

Enhancing brain age estimation with structural MRI and synthesized cerebral blood volume maps

Jordan Jomsky, Zongyu Li, Kay C Igwe, Yiren Zhang, Max Lashley, Tal Nuriel, Andrew Laine, Scott Small, Jia Guo, for the Frontotemporal Lobar Degeneration Neuroimaging Initiative and for the Alzheimer's Disease Neuroimaging Initiative

Enhancing brain age estimation with structural MRI and synthesized cerebral blood volume maps

Short Title: Brain Age Estimation with DeepCBV

Jordan Jomsky¹, Zongyu Li¹, Kay C. Igwe¹, Yiren Zhang¹, Max Lashley², Tal Nuriel^{3,4}, Andrew Laine¹, Scott Small^{3,5,6}, Jia Guo^{2,6}, for the Frontotemporal Lobar Degeneration Neuroimaging Initiative and for the Alzheimer’s Disease Neuroimaging Initiative*

Affiliations

- ¹Department of Biomedical Engineering, Columbia University, New York, NY, United States
- ²Zuckerman Institute, Columbia University, New York, NY, United States
- ³Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, United States
- ⁴Department of Pathology and Cell Biology, Columbia University, New York, NY, United States
- ⁵Department of Neurology, Columbia University, New York, NY, United States
- ⁶Department of Psychiatry, Columbia University, New York, NY, United States

*Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Corresponding author:

Jia Guo, PhD.
Zuckerman Institute
Columbia University in the City of New York
3227 Broadway, New York, NY 10027
United States of America
E-mail: jg3400@columbia.edu

Abstract

Brain age gap estimation (BrainAGE) is a promising imaging-derived biomarker of neurobiological aging and disease risk, yet current approaches rely predominantly on T1-weighted structural MRI, overlooking functional vascular changes that may precede tissue damage and cognitive decline. Deep learning-derived cerebral blood volume (DeepCBV) maps, synthesized from non-contrast MRI, offer a scalable alternative to contrast-enhanced perfusion imaging by capturing vascular information relevant to early neurodegeneration. We developed a multimodal BrainAGE framework that combines predictions from two separate 3D convolutional neural networks: one trained only on structural MRI scans and another trained only on DeepCBV maps generated by a pre-trained 3D patch-based deep learning model. Each model was trained and validated on 2,851 scans (1,507 females) from 13 open-source datasets and was evaluated for concordance with mild cognitive impairment (MCI) and Alzheimer's disease (AD) using 1,233 subjects. The combined model achieved the most accurate brain age gap for cognitively normal (CN) controls, with a mean absolute error (MAE) of 3.95 years ($R^2 = 0.943$), outperforming models trained on MRI (MAE = 4.10) or DeepCBV alone (MAE = 4.49). Saliency maps revealed complementary modality contributions: MRI emphasized white matter and cortical atrophy, while DeepCBV highlighted vascular-rich and periventricular regions implicated in hypoperfusion and early cerebrovascular dysfunction, consistent with known patterns of normal aging. Next, we observed that BrainAGE increased stepwise across diagnostic strata (CN < MCI < AD) and correlated with cognitive impairment (Clinical Dementia Rating Sum of Boxes $\rho = 0.403$; Mini-Mental State Examination $\rho = -0.310$). DeepCBV-based BrainAGE showed particularly strong separation between stable vs. progressive MCI (Mann-Whitney $U = 2.177 \times 10^4$, $p = 4.43 \times 10^{-8}$), suggesting sensitivity to prodromal vascular changes that precede overt atrophy. Integrating structural MRI with deep learning-derived vascular measures substantially enhances BrainAGE estimation and improves sensitivity to MCI and AD progression, supporting its potential role in risk stratification, early detection, and monitoring of therapeutic response. By enabling functional-like assessment from routine MRI, this approach lowers barriers to multimodal evaluation and provides a clinically actionable biomarker for large-scale aging and dementia studies.

Keywords: magnetic resonance imaging, brain age, deep learning, cerebral blood volume

Introduction

Aging is a universal heterogeneous biological process that contributes to an increased risk of developing neurodegenerative diseases, such as Alzheimer’s disease (AD).^{1–3} This inherent variability in biological aging, independent of chronological age, may be an indicator of pathological aging, reflecting an acceleration of biological mechanisms, such as inflammation and aberrant protein accumulation, also present in healthy aging.⁴ In the brain, pathogenic mechanisms of biological aging manifest as a collection of structural and functional changes, including cortical and subcortical atrophy,⁵ white matter hyperintensity (WMH) accumulation,^{6–8} widening perivascular spaces,^{9,10} brain perfusion,^{11,12} and functional connectivity dynamics,¹³ all of which are detectable through magnetic resonance imaging (MRI). These non-invasive neuroimaging tools are instrumental in quantifying biological age, a measure of the deviation from chronological age that reflects pathological changes. This deviation from a chronological age is known as brain age gap estimation (BrainAGE), serving as a crucial biomarker for aberrant aging.¹⁴

Brain age gap estimation (BrainAGE) has emerged as a promising neuroimaging-based biomarker for estimating a divergence from normative brain aging, potentially revealing underlying pathophysiological changes associated with neurodegeneration. T1-weighted (T1w) MRI has become a primary non-invasive imaging modality for assessing BrainAGE due to its utility in capturing age-related anatomical changes, its wide availability across large open-source neuroimaging databases, and its clinical utility. Various studies have demonstrated that machine and deep learning models trained solely on T1w images can accurately predict chronological age from brain scans of healthy individuals.^{15–20} In fact, Feng et al.¹⁶ were able to accurately predict chronological age (mean absolute error = 4.06, $R^2 = 0.941$) using a VGG-based deep learning model and the T1w brain image as input. While numerous recent BrainAGE models have reported lower mean absolute error using a variety of architectures and large-scale datasets, including transformer-based and minimally preprocessed approaches, direct comparisons across studies are complicated by differences in cohort composition, preprocessing pipelines, and evaluation strategies^{19–25}. We reference Feng et al.¹⁶ primarily because of the use of a similar VGG-based architecture and similar training data, which permits a more direct comparison. Importantly, the primary contribution of the present work is not state-of-the-art age prediction accuracy per se, but the demonstration that a learned vascular representation derived from non-contrast structural MRI provides complementary information that improves BrainAGE estimation and clinical discrimination.

Although T1w-based BrainAGE models are effective in predicting chronological age, this approach may fail to account for underlying biological processes that reflect brain function. Further, recent studies have demonstrated that vascular changes and cerebral blood flow, as measured by functional data modalities, provide earlier and more accurate biomarkers of AD onset.^{26,27} Combining structural and functional imaging modalities to determine BrainAGE may offer a more comprehensive clinical evaluation for aging and age-related neurodegenerative diseases and provide insight into the divergence of brain age prediction from chronological age.

Cerebral blood volume functional MRI (CBV-fMRI), which utilizes paired T1-weighted imaging acquired with and without gadolinium-based contrast agents (GBCAs), produces functional CBV maps that measure regional metabolism, providing a quantitative correlate of metabolic dysfunction. Further, the use of GBCAs provides a measure of blood-brain barrier (BBB) integrity through observation of contrast agent kinetics, all of which serve as critical measures for elucidating pathological aging. However, the use of GBCAs is invasive, requiring an injection, and is not feasible for large longitudinal studies. Additionally, CBV-fMRI is not as accessible in large-scale clinical datasets as T1w MRI.

We introduce Deep learning-derived Cerebral Blood Volume (DeepCBV) maps, a non-contrast surrogate for CBV-fMRI generated directly from standard T1-weighted MRI. This approach allows for the generation of high-resolution CBV maps without the use of contrast agents, making it suitable for large-scale and retrospective neuroimaging studies. We hypothesized that combining T1w MRI and DeepCBV maps into a single model may serve as a more accurate estimator of brain age and neurodegenerative risk, capturing both structural integrity and functional vascular changes that are critical in aging and Alzheimer's disease progression.

Our study had two main objectives for assessing the efficacy of our combined T1w and DeepCBV approach. First, we determined whether training a model on both a structural T1w image and a deep learning-derived CBV map could reveal a more robust BrainAGE metric than either modality alone. Second, we examined whether our combined model could identify established AD-specific relationships, which would demonstrate our model's clinical relevance.

Materials and methods

This study did not require ethical approval as it relied on publicly available anonymized MRI scans that were originally collected as part of studies already approved by the requisite standards.

Data Sources

In this study, we leveraged neuroimaging and clinical data from 13 publicly available datasets: Alzheimer's Disease Neuroimaging Initiative (ADNI), Brain Genomics Superstruct Project (BGSP),²⁸ Neuroimaging in Frontotemporal Dementia (NIFD), Southwest University Longitudinal Imaging Multimodal (SLIM) study,²⁹ Dallas Lifespan Brain Study (DLBS),³⁰ Southwest University Adult Lifespan Dataset (SALD),³¹ Australian Imaging, Biomarkers & Lifestyle Study of Aging (AIBL),³² Consortium for Reliability and Reproducibility (CoRR),³³ SchizConnect,³⁴ Information eXtraction from Images (IXI),³⁵ Open Access Series of Imaging Studies (OAS1 and OAS2),^{36,37} and the Parkinson's Progression Markers Initiative (PPMI).³⁸ Data used in the preparation of this article were obtained from the ADNI database (adni.loni.usc.edu). The ADNI database was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test MRI, positron emission

1
2
3 tomography (PET), other biological markers, and clinical and neuropsychological assessment can
4 be combined to measure the progression of mild cognitive impairment (MCI) and early AD.
5
6

7 The entire dataset included 2,851 T1w MRI images (1,507 females). All images were registered
8 to the MNI-152 template, widely used standard brain reference derived from averaging 152
9 structural MRI scans from healthy adults, using an affine transformation.³⁹ A diverse set of data
10 sources in this study were used to enhance the transferability and generalizability of our approach
11 to different datasets. The distribution of patient ages across the datasets, each study, and the data
12 split used to train and validate the model can be seen in Figure 1, where each plot highlights the
13 variability in age ranges across each study-specific dataset used to train our model.
14
15
16

17
18 To validate our model’s AD risk assessment, we utilized data from the ADNI, which includes
19 individuals clinically categorized as cognitively normal (CN), MCI, and AD. ADNI subjects
20 included in clinical analyses were fully excluded from BrainAGE model training, validation, and
21 regression fitting. All ADNI disease-related results were obtained by applying pretrained models
22 to ADNI data in inference mode only, ensuring strict separation between model development and
23 clinical validation. Within the MCI cohort, participants were further stratified into stable MCI
24 (SMCI) and progressive MCI (PMCI) based on longitudinal diagnostic trajectories. Cognitive and
25 clinical severity were quantified using three well-established measures: the Clinical Dementia
26 Rating Sum of Boxes (CDRSB), which captures functional decline across multiple domains and
27 is sensitive to the widespread consequences of brain atrophy⁴⁰; the Mini-Mental State Examination
28 (MMSE), a cognitive measure used for early detection of dementia⁴¹; and the Alzheimer’s Disease
29 Assessment Scale-Cognitive Subscale (ADAS13), a sensitive marker of cognitive decline across
30 the spectrum from MCI to AD.⁴²
31
32
33
34
35

36 DeepCBV Model

37
38 Conventional CBV-sensitive mapping in steady-state, gadolinium-based protocols relies on
39 acquiring paired T1-weighted scans before and after contrast and then deriving a voxel-wise
40 enhancement map that reflects contrast-related signal increase.⁴³ In practice, this enhancement
41 proxy is obtained by intensity-matching the paired scans and computing the difference between
42 the post-gadolinium and the pre-gadolinium image, which encodes the spatial pattern of
43 gadolinium-related uptake and therefore contains vascular information not present in the non-
44 contrast structural scan alone.⁴³
45
46
47

48
49 Because most large-scale, open neuroimaging datasets used for brain age modeling do not include
50 contrast-enhanced acquisitions, we generated DeepCBV maps using DeepContrast, a deep-
51 learning pipeline developed to substitute gadolinium administration at inference time by predicting
52 a voxel-wise contrast-uptake map from a single non-contrast T1w scan.^{43–46} In this manuscript,
53 DeepCBV refers to the DeepContrast-predicted voxel-wise relative cerebral blood volume (rCBV)
54
55
56

map computed from non-contrast T1w alone and used as an additional CBV-sensitive vascular feature for downstream BrainAGE modeling.

Training dataset and preprocessing for DeepContrast

As described in Liu et al.,⁴³ all DeepContrast model versions were trained on an in-house steady-state CBV-fMRI human dataset consisting of paired T1w and T1w-CE volumes acquired in the same session, with a train/validation/test split of 326/93/179 participants, respectively; receiver gain was held constant and offset set to zero so that voxel intensities remained physically comparable within each pair. Each T1w volume was intensity-compressed to the range [0, 1], and the corresponding T1w-CE volume was scaled by the same factor prior to computing the rCBV target by subtraction.⁴³ Loss evaluation was restricted to brain tissue using masks generated with FMRIB Software Library Brain Extraction Tool.⁴⁷

Supervision target

rCBV was defined as the voxelwise gadolinium-induced signal enhancement normalized by the enhancement measured from pure blood. The scalar-matched pre-contrast T1-weighted (T1w) and contrast-enhanced T1-weighted (T1w-CE) images were first spatially aligned. A gadolinium uptake map was then generated by subtracting the T1w image from the T1w-CE image.⁴³ Because the two images were scalar-matched, this signal difference was assumed to be proportional to local gadolinium uptake.

