluster level and FWE corrected $p = 0.05$ at the voxel level

Negative correlation between grey matter volume and age

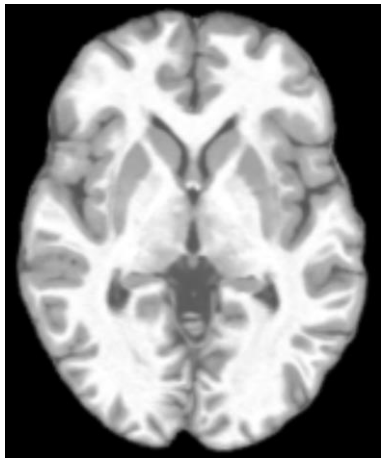


Negative correlation between white matter volume and age

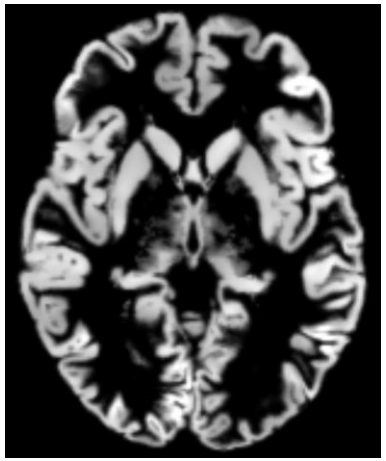


Brain Age Estimation

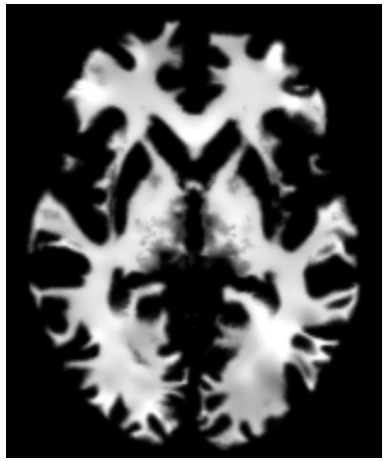
- Input data
  - Brain image after skull-stripping, intensity non-uniformity correction, and normalisation
  - Grey matter probability (partial volume fraction) image after segmentation and normalisation
  - White matter probability (partial volume fraction) image after segmentation and normalisation



