A novel deep learning-based brain age prediction framework for routine clinical MRI scans

Check for updates

Hyunwoong Kim[1,2], Seongbeom Park[1], Sang Won Seo[3,4,5,6], Duk L. Na[1,3], Hyemin Jang[7], Jun Pyo Kim[3,4], Hee Jin Kim[3,4,5,6], Sung Hoon Kang[8] & Kichang Kwak [1]✉

Physiological brain aging is associated with cognitive impairment and neuroanatomical changes. Brain age prediction of routine clinical 2D brain MRI scans were understudied and often unsuccessful. We developed a novel brain age prediction framework for clinical 2D T1-weighted MRI scans using a deep learning-based model trained with research grade 3D MRI scans mostly from publicly available datasets ($N = 8681$; age $= 51.76 \pm 21.74$). Our model showed accurate and fast brain age prediction on clinical 2D MRI scans from cognitively unimpaired (CU) subjects ($N = 175$) with MAE of 2.73 years after age bias correction (Pearson's $r = 0.918$). Brain age gap of Alzheimer's disease (AD) subjects was significantly greater than CU subjects ($p < 0.001$) and increase in brain age gap was associated with disease progression in both AD ($p < 0.05$) and Parkinson's disease ($p < 0.01$). Our framework can be extended to other MRI modalities and potentially applied to routine clinical examinations, enabling early detection of structural anomalies and improve patient outcome.

Brain aging is a complex and multifaceted process caused by the loss of neural circuits and synaptic plasticity, which is associated with cognitive decline and structural changes[1]. These changes are often associated with an increased susceptibility to neurodegenerative diseases[2] including Alzheimer's disease[3] (AD), Parkinson's disease[4], and Huntington's disease[5]. Structural and functional alterations associated with neurodegenerative diseases are currently considered largely irreversible. Furthermore, validated diagnostic biomarkers remain scarce; while a few exist for diseases with a known genetic etiology[6], those for sporadic forms are only beginning to develop or gain early regulatory approval[7,8]. AD develops at the culmination of a chronic pathophysiological process, with preclinical stages appearing decades earlier in life[9]. Therefore, understanding the links between brain aging process and the mechanism of neurodegenerative disease can significantly improve the early detection of these conditions.

Moreover, AD is believed to be a slowly evolving process that begins years to decades prior to the appearance of clinical symptoms[10]. This highlights the need of a cheap and efficient biomarker for AD and brain aging. Underlying structural changes may exist even in cognitively unimpaired individuals, and longitudinal studies using structural magnetic resonance imaging (MRI) indicate indicates these changes are associated with an increased risk of cognitive decline[3,11]. A novel biomarker to measure subtle changes of the brain structure can help the detection and early intervention of AD. While brain aging itself is a primary risk factor for the onset of these debilitating conditions, the exact mechanisms underlying the transition from normal brain aging to pathological neurodegeneration remain unclear.

Advanced neuroimaging tools allowed delineation of complex structural and functional changes of the brain[12]. MRI is a noninvasive structural imaging tool of the brain that is routinely used for both research and clinical purposes. While high-resolution T1 weighted MRI scans maximize the contrast difference between components of the brain and are commonly used to study brain structure[12], 2D MRI scans commonly available in clinical practice are essential for the diagnosis and monitoring of various brain conditions. In clinical practice, radiologists and clinicians identify brain atrophy by visually assessing the enlargement of the lateral ventricles, the sulcal widening between gyri, and the width of the temporal horn around the hippocampus, which increase linearly with age in healthy subjects, irrespective of sex[13]. However, these visual assessments lack accuracy compared

[1]BeauBrain Healthcare, Inc, Seoul, South Korea. [2]CHA University School of Medicine, Seongnam, South Korea. [3]Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea. [4]Alzheimer's Disease Convergence Research Center, Samsung Medical Center, Seoul, South Korea. [5]Department of Health Sciences and Technology, SAIHST, Sungkyunkwan University, Seoul, South Korea. [6]Department of Digital Health, SAIHST, Sungkyunkwan University, Seoul, South Korea. [7]Department of Neurology, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, South Korea. [8]Department of Neurology, Korea University Guro Hospital, Korea University College of Medicine, Seoul, South Korea. ✉e-mail: kichang.kwak@beaubrain.bio

to quantitative analyses due to the absence of objective metrics. Consequently, there is a need for accessible quantitative method to analyze dynamics of changes in the aging brain using 2D T1 imaging in clinical practice.