To obtain rCBV, the gadolinium uptake map was normalized by the mean signal enhancement measured within a pure blood region of interest, such as a large vessel or venous sinus. The resulting dimensionless map represents blood-normalized gadolinium enhancement and was interpreted as a relative CBV measure.^{43,44} Higher rCBV values indicate greater blood-volume-related gadolinium uptake relative to the blood reference signal.

DeepContrast Models

To ensure methodological completeness, we summarize the four DeepContrast model generations that underpin our DeepCBV generation pipeline. We provide a brief architectural walkthrough to contextualize the model used for generating voxel-wise rCBV maps as input for the BrainAGE model.

2D RA-UNet and 3D Extension

The first working DeepContrast model version was constructed using a residual attention U-Net style encoder-decoder trained on coronal slices.⁴³ Motivated by the inherently 3D continuity of anatomy and vascular patterns, we extended this model architecture to a volumetric residual-attention U-Net that processes 3D context directly.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Transformer-enhanced U-Net

We next evaluated the Transformer U-Net hybrid TransUNet encoder representations to better capture long-range dependencies that can be difficult for purely convolutional receptive fields,⁴⁸ while maintaining an image-to-image regression head that outputs voxel-wise rCBV maps. This direction is supported by prior work showing that Transformer-based representations can improve robustness and cross-dataset generalization in brain MRI tasks.⁴⁹

Patch-based 3D Transformer

We implemented a patch-based 3D Transformer that predicts voxel-wise rCBV from overlapping subvolumes and reconstructs a full-volume map by averaging overlapping predictions. This design is inspired by Transformer U-Net hybrid approaches that leverage long-range Transformer context in brain MRI.^{48,50} Each T1w scan was conformed to MNI template space and resampled to 1 mm isotropic resolution, then padded or cropped to $192 \times 192 \times 160$ voxels.³⁹ Overlapping $96 \times 96 \times 96$ patches were extracted with a sliding-window strategy with a stride of 6 voxels, each patch was independently mapped to predict a rCBV patch, and overlapping predictions were averaged voxel-wise to reduce blocking artifacts and improve spatial consistency.⁵⁰

DeepContrast Training Protocol

All models were optimized using stochastic gradient descent with a batch size of 4 and a robust adaptive regression loss. Evaluation was performed with the brain extracted scans. DeepContrast synthesis fidelity was quantified by whole-brain, voxel-wise comparisons between predicted maps and paired ground-truth rCBV maps using peak signal-to-noise ratio (PSNR, decibels [dB]) and structural similarity index measure (SSIM).

Full-cohort DeepCBV generation for BrainAGE

Following synthesis-fidelity benchmarking across model generations, the selected DeepContrast model was applied at inference time to the full multi-dataset cohort to generate DeepCBV maps from non-contrast T1w input alone, enabling CBV-sensitive vascular proxy feature extraction in datasets lacking contrast-enhanced scans.

BrainAGE Model

Our BrainAGE model was implemented using a VGG8 architecture in PyTorch.⁵¹ Each block consisted of a 3D convolutional layer, followed by 3D batch normalization, a rectified linear unit (ReLU) activation layer, and a 3D max pooling layer, with 5 repeated blocks making up the encoder portion of the model. The number of output channels in each convolutional layer was 16, 32, 64, 128, and 256. The final 3 fully connected layers map the flattened encoder output to 512, then 128, and finally 1 unit to produce the final brain age prediction.

A linear regression model was implemented in R (Version 4.4.1) to combine the predictions from each model. Each model's prediction was combined and used as a feature to predict the actual age of the patient. This combined approach integrates structural and functional age estimates, leveraging the unique strengths of each model. Additionally, we used sex as an additional covariate in our final model. Figure 2 showcases the full model pipeline from data collection to inference.

The dataset was first stratified based on age. Specifically, the age distribution was divided into deciles, which identified the 10th to the 90th percentiles of the age variable. The resulting quantiles, along with the minimum and maximum age values, were used to define the bin edges, and each scan was accordingly assigned to these age bins. The dataset was partitioned into ten separate stratified partitions based on age bins as well as the project origin of the scan. The dataset of 2,851 scans was split in an 8:1:1 ratio into training, validation, and test sets. The validation and test set were selected from the ten partitions with minimal Kullback-Leibler (KL) divergence compared to the dataset overall to ensure generalizable performance.

Preprocessing steps were previously described in Feng et al.,¹⁶ and were applied without modification. Briefly, T1w MRI images from all cohorts were first intensity bias field corrected,⁵² then brain extracted,⁵³ and finally registered to the standardized MNI152 affine space.³⁹ All images were manually checked at each stage to ensure image quality, whole-brain coverage, and the integrity of intermediate preprocessing outputs. We applied a modified intrascan min-max normalization procedure to each T1w image, where each voxel is divided by the mean of the top 1% voxel intensities in that scan. Images were excluded only if the original T1w scan had insufficient whole-brain coverage, image-obstructing artifacts were present, or a poor-quality image after any preprocessing step. For this current study, no images were excluded from the dataset we used.

The VGG model for both T1w and DeepCBV scans was trained for 100 epochs with early stopping if the mean absolute error (MAE) did not improve after 10 epochs. An AdamW optimizer with a learning rate of 1×10^{-4} and MAE loss was used to train the model with a batch size of 3.⁵⁴ All models were trained on the NVIDIA RTX A6000 GPU. After training the modality-specific deep learning models, predictions were generated separately for the training, validation, and held-out test sets. Predictions from the training and validation sets were used to fit a linear regression model combining T1w- and DeepCBV-derived age estimates. The fitted regression parameters were then applied to predictions from the held-out test set to generate final BrainAGE estimates for performance evaluation. All models and weights can be obtained from the following GitHub repository: <https://github.com/jzjomsky/DeepCBV-BrainAGE>.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Statistical Analysis

Linear regression analysis was performed using the chronological age of the patient as the dependent variable and the predicted age and the sex of the patient from each model as features with male coded as 1 and female coded as 0. We computed these results along with their residuals and coefficient of determination (R^2). In order to determine if the multimodal models showed statistically significant improvement over the unimodal models, we also performed a bootstrap analysis of the residual sum of squares (RSS).

To examine the validity of our BrainAGE estimates with known biomarkers of AD, we computed Spearman correlation coefficients (ρ) between BrainAGE values and each cognitive score. We then evaluated whether BrainAGE estimates differed across clinical stages of AD progression. To assess this, a Kruskal-Wallis test was used to detect whether there was a significant difference in brain age gaps between CN, MCI, and AD BrainAGE estimates from each model. After ensuring statistically significant results from these initial tests, pairwise Mann-Whitney tests were performed for each combination of AD diagnosis strata within each model. Similar Mann-Whitney tests were used to determine statistical significance in brain age gaps for each model in the SMCI and PMCI groups. Bonferroni correction was applied to correct for multiple comparisons ($\alpha = 0.05$). Additionally, we assessed group-level differences in BrainAGE using Cohen’s d effect sizes. Together, these analyses were used to evaluate both the continuous associations between BrainAGE and cognitive decline as well as the discriminative ability of BrainAGE to differentiate between clinical stages of disease progression.

Downloaded from <https://academic.oup.com/braincomms/advance-article/doi/10.1093/braincomms/fcag283/6740085> by guest on 28 July 2026

Results

DeepCBV synthesis across DeepContrast model generations

Comparison Pair / DeepCBV Method	PSNR (dB)	SSIM
rCBV vs. T1w	10.13	0.778
rCBV vs. DeepCBV (2D RA-UNet ⁴³)	29.64	0.879
rCBV vs. DeepCBV (3D RA-UNet ⁴⁸)	29.04	0.842
rCBV vs. DeepCBV (2D TransUNet ⁴⁸)	29.46	0.875
rCBV vs. DeepCBV (Patch-based 3D TransUNet ⁴⁸)	30.58	0.912

Table 1. Quantitative model performance comparison across DeepContrast model architectures. The comparison includes two-dimensional (2D) and 3D RA-UNet and TransUNet variants. Performance is reported using peak signal-to-noise ratio (PSNR, decibels [dB]) and structural similarity index measure (SSIM). Higher PSNR and SSIM indicate better agreement with the ground-truth rCBV maps. The rCBV vs. T1w MRI comparison provides a baseline for comparison of metrics. The patch-based 3D TransUNet achieved the highest performance (PSNR = 30.58 dB; SSIM = 0.912). Bold values indicate best performance across all models.

We evaluated DeepCBV synthesis fidelity by comparing DeepContrast-predicted maps against the ground-truth voxel-wise rCBV targets derived from paired pre-contrast and post-contrast T1w acquisitions. As summarized in Table 1, synthesis fidelity improved across successive DeepContrast model generations, consistent with architectural changes that progressively increase the ability to model long-range context and global structure beyond local convolutional neighborhoods.^{48,50}

The patch-based 3D Transformer model achieved the strongest overall fidelity by PSNR and SSIM and better recovered vascular patterns aligned with the rCBV target in representative slices shown in Figure 3. The stronger agreement between prediction and ground truth, relative to the raw T1w comparison, supports that the predicted DeepCBV maps encode contrast-related structure not directly present in the non-contrast image.⁴³ Based on these results, our patch-based 3D TransUNet DeepContrast model was selected as the DeepCBV generator and used to produce DeepCBV inputs for downstream BrainAGE analyses.

Overall Performance

Model	Mean Absolute Error (MAE)	Mean Squared Error (MSE)	R ²	Comparison	RSS Difference 95% CI	Effect Size of RSS Difference (Cohen's d)
T1w only (T-model)	4.10	29.02	0.936	T vs TD	[349.02, 1718.76]*	2.63
DeepCBV only (D-model)	4.49	33.34	0.927	D vs TD	[1122.83, 3081.36]*	4.02
T1w + DeepCBV (TD-model)	3.96	26.07	0.943	TD vs TDS	[1.13, 388.66]*	1.19
T1w + DeepCBV + sex (TDS-model)	3.95	25.98	0.943			

Table 2: Results by model. Performance metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), and R². Model improvements were assessed using bootstrap analyses of the Residual Sum of Squares (RSS). RSS differences are reported with 95% confidence intervals (CI); * denotes CI that exclude zero with statistically significant improvement (p < 0.05). Cohen's d reflects the standardized magnitude of RSS reduction for each model comparison. Bold values indicate best performance across all models.

We compared the results on the test dataset of four linear regression models created from our model architecture: (1) T1w-predicted age only (T-model), (2) DeepCBV-predicted age only (D-model), (3) combined T1w- and DeepCBV-predicted ages (TD-model), and (4) combined T1w- and DeepCBV-predicted ages with sex as an additional covariate (TDS-model), as shown in Table 2.

Table 2 contains the full results of each model created from the T1w-predicted age, DeepCBV-predicted age, and sex features with MAE, Mean Squared Error (MSE), and R². The T-model yielded a mean absolute error (MAE) of 4.10 years and R² of 0.936. The D-model (DeepCBV only) achieved an MAE of 4.49 years and R² of 0.927. Integrating both modalities in a linear regression model (combined TD-model) improved performance over the T-model and D-model,

achieving an MAE of 3.96 and R^2 of 0.943. Adding sex as a feature in the TDS-model marginally improved performance, reducing the MAE to 3.95 and keeping the same R^2 of 0.943.

In Figure 4, residual plots reveal a mild age-dependent bias, with predictions tending toward the mean age at the extremes of the age range, a pattern commonly observed in brain age estimation models. This behavior was consistent across all models and therefore does not affect relative performance comparisons. No explicit age-bias correction was applied in this analysis.

To assess the practical significance of performance differences between models, we performed nonparametric bootstrap resampling of the test set 1,000 times to estimate uncertainty in residual sum of squares (RSS) and differences in RSS between models. These results are also in Table 2. In all bootstrap comparisons, the 95% confidence intervals for RSS differences excluded zero, indicating statistically significant improvements for multimodal models.

The combined T1w + DeepCBV model (TD-model) showed a substantial reduction in RSS compared with the T1w-only model (RSS difference 95% CI = [349, 1719]; bootstrap $p = 0.001$), corresponding to a large effect size (Cohen's d (d) = 2.63, 95% CI: 0.98-4.80). Similarly, the combined model markedly outperformed the DeepCBV-only model (RSS difference = 1123-3081, 95% CI; $p = 0.001$), with a very large standardized effect ($d = 4.02$, 95% CI: 2.20-6.05).