Brain image



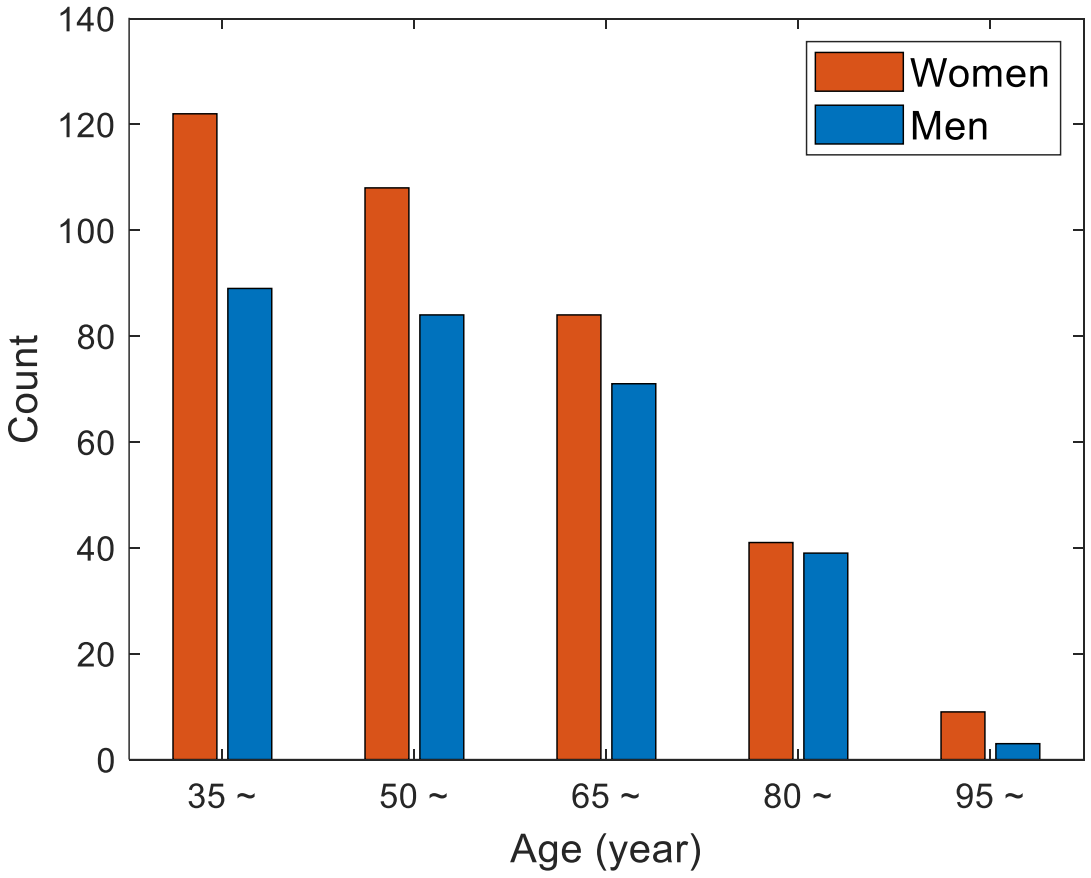
Grey matter
probability image



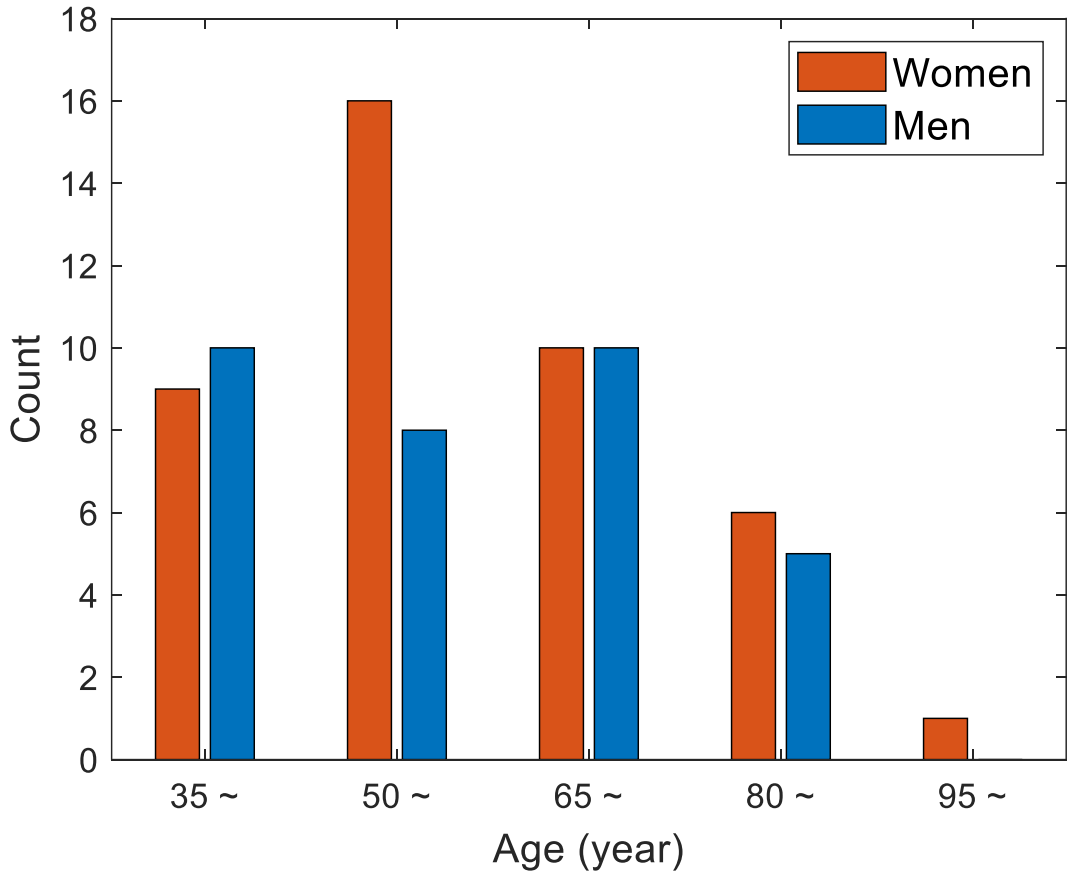
White matter
probability image

Input data for brain age estimation

- Training and test datasets
  - Training dataset: $n = 650$
    - Brain image: Brain/train/001-650.nii.gz
    - Grey matter probability image: GM/train/001-650.nii.gz
    - White matter probability image: WM/train/001-650.nii.gz
  - Test dataset: $n = 75$
    - Brain image: Brain/test/001-075.nii.gz
    - Grey matter probability image: GM/test/001-075.nii.gz
    - White matter probability image: WM/test/001-075.nii.gz