Recently, the term 'brain age' is widely being used as the predicted age from structural MRI scans of the brain[14,15]. The development of deep learning models allowed model architectures that perform well on natural images to be applied on domain-specific tasks. Most notably, convolutional neural network (CNN) models showed good performance on MRI scans[14,16]. A positive brain age gap or delta, defined as the age difference between predicted brain age and chronological age, is shown to indicate deviation from healthy aging in multiple studies[17,18]. Research-grade 3D T1 weighted MRI scans are mainly used to train brain age prediction models[16]. Only recently the importance of models targeted to two dimensional (2D) MRI scans more commonly used in clinical settings were highlighted[19–21]. However, these models were trained using in-house 2D MRI scans, and focused on modalities other than T1 weighted images[19] or T1 weighted images of healthy children aged 0 to 5 years old[21]. When a 2D CNN architecture (VGG16[22])-based brain age model trained successfully with 3D T1 weighted brain MRI scans[23] was applied to 2D T1 weighted scans, it showed significant brain age bias and mean absolute error (MAE) of 8.12 [7.19, 9.29] years even after age bias correction[20]. Another study suggested an alternative approach to train the same VGG16-based brain age prediction model[23] by using a mixture of actual research-grade 3D MRI scans and synthetic 1 mm isotropic MP-RAGE from clinical 2D MRI scans of various modalities[24]. Although the model failed to predict brain age of clinical MRIs accurate enough for clinical use, the study demonstrated that it is feasible to incorporate research-grade and clinical MRI scans for model training. Therefore, to accurately train the brain age prediction model targeted for clinical 2D T1 weighted MRI scans, development of both the model and an adequate MRI scan preprocessing pipeline for training are required.

In this study, we aim to build a 3D CNN based brain age prediction model specialized for clinical 2D axial T1 weighted MRI scans, by training it with publicly available research grade 3D T1 weighted MRI scans. To achieve this, we created a novel preprocessing pipeline that first slices research grade 3D scans into 2D axial scans, then interpolates the 2D axial scans back into 3D scans to use for training. This is intended to overcome the heterogenic nature of 2D axial scans caused by the larger slice thickness compared to other dimensions. We then tested the model on an independent test set of actual clinical 2D scans from healthy subjects. When the model is applied on actual clinical 2D scans, the scans are again interpolated into 3D scans. Finally, the model was tested on neurodegenerative disease cohorts to evaluate the association of brain age gap with different stages of neurodegenerative diseases.

Results

Training brain age prediction model with research grade 3D MRI scans

The biggest challenge of training models intended for clinical grade 2D MRI scans is the lack of publicly available routine clinical 2D MRI datasets. To overcome this problem, we collected 8681 research grade 3D MRI scans from the Samsung Medical Center (SMC) and 24 publicly available datasets (age mean ± SD: 51.76 ± 21.74, range: 16.22–95) and sliced the scans with axial gaps larger than 7 mm to mimic clinical grade 2D scans. The overall framework for training the model using 3D MRI scans and applying it to routine clinical 2D MRI scans are shown in Fig. 1.

The 3D DenseNet-169 based model for brain age prediction was successfully trained using interpolated 2D axial T1-weighted MRI scans created from slicing research grade 3D T1-weighted MRI scans (Fig. 2a). When the best model was selected using the validation dataset, MAE averaged over 5-fold cross-validation was 1.53 years with 95% CI [1.22, 1.84] and showed high correlation with chronological age (Pearson's r = 0.996 [0.994, 0.998]) during training and MAE 3.66 years with 95% CI [3.50, 3.81], r = 0.974 [0.970, 0.978] during validation. The trained model demonstrated accurate brain age prediction for the test dataset with MAE 3.68 years with

95% CI [3.52, 3.83] and showed high correlation with chronological age (r = 0.973 [0.970, 0.977]) (Fig. 2b). Minimal brain age bias by chronological age was observed during the training process (Fig. 2c). When the MAE for the test dataset was compared across the source datasets, we observed minimal variation for the performance of the model indicating the model was able to generalize across different age groups or demographics (Supplementary Table 1).