Adding sex as a covariate to the combined model resulted in a smaller but measurable further reduction in RSS (RSS difference = 1-389, 95% CI; $p = 0.001$), corresponding to a moderate-to-large effect size ($d = 1.19$, 95% CI: 0.01-3.70). Together, these results indicate that while point-estimate improvements in MAE were numerically modest, the underlying reduction in prediction error was consistent, statistically robust, and associated with large effect sizes.

Visualizing Performance

To enhance clinical interpretability, we applied Grad-CAM to visualize the regions of T1w MRI and DeepCBV inputs that most influenced the model's predictions. For each modality, we extracted the top 20% of gradient values from the final convolutional layer and overlaid them on average T1w scans across 10-year age bins.

As shown in Figure 5, the T1w and DeepCBV encoders focus on distinct, complementary brain regions. For each modality and age group, slices were selected at the location of maximal Grad-CAM activation to highlight the brain regions most influential for the model's predictions. Slice locations vary across panels. A 3D Grad-CAM Visualization of top 20% gradient values averaged for CN participants from the test set of ages 65-100 from the MRI and DeepCBV CNNs to qualitatively characterize the different regions each modality emphasizes is located in the

Supplementary Material. The T1w encoder emphasizes white matter structures, while the DeepCBV encoder highlights vascular-rich areas, particularly in the central and lower brain.

Applicability to AD

To validate the clinical relevance of our BrainAGE models, we performed a series of analyses on 1,233 subjects from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort, focusing on the models’ ability to capture AD-related pathology.

First, we investigated the association between BrainAGE and the diagnostic status of AD. As shown in Figure 6, a stepwise increase in BrainAGE was observed with advancing disease severity across all models. Across diagnostic groups (CN, MCI, AD), Kruskal-Wallis tests demonstrated significant group differences for all models (T-model: $H = 30.18$, $p = 2.80 \times 10^{-7}$; D-model: $H = 33.62$, $p = 5.00 \times 10^{-8}$; TDS-model: $H = 40.89$, $p = 1.33 \times 10^{-9}$), supporting robust separation across disease stages.

CN individuals displayed the lowest BrainAGE, followed by those with MCI, and finally, patients with AD, who exhibited the highest BrainAGE. Post-hoc Mann-Whitney tests with Bonferroni correction demonstrated significant differences between CN and MCI for the T-model ($U = 8.037 \times 10^4$, $p = 7.73 \times 10^{-4}$), D-model ($U = 8.439 \times 10^4$, $p = 3.57 \times 10^{-2}$), and TDS-model ($U = 8.004 \times 10^4$, $p = 5.37 \times 10^{-4}$). We also report the effect size calculated using Cohen’s d (d) as 0.27, 0.20, and 0.28 for the T-model, D-model, and TDS-model respectively. Comparisons between CN and AD were highly significant across all models, including the T-model ($d=0.49$, $U = 2.128 \times 10^4$, $p = 1.55 \times 10^{-6}$), D-model ($d=0.55$, $U = 2.026 \times 10^4$, $p = 3.68 \times 10^{-8}$), and TDS-model ($d=0.58$, $U = 1.965 \times 10^4$, $p = 3.34 \times 10^{-9}$).

For MCI versus AD, significant differences were observed for the D-model ($d=0.33$, $U = 4.846 \times 10^4$, $p = 7.02 \times 10^{-4}$) and TDS-model ($d=0.30$, $U = 4.913 \times 10^4$, $p = 1.82 \times 10^{-3}$), whereas the T-model did not show a statistically significant difference between these groups ($d=0.23$, $U=5.21 \times 10^4$, $p=0.0661$). Overall, the TDS-model demonstrated the strongest and most consistent statistical separation across diagnostic strata.

Next, we evaluated the prognostic potential of the models to differentiate between individuals with MCI who remained stable (SMCI) versus those who progressed to AD (PMCI). Focusing specifically on MCI progression, Mann-Whitney U tests revealed that baseline BrainAGE was significantly higher in the PMCI group compared to the SMCI group across all three models. The T-model showed a significant difference ($d=0.30$, $U = 2.64 \times 10^4$, $p = 1.26 \times 10^{-2}$). The D-model demonstrated the strongest effect ($d=0.54$, $U = 2.18 \times 10^4$, $p = 4.43 \times 10^{-8}$), followed by the TDS-model ($d=0.46$, $U = 2.33 \times 10^4$, $p = 7.25 \times 10^{-6}$), highlighting the added value of incorporating DeepCBV-derived information for predicting disease progression.

Variable	BrainAGE Model	Spearman ρ	p-value
CDRSB	T-Model	0.477*	0.000609 *
	D-Model	0.241	0.0995
	TDS-Model	0.403*	0.00456*
MMSE	T-Model	-0.186	0.196
	D-Model	-0.369*	0.00828*
	TDS-Model	-0.310*	0.0286*
ADAS13	T-Model	0.175	0.225
	D-Model	0.215	0.134
	TDS-Model	0.190	0.187

Table 3: Cognitive tests and model associations. The Spearman correlation coefficients ρ and associated p-values for the relationship between BrainAGE values and various cognitive scores, including the Clinical Dementia Rating Sum of Boxes (CDRSB), the Mini-Mental State Examination (MMSE), and the Alzheimer's Disease Assessment Scale (ADAS).

Finally, we assessed the relationship between BrainAGE and established clinical scores using Spearman correlation analysis (Table 3). The BrainAGE values demonstrated significant associations with measures of dementia severity and cognitive function. BrainAGE was positively correlated with the Clinical Dementia Rating Sum of Boxes (CDRSB), indicating that an older-appearing brain corresponds to greater functional impairment. This relationship was significant for the T-model ($\rho = 0.477$, $p < 0.001$) and TDS-model ($\rho = 0.403$, $p = 0.005$) models. Conversely, BrainAGE was negatively correlated with the Mini-Mental State Examination (MMSE) score, where a higher BrainAGE is associated with poorer cognitive performance. This was significant for the DeepCBV BrainAGE ($\rho = -0.369$, $p = 0.008$) and Combo BrainAGE ($\rho = -0.310$, $p = 0.029$) models. No significant correlations were observed for the ADAS13 scores.

Discussion

In this study, we developed and validated a multimodal deep learning framework for BrainAGE that synergistically combines structural T1w MRI with DeepCBV maps. The results offer profound insights into the distinct and complementary biological processes that characterize brain aging. The primary contribution of this study is the demonstration that a learned vascular representation derived from non-contrast T1-weighted MRI can be used to generate biologically meaningful BrainAGE estimates. DeepCBV-based BrainAGE captures age-related signal and provides strong discrimination between stable and progressive mild cognitive impairment, supporting its relevance for modeling disease-related aging processes. By integrating DeepCBV with structural MRI, this framework highlights the value of vascular information for BrainAGE and clinical AD risk

stratification. This joint modeling of complementary imaging information more closely captures brain age and pathological effects on brain age, which allows the output of the TDS-model to serve as a biomarker for early detection of AD progression. The model assigned a small but statistically significant negative coefficient ($\beta = -0.168, p=0.02$) to the male sex, suggesting that male brains are consistently predicted as being biologically older than female brains of the same chronological age. This result aligns with established clinical evidence of sex-differentiated brain aging trajectories, further grounding our model's predictions in known physiological patterns.^{55–57}

The rationale for this multimodal approach is grounded in the limitations of unimodal imaging. T1w MRI is effective at capturing macroscopic anatomical changes like cortical and subcortical atrophy, which are well-established hallmarks of brain aging.⁵⁸ However, these structural alterations are often late-stage manifestations of underlying aging-related signals.^{59,60} The subtle physiological processes that precede them, such as changes in vasculature, remain largely invisible to conventional structural scans.^{61–63} The DeepCBV maps directly address this gap by providing a proxy for these functional dynamics. Our framework generates this vital functional information from standard, non-contrast T1w scans, thereby circumventing the significant clinical and logistical barriers associated with traditional functional imaging techniques. By leveraging a deep learning model to synthesize CBV maps, our approach democratizes access to functional brain information, making a multimodal assessment scalable, safe for repeated measures, and applicable to open-source neuroimaging databases.

To understand the mechanistic basis for this enhanced performance, we employed gradient-based class activation maps (Grad-CAM) to visualize the brain regions that most influenced each model's predictions. The T1w encoder consistently focused its attention on white matter structures, learning to detect age-related changes in white matter integrity and cortical loss.⁶⁴ In contrast, the DeepCBV encoder preferentially highlighted vascular-rich areas, particularly in the periventricular areas, capturing information about blood flow and metabolic activity. This confirms that the two modalities used in the TDS-model provide unique, non-redundant information, explaining why the combined model outperforms its unimodal counterparts.

The Grad-CAM analysis uncovered a biologically plausible dynamic shift in regional focus across the human lifespan, a finding that aligns with prominent neurocognitive theories of aging.⁶⁵ For the T1w-only model, the predictive focus in younger individuals was predominantly in the medial temporal lobe. As age increased, however, the model's attention transitioned to the inferior frontal cortices (IFC). The increased reliance on the IFC in older age suggests the model has learned that this region becomes a more salient predictor of age, perhaps following patterns of atrophy concordant with aging theory.⁵⁸

A parallel and compelling age-related transition was observed for the DeepCBV model. In younger populations, the focus was on the prefrontal cortex, likely reflecting the high metabolic demands

of executive functions. In older adults, however, the model's predictive focus shifted to deep brain structures, particularly the medial temporal region, indicative of the model learning to detect key patterns of age-related cerebrovascular decline.⁶¹ This resonates with theories of neurocognitive aging, namely Posterior-to-Anterior Shift in Aging (PASA), which is a functional model describing how older adults compensate for posterior sensory deficits by recruiting anterior prefrontal regions. In fact, while adhering to PASA,⁶⁶ the DeepCBV model's shift in focus also follows patterns of vascular damage in periventricular regions where vasculature vulnerability increases with age. Further, the periventricular region is a classic location for white matter hyperintensities (WMH), which are lesions indicative of chronic ischemia that are associated with more advanced age and cognitive decline.⁶⁻⁸ The enlargement of the lateral ventricles with age is a well-known hallmark of aging,⁵⁸ and the model's increased focus on the adjacent tissue suggests a particular sensitivity to vascular damage. Specifically, the DeepCBV model learned to identify regions of compromised vasculature and likely hypoperfusion as powerful predictors of brain aging. This validates that the DeepCBV-based predictions capture distinct and clinically relevant markers of cerebrovascular health, providing functional insights that complement the purely structural features identified by the T1w model.

Our combined TDS-model appears to offer a heightened sensitivity to the earliest signs of neurodegeneration. In diseases like AD, reduced cerebral blood flow and brain atrophy both contribute to the onset of neurodegeneration.⁶⁷⁻⁶⁹ A central aim of this study was to move beyond creating a statistically accurate age prediction model and develop a tool with tangible clinical relevance for age-related neurodegenerative diseases. By applying our models to a large, well-characterized cohort from ADNI, we have demonstrated that integrating DeepCBV with traditional T1w structural MRI provides a substantially more powerful biomarker for AD than either modality alone. The clear increase in BrainAGE from CN individuals to those with MCI and finally to patients with diagnosed AD aligns with the conceptualization of AD as a continuum of advancing neuropathology.

The T1w-based BrainAGE correlated most strongly with the CDRSB, which evaluates dementia severity, while the DeepCBV-based BrainAGE correlated most strongly with MMSE, a direct measure of cognitive status. This suggests that the vascular information captured by DeepCBV may be more tightly linked to the subtle network disruptions that manifest as cognitive deficits, particularly in the earlier stages of the disease, while dementia progression may be tightly coupled to structural changes. However, only the combined BrainAGE model was able to demonstrate significant correlations with both the CDRSB and the MMSE, which indicates that a multimodal approach successfully integrates the strengths of both models, presenting a biomarker that potentially reflects cognitive impairment.

A potentially significant clinical contribution of this work is the model's enhanced ability to predict which individuals with MCI progress to AD. Our results show that while the T-model could

significantly differentiate between SMCI and PMCI, its predictive power was dwarfed by the models incorporating DeepCBV. The D-model was the single strongest predictor by a substantial margin, with the combined model also showing improvement over the T-model as a result. This finding strongly suggests that at the MCI stage, functional and vascular changes are more sensitive markers of impending decline than the structural changes alone. However, the combined TDS-model demonstrated improved risk stratification in AD, integrating the strengths of both single modality models, which may allow for the enrichment of clinical trials with high-risk individuals and enable earlier, more targeted patient management.

Although the absolute improvements in MAE between models were modest and do not exceed current benchmarks,¹⁹ prior work has shown that brain-age models optimized exclusively for minimal age prediction error may exhibit reduced sensitivity to disease-related effects, whereas models with modestly lower age prediction accuracy can show enhanced sensitivity for disease detection and group discrimination.⁷⁰ Importantly, bootstrap analyses revealed that these small gains in MAE still corresponded to large and consistent reductions in residual error with effect sizes also indicating substantial improvements in model performance. From a clinical perspective, these observed improvements reflect not mere incremental optimization of age prediction accuracy, but increased reliability and clinical utility of BrainAGE for AD risk stratification.