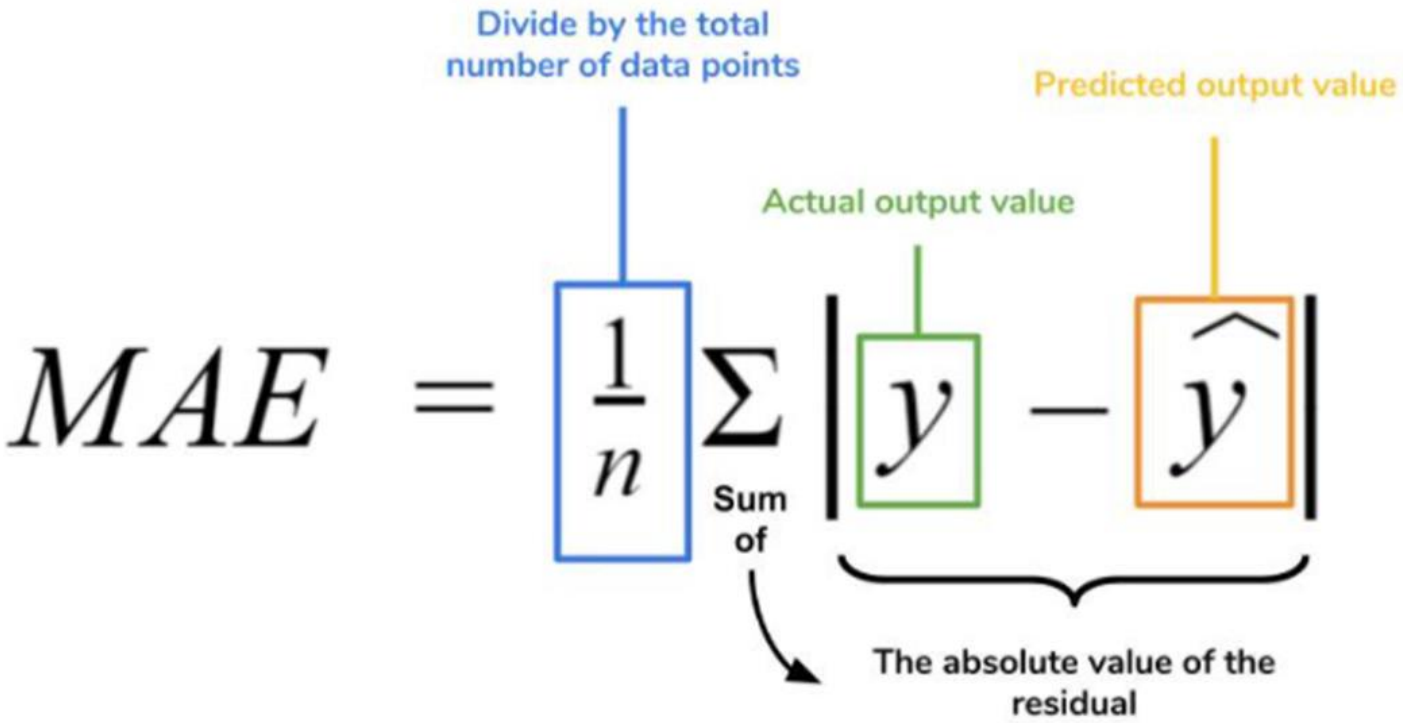
Training dataset



Test dataset

Distribution of age and sex for training and test datasets

- Brain age estimation performance
  - Mean absolute error (MAE)
    - (sum of absolute errors)/(sample size)



[https://medium.com/@polanitzer/the-minimum-mean-absolute-error-mae-challenge-928dc081f031]