A key limitation of this study is the use of generated vascular data. DeepCBV maps are synthesized from non-contrast T1-weighted MRI and therefore represent a learned vascular proxy rather than an independently acquired perfusion measurement. DeepCBV is derived from structural MRI to recreate contrast-enhanced data, therefore may still partially re-encode anatomical information already present in the T1-weighted image rather than capturing purely physiological vascular signal. Accordingly, DeepCBV-based BrainAGE should be interpreted as reflecting vascularly informed patterns of brain aging rather than direct cerebral blood volume measurements. Although our analyses demonstrated that DeepCBV contributes clinically relevant vascular information, future work incorporating independently acquired perfusion imaging may be necessary to further characterize the physiological efficacy and specificity of generated vascular maps.

In this study, we demonstrated that integrating anatomical (T1w MRI) and functional (DeepCBV) data improves BrainAGE prediction accuracy, validating the synergistic value of multimodal fusion of structural and functional imaging data. Future studies will further validate DeepCBV as a clinically deployable surrogate for CBV-fMRI, enabling non-contrast assessment of regional basal brain activity and metabolic demand. Combining robust modeling strategies with translational validation will help bridge the gap between research and clinical practice, advancing BrainAGE as a clinically actionable biomarker.

Data Availability

Data were provided in part by IXI, accessed from <http://brain-development.org/ixi-dataset>.

Data were provided in part by OASIS. OASIS Cross-Sectional: Principal Investigators: D. Marcus, R. Buckner, J. Csernansky J. Morris; P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382 OASIS: Longitudinal: Principal Investigators: D. Marcus, R. Buckner, J. Csernansky, J. Morris; P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382.

Data were provided in part by the Brain Genomics Superstruct Project of Harvard University and the Massachusetts General Hospital (Principal Investigators: Randy Buckner, Joshua Roffman, and Jordan Smoller), with support from the Center for Brain Science Neuroinformatics Research Group, the Athinoula A. Martinos Center for Biomedical Imaging, and the Center for Human Genetic Research. 20 individual investigators at Harvard and MGH generously contributed data to the overall project.

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern

California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data used in the preparation of this article was obtained from the Australian Imaging Biomarkers and Lifestyle flagship study of ageing (AIBL) funded by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) which was made available at the ADNI database (adni.loni.usc.edu). The AIBL researchers contributed data but did not participate in analysis or writing of this report. AIBL researchers are listed at www.aibl.csiro.au.

Data used in preparation of this article were obtained from the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) database (<https://ida.loni.usc.edu/login.jsp?project=NIFD>). The investigators at NIFD/FTLDNI contributed to the design and implementation of FTLDNI and/or provided data, but did not participate in analysis or writing of this report (unless otherwise listed).

Data collection and sharing for this project was funded by the Frontotemporal Lobar Degeneration Neuroimaging Initiative (National Institutes of Health Grant R01 AG032306). The study is coordinated through the University of California, San Francisco, Memory and Aging Center. FTLDNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (<https://www.ppmi-info.org/access-data-specimens/data>). For up-to-date information on the study, visit www.ppmi-info.org. PPMI - a public-private partnership - is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, the full names of all of the PPMI funding partners can found at <https://www.ppmi-info.org/about-ppmi/who-we-are/study-sponsors>.

Data used in preparation of this article were obtained from the SchizConnect database (<http://schizconnect.org>) As such, the investigators within SchizConnect contributed to the design and implementation of SchizConnect and/or provided data but did not participate in analysis or writing of this report. Data collection and sharing for this project was funded by NIMH cooperative agreement 1U01MH097435.

All data used during the study are available in the International Data-sharing Initiative (INDI, http://fcon_1000.projects.nitrc.org/) provided by the Southwest University Longitudinal Imaging Multimodal (SLIM) study.

Funding Sources

This research was conducted while Jia Guo was a Hevolution/American Federation for Aging Research New Investigator Awardee in Aging Biology and Geroscience Research [PT# 24-4274 PG014621]. This work was supported by a grant from the Simons Foundation [811245, S.A.S.] and a grant from the Simons Foundation International [SFI-AN-NC-AB-Research-00018029, S.A.S.].

Competing Interest

The authors have no conflicts of interest to disclose.

Appendix

Data used in preparation of this article were obtained from the Frontotemporal Lobar Degeneration Neuroimaging Initiative (FTLDNI) database. The investigators at NIFD/FTLDNI contributed to the design and implementation of FTLDNI and/or provided data, but did not participate in analysis or writing of this report (unless otherwise listed). The FTLDNI investigators included the following individuals:

Howard Rosen, Bradford C. Dickerson, Kimoko Domoto-Reilly, David Knopman, Bradley F. Boeve, Adam L. Boxer, John Kornak, Bruce L. Miller, William W. Seeley, Maria-Luisa Gorno-Tempini, Scott McGinnis and Maria Luisa Mandelli.

References

1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The Hallmarks of Aging. *Cell*. 2013;153(6):1194-1217. doi:10.1016/j.cell.2013.05.039
2. Hou Y, Dan X, Babbar M, et al. Ageing as a risk factor for neurodegenerative disease. *Nat Rev Neurol*. 2019;15(10):565-581. doi:10.1038/s41582-019-0244-7
3. Hahn O, Foltz AG, Atkins M, et al. Atlas of the aging mouse brain reveals white matter as vulnerable foci. *Cell*. 2023;186(19):4117-4133.e22. doi:10.1016/j.cell.2023.07.027
4. Dias MF, Duarte JV, de Carvalho P, Castelo-Branco M. Unravelling pathological ageing with brain age gap estimation in Alzheimer's disease, diabetes and schizophrenia. *Brain Commun*. 2025;7(2):fcaf109. doi:10.1093/braincomms/fcaf109
5. Brickman AM, Tosto G, Gutierrez J, et al. An MRI measure of degenerative and cerebrovascular pathology in Alzheimer disease. *Neurology*. 2018;91(15):e1402-e1412. doi:10.1212/WNL.00000000000006310

6. Laing KK, Simoes S, Baena-Caldas GP, et al. Cerebrovascular disease promotes tau pathology in Alzheimer's disease. *Brain Commun.* 2020;2(2):fcaa132. doi:10.1093/braincomms/fcaa132

7. Rizvi B, Lao PJ, Chesebro AG, et al. Association of Regional White Matter Hyperintensities With Longitudinal Alzheimer-Like Pattern of Neurodegeneration in Older Adults. *JAMA Netw Open.* 2021;4(10):e2125166. doi:10.1001/jamanetworkopen.2021.25166

8. Lee S, Viqar F, Zimmerman ME, et al. White matter hyperintensities are a core feature of Alzheimer's disease: Evidence from the dominantly inherited Alzheimer network. *Ann Neurol.* 2016;79(6):929-939. doi:10.1002/ana.24647

9. Lara FR, Scruton AL, Pinheiro A, et al. Aging, prevalence and risk factors of MRI-visible enlarged perivascular spaces. *Aging.* 2022;14(17):6844-6858. doi:10.18632/aging.204181

10. Wardlaw JM, Benveniste H, Nedergaard M, et al. Perivascular spaces in the brain: anatomy, physiology and pathology. *Nat Rev Neurol.* 2020;16(3):137-153. doi:10.1038/s41582-020-0312-z

11. Leidhin CN, McMorrow J, Carey D, et al. Age-related normative changes in cerebral perfusion: Data from The Irish Longitudinal Study on Ageing (TILDA). *NeuroImage.* 2021;229:117741. doi:10.1016/j.neuroimage.2021.117741

12. Hu J, Craig MS, Knight SP, et al. Regional changes in cerebral perfusion with age when accounting for changes in gray-matter volume. *Magn Reson Med.* 2025;93(4):1807-1820. doi:10.1002/mrm.30376

13. Markov NT, Lindbergh CA, Staffaroni AM, et al. Age-related brain atrophy is not a homogenous process: Different functional brain networks associate differentially with aging and blood factors. *Proc Natl Acad Sci U S A.* 2022;119(49):e2207181119. doi:10.1073/pnas.2207181119

14. Franke K, Ziegler G, Klöppel S, Gaser C, Alzheimer's Disease Neuroimaging Initiative. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *NeuroImage.* 2010;50(3):883-892. doi:10.1016/j.neuroimage.2010.01.005

15. Lancaster J, Lorenz R, Leech R, Cole JH. Bayesian Optimization for Neuroimaging Pre-processing in Brain Age Classification and Prediction. *Front Aging Neurosci.* 2018;10:28. doi:10.3389/fnagi.2018.00028

16. Feng X, Lipton ZC, Yang J, et al. Estimating brain age based on a uniform healthy population with deep learning and structural magnetic resonance imaging. *Neurobiol Aging.* 2020;91:15-25. doi:10.1016/j.neurobiolaging.2020.02.009

17. Dinsdale NK, Bluemke E, Smith SM, et al. Learning patterns of the ageing brain in MRI using deep convolutional networks. *NeuroImage*. 2021;224:117401. doi:10.1016/j.neuroimage.2020.117401
18. Dörfel RP, Arenas-Gomez JM, Fisher PM, et al. Prediction of brain age using structural magnetic resonance imaging: A comparison of accuracy and test-retest reliability of publicly available software packages. *Hum Brain Mapp*. 2023;44(17):6139-6148. doi:10.1002/hbm.26502
19. Yin C, Imms P, Cheng M, et al. Anatomically interpretable deep learning of brain age captures domain-specific cognitive impairment. *Proc Natl Acad Sci*. 2023;120(2):e2214634120. doi:10.1073/pnas.2214634120
20. Dartora C, Marseglia A, Mårtensson G, et al. A deep learning model for brain age prediction using minimally preprocessed T1w images as input. *Front Aging Neurosci*. 2023;15:1303036. doi:10.3389/fnagi.2023.1303036
21. Peng H, Gong W, Beckmann CF, Vedaldi A, Smith SM. Accurate brain age prediction with lightweight deep neural networks. *Med Image Anal*. 2021;68:101871. doi:10.1016/j.media.2020.101871
22. Siegel NT, Kainmueller D, Deniz F, Ritter K, Schulz M. Do Transformers and CNNs Learn Different Concepts of Brain Age? *Hum Brain Mapp*. 2025;46(8):e70243. doi:10.1002/hbm.70243
23. Yin C, Imms P, Chowdhury NF, et al. Deep learning to quantify the pace of brain aging in relation to neurocognitive changes. *Proc Natl Acad Sci*. 2025;122(10):e2413442122. doi:10.1073/pnas.2413442122
24. Liu G, Chaudhari N, Kanakaris N, Yin C, Bogdan P, Irimia A. Controllable Generative Model for Brain Evolution. In: 2025:1-5. doi:10.1109/ICASSP49660.2025.10888742
25. Cheng A, Duan S, Li S, et al. ERMoe: Eigen-Reparameterized Mixture-of-Experts for Stable Routing and Interpretable Specialization. *arXiv*. Preprint posted online November 14, 2025:arXiv:2511.10971. doi:10.48550/arXiv.2511.10971
26. Sweeney MD, Montagne A, Sagare AP, et al. Vascular dysfunction—The disregarded partner of Alzheimer’s disease. *Alzheimers Dement*. 2019;15(1):158-167. doi:10.1016/j.jalz.2018.07.222
27. Marmarelis V, Billinger S, Joe E, et al. Dysregulation of cerebral perfusion dynamics is associated with Alzheimer’s disease. *Alzheimers Dement Diagn Assess Dis Monit*. 2025;17(3):e70134. doi:10.1002/dad2.70134
28. Holmes AJ, Hollinshead MO, O’Keefe TM, et al. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Sci Data*. 2015;2(1):150031. doi:10.1038/sdata.2015.31

29. Liu W, Wei D, Chen Q, et al. Longitudinal test-retest neuroimaging data from healthy young adults in southwest China. *Sci Data*. 2017;4(1):170017. doi:10.1038/sdata.2017.17
30. Park DC, Hennessee JP, Smith ET, et al. The Dallas Lifespan Brain Study: A Comprehensive Adult Lifespan Data Set of Brain and Cognitive Aging. *Sci Data*. 2025;12(1):846. doi:10.1038/s41597-025-04847-7
31. Wei D, Zhuang K, Ai L, et al. Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Sci Data*. 2018;5(1):180134. doi:10.1038/sdata.2018.134
32. Ellis KA, Bush AI, Darby D, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr*. 2009;21(4):672-687. doi:10.1017/S1041610209009405
33. Zuo XN, Anderson JS, Bellec P, et al. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci Data*. 2014;1(1):140049. doi:10.1038/sdata.2014.49
34. Wang L, Alpert KI, Calhoun VD, et al. SchizConnect: Mediating neuroimaging databases on schizophrenia and related disorders for large-scale integration. *NeuroImage*. 2016;124(Pt B):1155-1167. doi:10.1016/j.neuroimage.2015.06.065
35. IXI Dataset – Brain Development. Accessed January 29, 2026. <https://brain-development.org/ixi-dataset/>
36. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *J Cogn Neurosci*. 2007;19(9):1498-1507. doi:10.1162/jocn.2007.19.9.1498
37. Marcus DS, Fotenos AF, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults. *J Cogn Neurosci*. 2010;22(12):2677-2684. doi:10.1162/jocn.2009.21407
38. Parkinson Progression Marker Initiative. The Parkinson Progression Marker Initiative (PPMI). *Prog Neurobiol*. 2011;95(4):629-635. doi:10.1016/j.pneurobio.2011.09.005
39. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*. 2002;17(2):825-841. doi:10.1016/s1053-8119(02)91132-8
40. Williams MM, Storandt M, Roe CM, Morris JC. Progression of Alzheimer disease as measured by Clinical Dementia Rating sum of boxes scores. *Alzheimers Dement J Alzheimers Assoc*. 2013;9(10):S39-S44. doi:10.1016/j.jalz.2012.01.005

41. Arevalo-Rodriguez I, Smailagic N, Roqué-Figuls M, et al. Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev*. 2021;2021(7):CD010783. doi:10.1002/14651858.CD010783.pub3
42. Skinner J, Carvalho JO, Potter GG, et al. The Alzheimer's Disease Assessment Scale-Cognitive-Plus (ADAS-Cog-Plus): an expansion of the ADAS-Cog to improve responsiveness in MCI. *Brain Imaging Behav*. 2012;6(4):489-501. doi:10.1007/s11682-012-9166-3
43. Liu C, Zhu N, Sun H, et al. Deep learning of MRI contrast enhancement for mapping cerebral blood volume from single-modal non-contrast scans of aging and Alzheimer's disease brains. *Front Aging Neurosci*. 2022;14. doi:10.3389/fnagi.2022.923673
44. Sun H, Liu X, Feng X, et al. Substituting Gadolinium in Brain MRI Using DeepContrast. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. 2020:908-912. doi:10.1109/ISBI45749.2020.9098323
45. Guo J, Small SA. System, method and computer-accessible medium for detecting functional disorder (s) or aging progression in magnetic resonance imaging. Published online December 31, 2024. Accessed February 8, 2026. <https://patents.google.com/patent/US12182969B2/en>
46. Guo J, Small SA. System, method and computer-accessible medium for detecting structural disorder (s) using magnetic resonance imaging. Published online September 24, 2024. Accessed February 8, 2026. <https://patents.google.com/patent/US12100121B2/en>
47. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002;17(3):143-155. doi:10.1002/hbm.10062
48. Chen J, Lu Y, Yu Q, et al. TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. *arXiv*. Preprint posted online February 8, 2021;arXiv:2102.04306. doi:10.48550/arXiv.2102.04306
49. Rao VM, Wan Z, Arabshahi S, et al. Improving across-dataset brain tissue segmentation for MRI imaging using transformer. *Front Neuroimaging*. 2022;1. doi:10.3389/fnimg.2022.1023481
50. Cao A, Rao VM, Liu K, Liu X, Laine AF, Guo J. TABSurfer: a Hybrid Deep Learning Architecture for Subcortical Segmentation. *arXiv*. Preprint posted online December 13, 2023;arXiv:2312.08267. doi:10.48550/arXiv.2312.08267
51. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv*. Preprint posted online April 10, 2015;arXiv:1409.1556. doi:10.48550/arXiv.1409.1556

52. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998;17(1):87-97. doi:10.1109/42.668698

53. Ségonne F, Dale AM, Busa E, et al. A hybrid approach to the skull stripping problem in MRI. *NeuroImage*. 2004;22(3):1060-1075. doi:10.1016/j.neuroimage.2004.03.032

54. Loshchilov I, Hutter F. Decoupled Weight Decay Regularization. *arXiv*. Preprint posted online January 4, 2019;arXiv:1711.05101. doi:10.48550/arXiv.1711.05101

55. Coffey CE, Lucke JF, Saxton JA, et al. Sex Differences in Brain Aging: A Quantitative Magnetic Resonance Imaging Study. *Arch Neurol*. 1998;55(2):169-179. doi:10.1001/archneur.55.2.169

56. Király A, Szabó N, Tóth E, et al. Male brain ages faster: the age and gender dependence of subcortical volumes. *Brain Imaging Behav*. 2016;10(3):901-910. doi:10.1007/s11682-015-9468-3

57. Armstrong NM, An Y, Beason-Held L, et al. Sex differences in brain aging and predictors of neurodegeneration in cognitively healthy older adults. *Neurobiol Aging*. 2019;81:146-156. doi:10.1016/j.neurobiolaging.2019.05.020

58. Blinkouskaya Y, Caçoilo A, Gollamudi T, Jalalian S, Weickenmeier J. Brain aging mechanisms with mechanical manifestations. *Mech Ageing Dev*. 2021;200:111575. doi:10.1016/j.mad.2021.111575

59. Feng X, Guo J, Sigmon HC, et al. Brain regions vulnerable and resistant to aging without Alzheimer's disease. *PLoS One*. 2020;15(7):e0234255. doi:10.1371/journal.pone.0234255

60. Zimmerman B, Rypma B, Gratton G, Fabiani M. Age-related changes in cerebrovascular health and their effects on neural function and cognition: A comprehensive review. *Psychophysiology*. 2021;58(7):e13796. doi:10.1111/psyp.13796

61. Zimmerman B, Rypma B, Gratton G, Fabiani M. Age-related changes in cerebrovascular health and their effects on neural function and cognition: A comprehensive review. *Psychophysiology*. 2021;58(7):e13796. doi:10.1111/psyp.13796

62. Miralbell J, Soriano JJ, Spulber G, et al. Structural brain changes and cognition in relation to markers of vascular dysfunction. *Neurobiol Aging*. 2012;33(5):1003.e9-17. doi:10.1016/j.neurobiolaging.2011.09.020

63. Vemuri P, Decarli C, Duering M. Imaging Markers of Vascular Brain Health: Quantification, Clinical Implications, and Future Directions. *Stroke*. 2022;53(2):416-426. doi:10.1161/STROKEAHA.120.032611

64. Irimia A. Cross-Sectional Volumes and Trajectories of the Human Brain, Gray Matter, White Matter and Cerebrospinal Fluid in 9473 Typically Aging Adults. *Neuroinformatics*. 2021;19(2):347-366. doi:10.1007/s12021-020-09480-w
65. McDonough IM, Nolin SA, Visscher KM. 25 years of neurocognitive aging theories: What have we learned? *Front Aging Neurosci*. 2022;14:1002096. doi:10.3389/fnagi.2022.1002096
66. Davis SW, Dennis NA, Daselaar SM, Fleck MS, Cabeza R. Que PASA? The posterior-anterior shift in aging. *Cereb Cortex*. 2008;18(5):1201-1209. doi:10.1093/cercor/bhm155
67. Zonneveld HI, Loehrer EA, Hofman A, et al. The bidirectional association between reduced cerebral blood flow and brain atrophy in the general population. *J Cereb Blood Flow Metab*. 2015;35(11):1882-1887. doi:10.1038/jcbfm.2015.157
68. Bracko O, Cruz Hernández JC, Park L, Nishimura N, Schaffer CB. Causes and consequences of baseline cerebral blood flow reductions in Alzheimer's disease. *J Cereb Blood Flow Metab*. 2021;41(7):1501-1516. doi:10.1177/0271678X20982383
69. Benedictus MR, Leeuwis AE, Binnewijzend MAA, et al. Lower cerebral blood flow is associated with faster cognitive decline in Alzheimer's disease. *Eur Radiol*. 2017;27(3):1169-1175. doi:10.1007/s00330-016-4450-z
70. Schulz MA, Siegel NT, Ritter K. Brain-age models with lower age prediction accuracy have higher sensitivity for disease detection. *PLOS Biol*. 2025;23(10):e3003451. doi:10.1371/journal.pbio.3003451

Figure Legends

Figure 1. Age distributions across datasets and model splits. A) Distribution of participant age across the full dataset (N = 2,851 MRI scans), shown as counts per 10-year interval. B) Distribution of age (years) by study cohort and sex across all included datasets. Each violin represents the distribution of individual MRI scans within that cohort (N per cohort shown on the x-axis). C) Distribution of age (years) across the training, validation, and test splits used for model development (total N = 2,851; N per split shown on the x-axis). Each violin represents individual MRI scans.

Figure 2. Brain age estimation pipeline. Overview of the multimodal framework used to estimate brain age from structural MRI. Raw T1-weighted (T1w) MRI scans are acquired in Digital Imaging and Communications in Medicine (DICOM) format and converted to Neuroimaging Informatics Technology Initiative (NIfTI) format following anonymization. Preprocessing includes affine registration to a standardized template space, skull stripping (removal of non-brain tissue), and final affine alignment. The preprocessed T1w volume is input into two parallel pathways. In the first pathway, the T1w scan is passed through a three-dimensional (3D) convolutional encoder to generate a predicted structural brain age (“T1 Age”). In the second pathway, a deep learning-based contrast synthesis model (DeepC) generates deep learning-derived cerebral blood volume maps (DeepCBV) from the same T1w scan. The DeepCBV volume is then processed through a separate 3D encoder to produce a predicted vascular-based brain age (“DeepCBV Age”). Predicted ages from the T1w and DeepCBV encoders are combined using a linear regression model. The symbols $\times a$ and $\times b$ denote learned regression coefficients applied to the T1 Age and DeepCBV Age predictions, respectively. Sex is included as an additional covariate (coded as 0/1) and multiplied by coefficient $\times c$. The green “+” node represents linear combination of these weighted inputs. The final output (“Combo Age”) represents the multimodal brain age prediction integrating structural, vascular, and sex-related information.

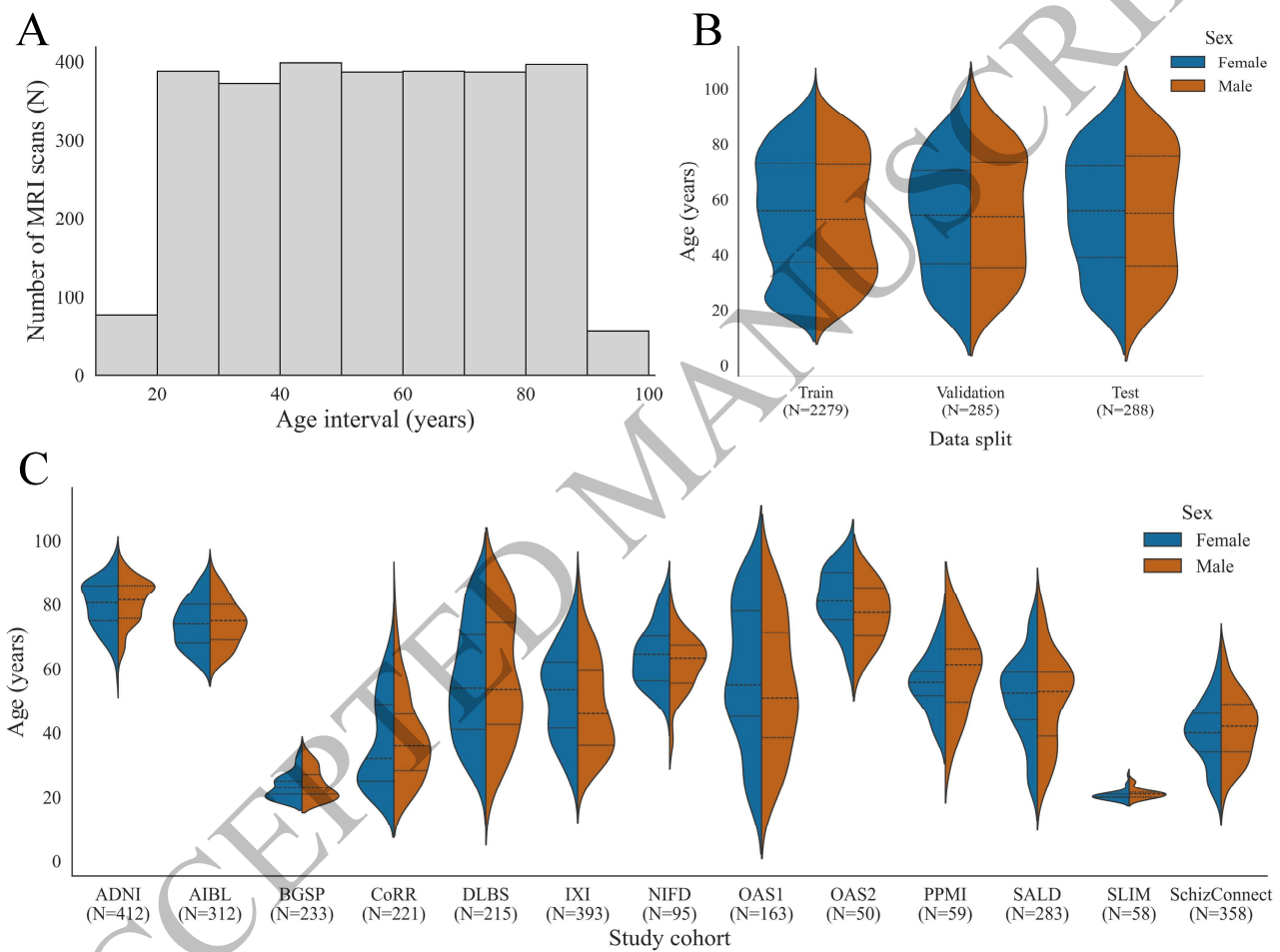
Figure 3. DeepCBV synthesis performance evaluation. Representative axial, coronal, and sagittal slices demonstrating the DeepCBV synthesis pipeline. Columns show (left to right): T1-weighted (T1w) magnetic resonance imaging (MRI) input, contrast-enhanced T1-weighted reference (T1w-CE), relative cerebral blood volume (rCBV) map used for supervision, and the predicted deep learning-derived cerebral blood volume map (DeepCBV) generated by the best-performing model (patch-based three-dimensional [3D] TransUNet). Intensity values are displayed in arbitrary units (a.u.) with the corresponding color scale.

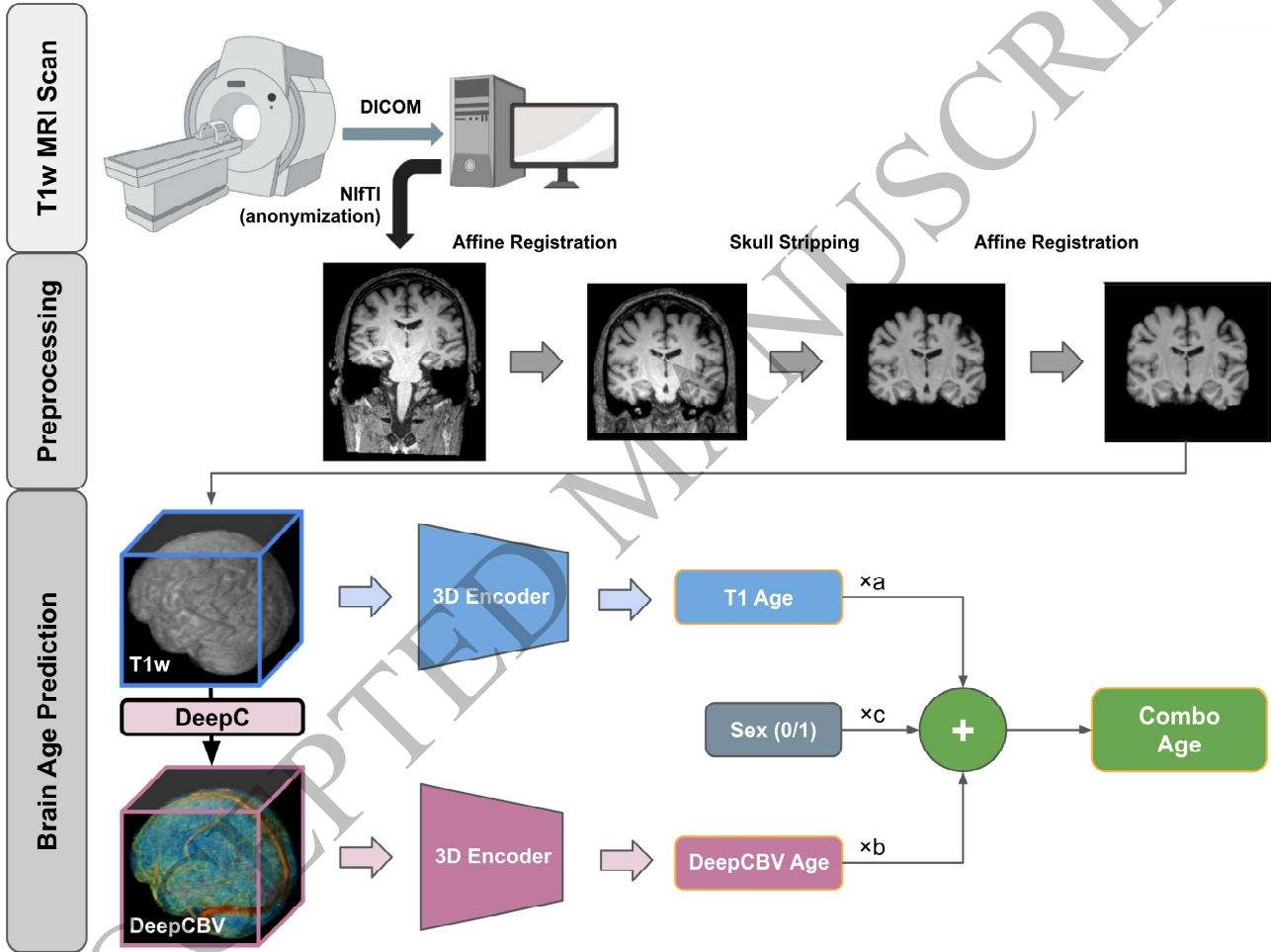
Figure 4: Model performance for brain age prediction across modalities. Predicted age versus chronological age (left column) and corresponding residual plots (right column) for the T-model (T1-weighted MRI), D-model (deep learning-derived cerebral blood volume; DeepCBV), TD-model (combined T1w + DeepCBV), and TDS-model (combined T1w + DeepCBV + sex). In the

scatter plots, the black solid line represents the identity line ($y = x$), and the red dashed line represents the fitted ordinary least squares (OLS) regression line. The slope and intercept of each fitted line as well as the mean (μ) and standard deviation (σ) of the fitted residuals are denoted for each subplot. All slopes and intercepts were deemed significant with $\alpha < 0.05$. Residual plots display prediction error (predicted age - chronological age) with a locally weighted smoothing curve (red) and a horizontal reference line at zero (black dashed).

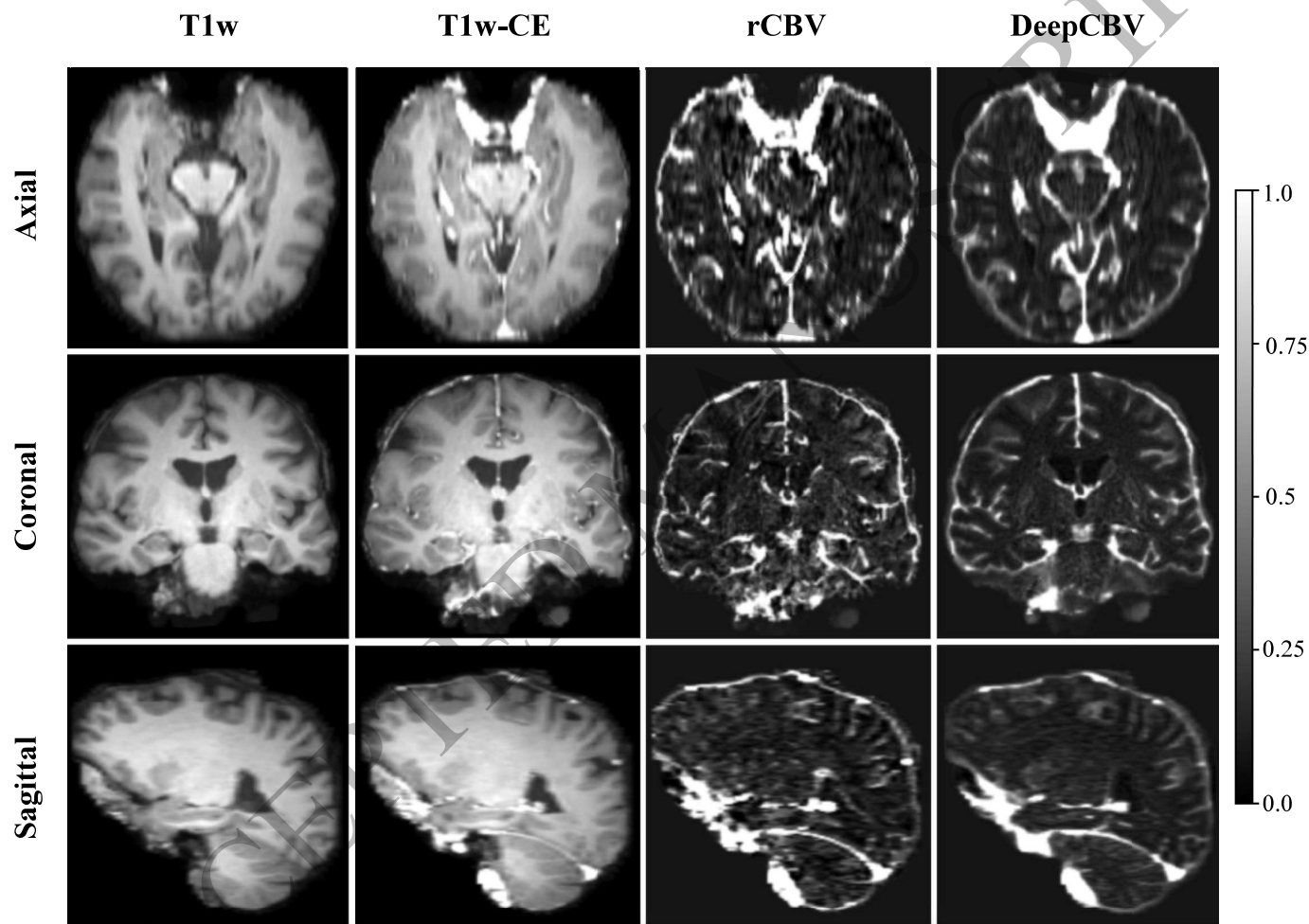
Figure 5. Modality-specific saliency maps across age groups. Average gradient-based saliency maps for the T1-weighted (T1w) and DeepCBV encoders across 10-year age intervals. Saliency maps were generated using gradient-weighted class activation mapping (Grad-CAM) applied to the final convolutional layer of each three-dimensional (3D) encoder. For each modality, the top 20% of voxel-wise gradient magnitudes were extracted and overlaid on the corresponding average T1w magnetic resonance imaging (MRI) scan within each age bin. For each modality and age group, slices were selected at the location of maximal Grad-CAM activation to highlight the brain regions most influential for the model's predictions and presented in radiological view. Higher gradient intensities indicate greater contribution of that region to the predicted brain age. DeepCBV denotes deep learning-derived cerebral blood volume maps synthesized from T1w MRI. Age groups are shown in 10-year intervals.

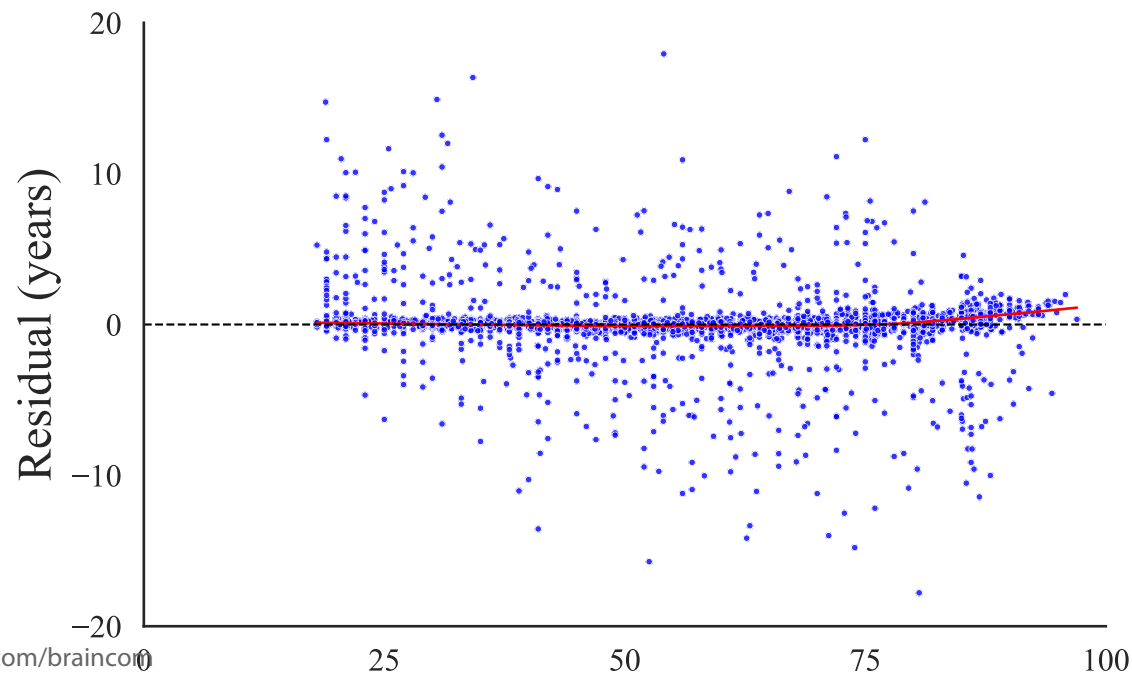
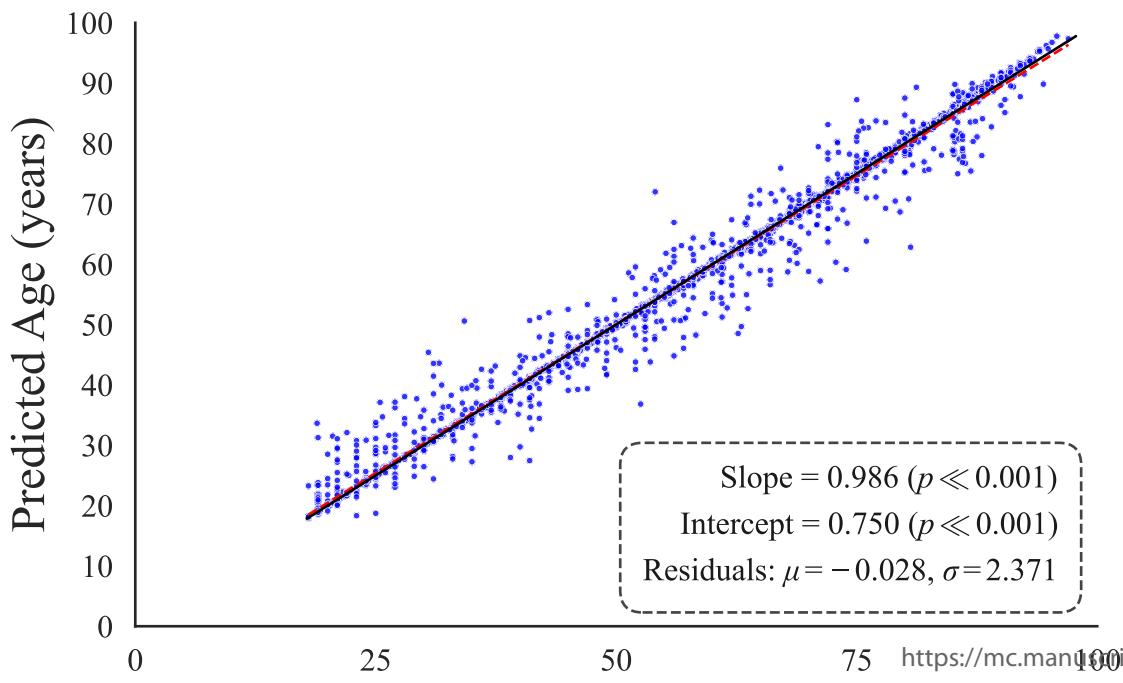
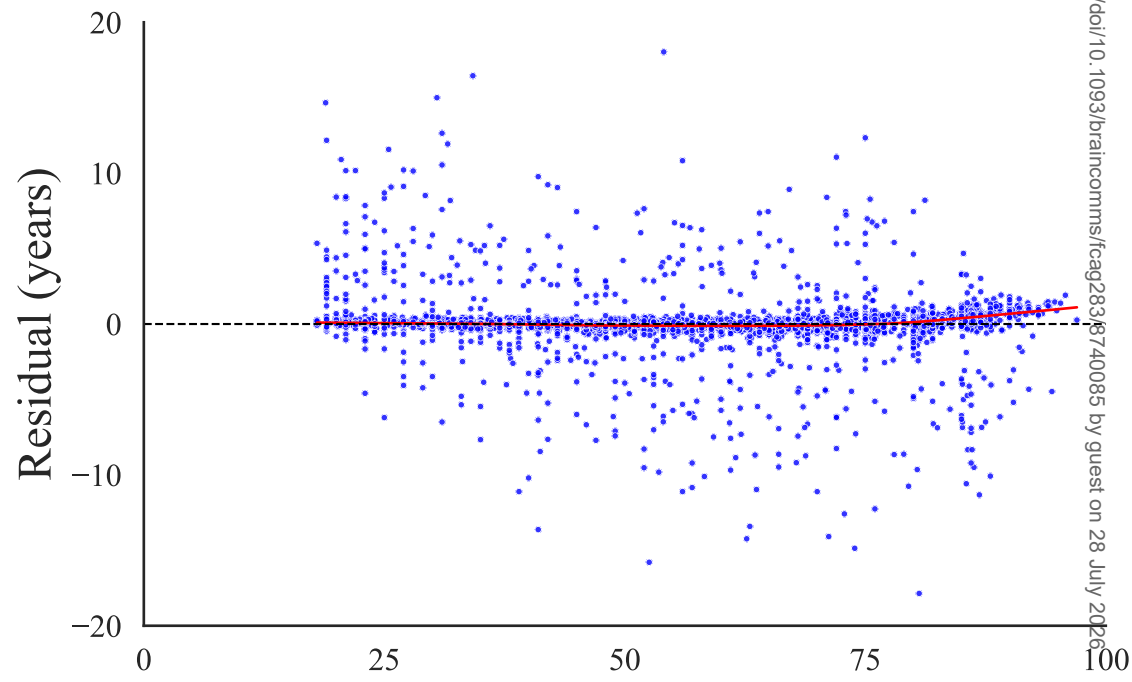
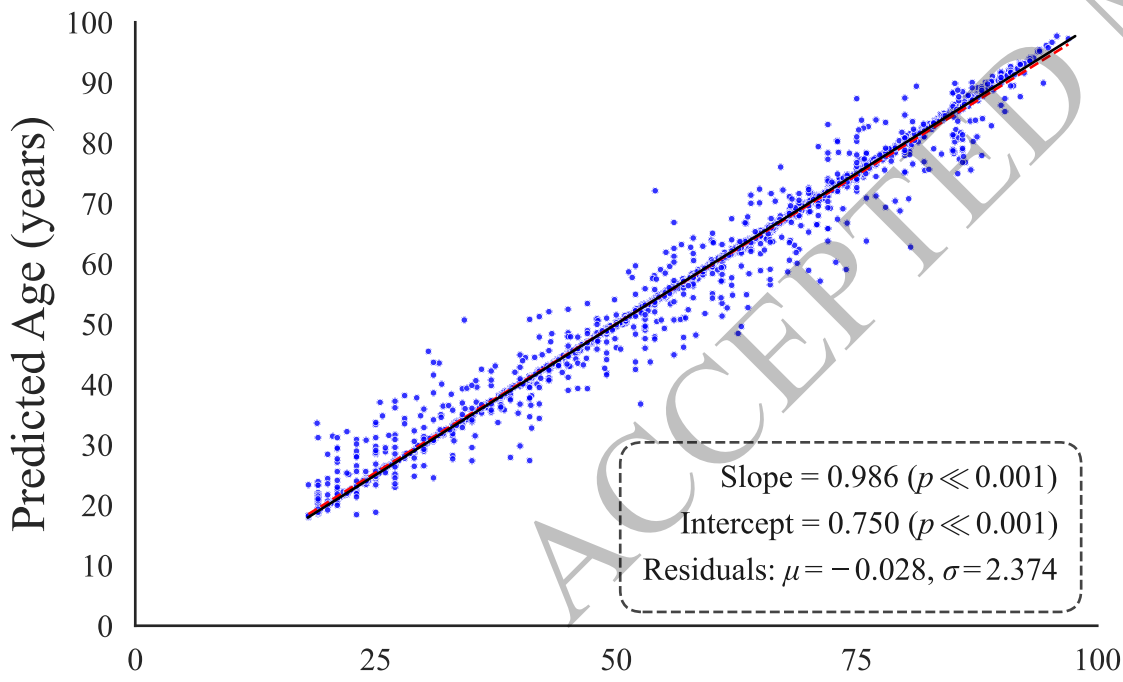
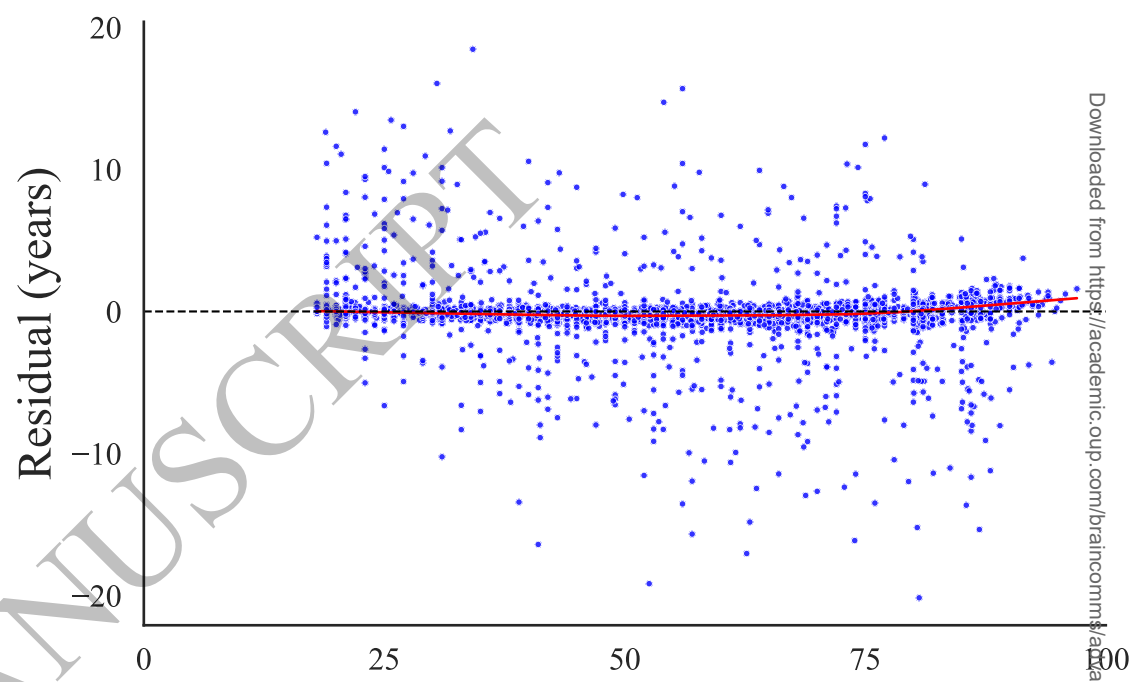
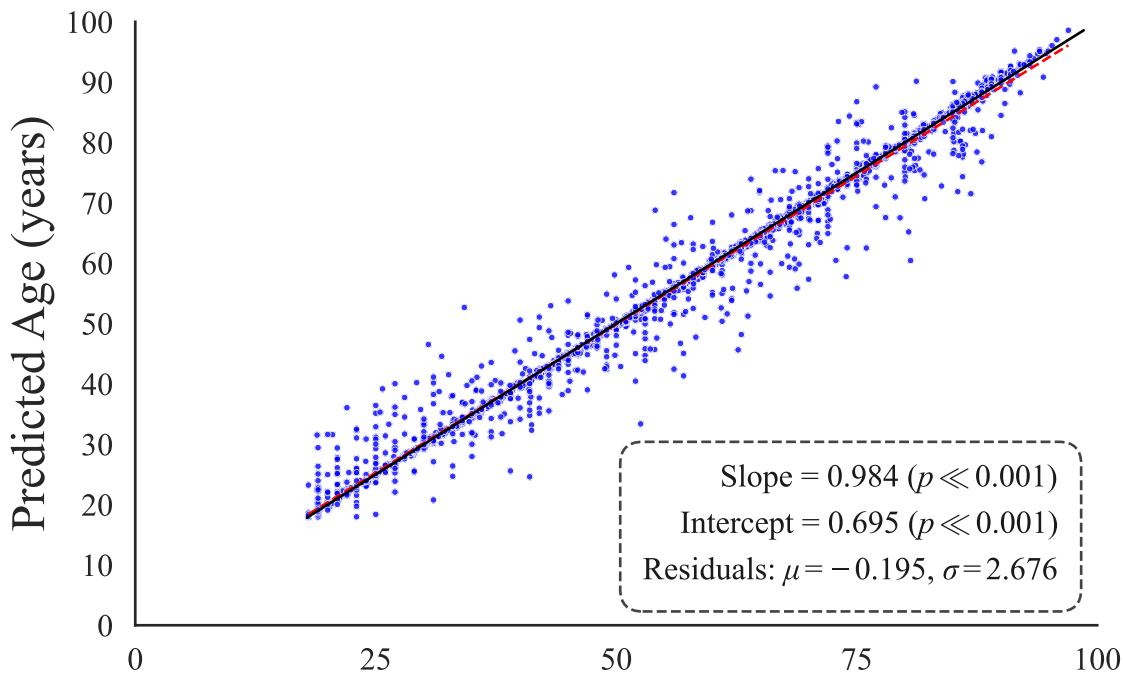
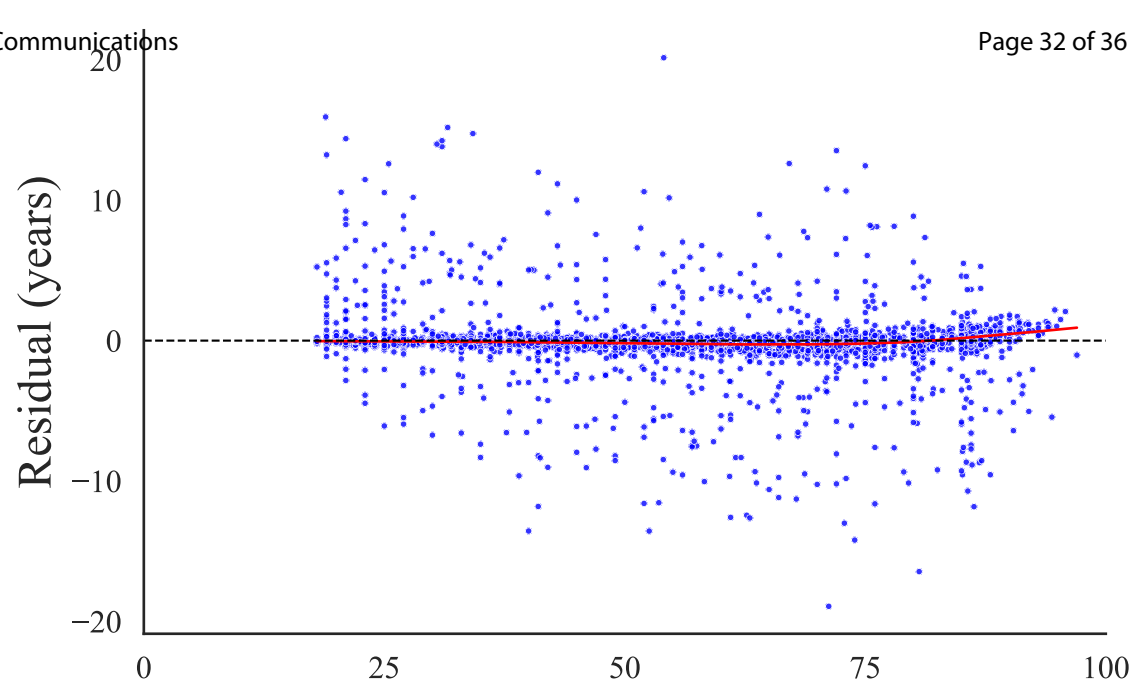
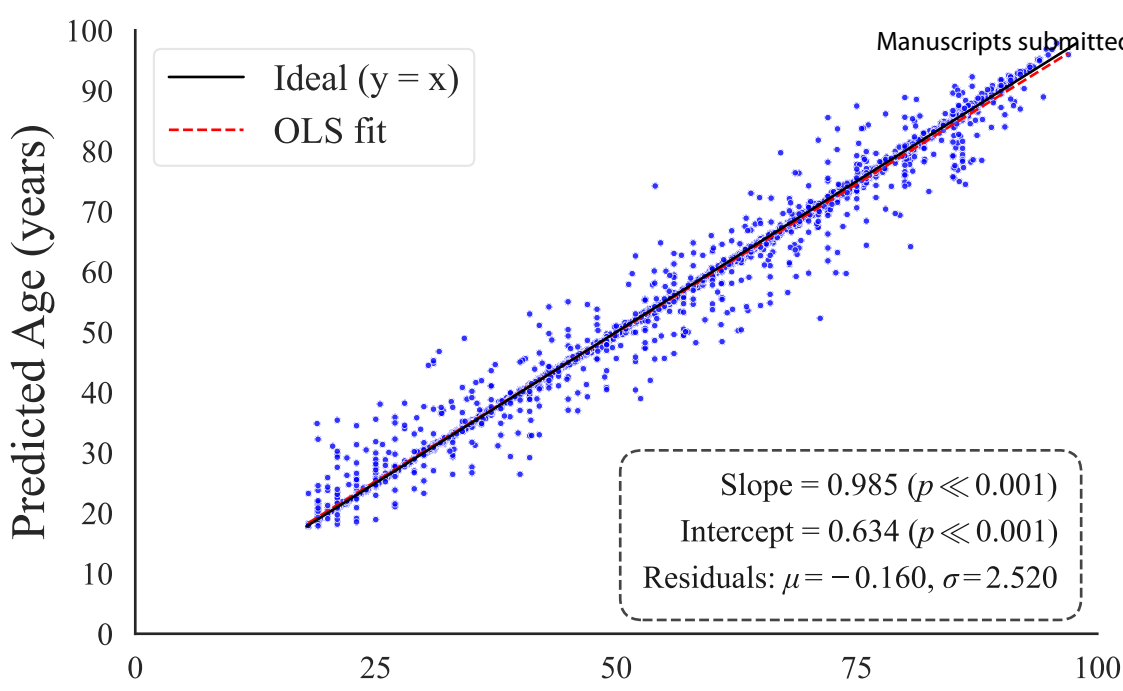
Figure 6: Boxplot for BrainAGE scores by Alzheimer's disease risk. A boxplot showing the distribution of BrainAGE scores for the T-model, D-model, and TDS-model across A) cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's Disease (AD) and B) stable MCI (SMCI) and progressive MCI (PMCI) with respective post-hoc Mann-Whitney tests with Bonferroni correction to show statistical significance. Statistical significance is indicated as follows: $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), and $p < 0.0001$ (****).





1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

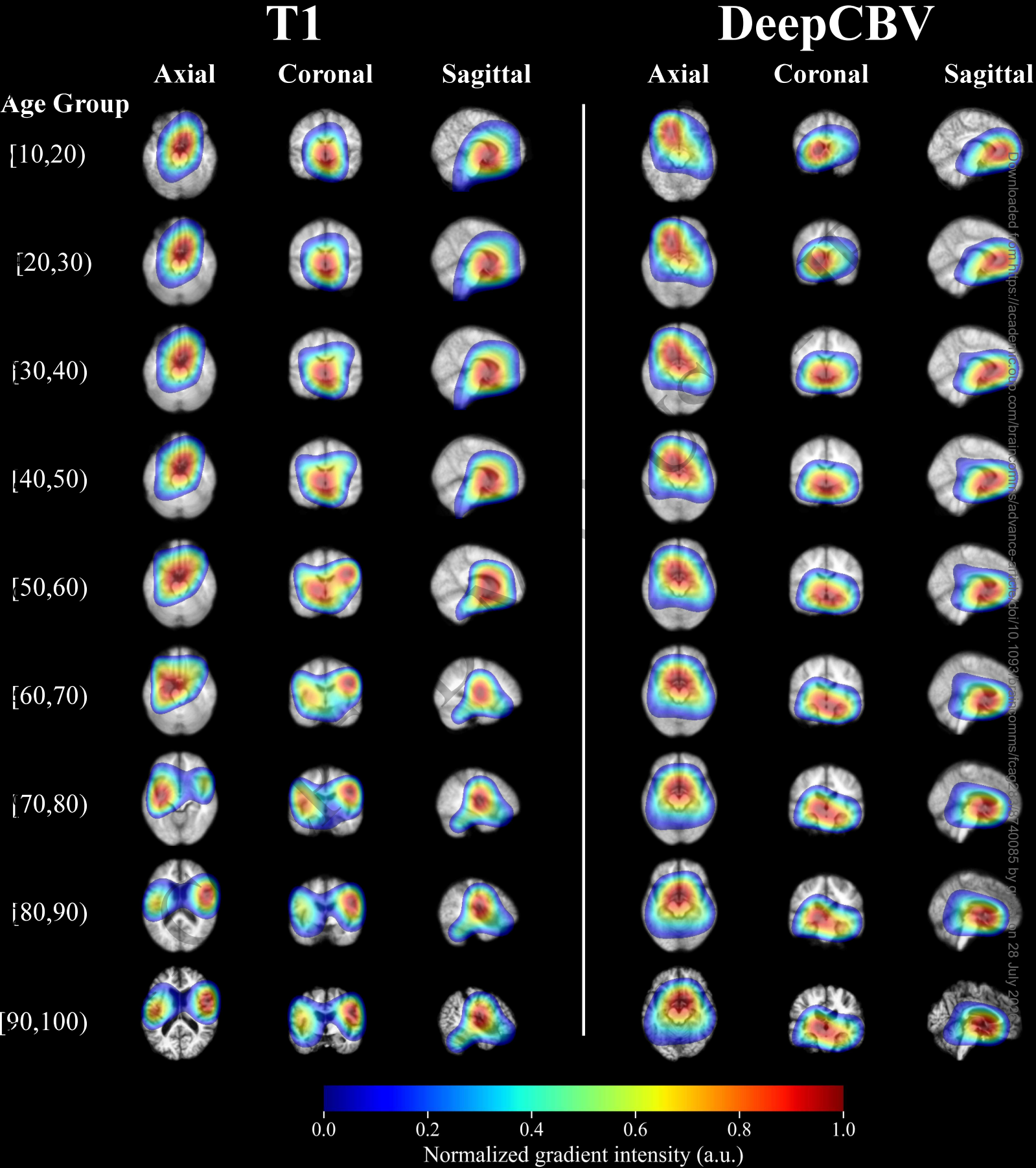


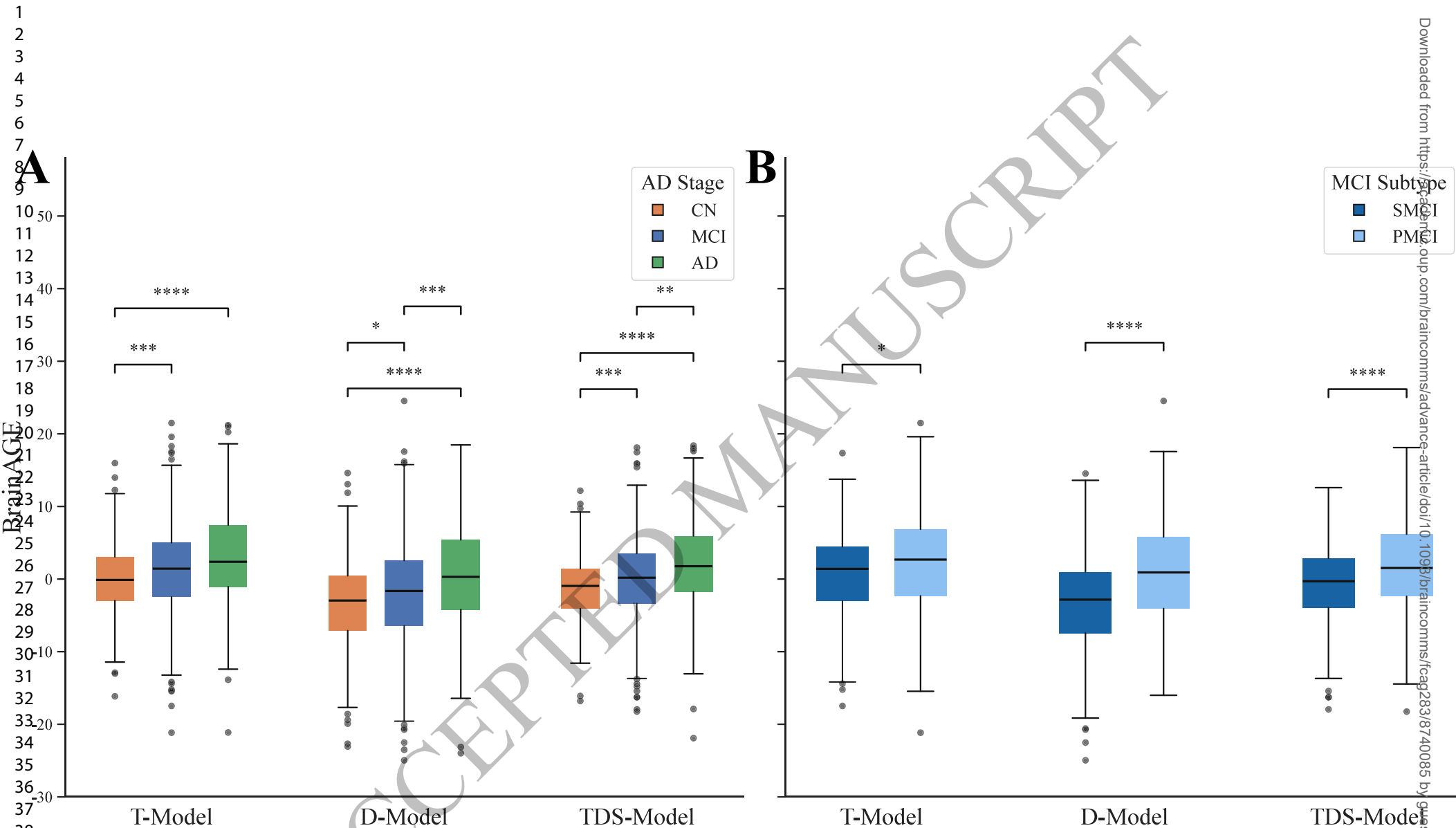


Actual Age (years)

Actual Age (years)

1





Combining structural MRI with derived cerebral blood volume maps in brain age gap estimation (BrainAGE) predicts pathological brain aging and Alzheimer's Disease (AD) progression from mild cognitive impairment (MCI)