Clinical grade axial 2D T1-weighted MRI scans from SMC

The model also successfully predicted brain age from clinical grade axial 2D T1-weighted MRI scans of cognitively unimpaired subjects (Fig. 3a). For 175 axial 2D scans, the model achieved average 5-fold test MAE of 2.73 years with 95% CI [2.54, 2.92] after bias correction (r = 0.918 [0.901, 0.936]), and when using the ensemble of all 5 models from each CV runs the MAE was further reduced to 2.23 years. When applied to 199 axial 2D scans from patients diagnosed with AD, the model predicted positive brain age gap where mean corrected brain age gap was 3.10 [2.52, 3.67] years (Fig. 3b). Brain age gap was significantly higher in patients diagnosed with AD than cognitively unimpaired subjects (Fig. 3c; $p < 0.001$, Two sample Z-test). Brain age gap showed minimal brain age bias after linear bias correction (Fig. 3d), with mean corrected brain age gap of 0.09 [−0.66, 0.85] years. The trend of positive brain age gap was stronger in relatively younger individuals (Fig. 3e).

Guided backpropagation analysis showed that the model focus on cerebrospinal fluid (CSF) regions and their adjacent regions to predict brain age (Fig. 4a). This includes the lateral ventricle, the 3rd and 4th ventricles, and region involving the insula. For MRI scans from patients with Alzheimer's disease, the model correctly overestimates brain age by focusing on the larger CSF volumes due to brain atrophy (Fig. 4b). Full visualization of the guided backpropagation analysis is shown in Supplementary Fig. 1.
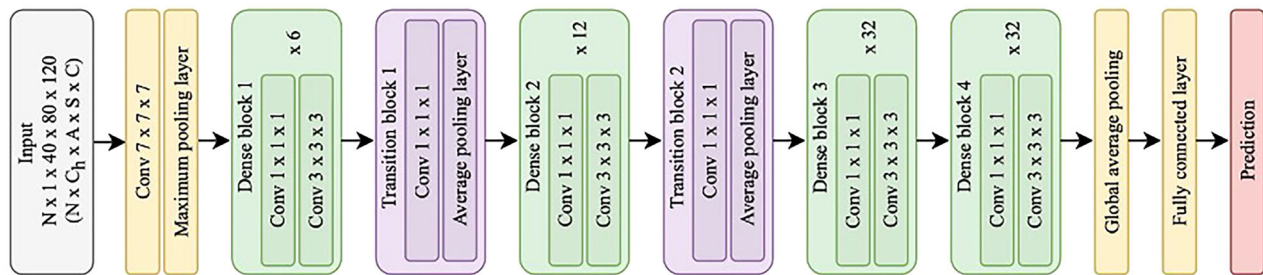
Brain age prediction on neurodegenerative disease cohorts

For the ADNI dataset, our model predicted brain age MAE of 2.05 years after age bias correction for cognitively unimpaired (CU) subjects with mean corrected brain age gap of 0.57 years (Fig. 5a). Brain age was then predicted for subjects in the mild cognitive impairment (MCI) and Alzheimer's disease (AD) cohorts with bias correction using coefficients learned from the CU cohort. The cohorts showed significantly different mean brain age gap ($p < 0.001$, Welch's ANOVA; brain age gap: 0.57 vs 2.15 vs 2.47 years in CU vs MCI vs AD cohorts). Further post hoc analysis showed the brain age gap of patients diagnosed with MCI or AD were significantly greater than CU subjects ($p < 0.001$, Pairwise Games–Howell test). Brain age gap was also significantly greater in patients with AD compared to patients with MCI ($p < 0.05$, Pairwise Games–Howell test).

To study further clinical implications of the model, we tested correlations between predicted brain age gap and the Mini-Mental State Examination (MMSE) score in patients with AD. From 2D MRI scans of patients with AD obtained from SMC, we observed a significant negative correlation between brain age gap and MMSE score (Supplementary Fig. 2a; r = -0.25, $p < 0.001$). However, this correlation was not statistically significant in 3D MRI scans of patients with AD in the ADNI dataset (Supplementary Fig. 2b; r = -0.066, $p = 0.18$)
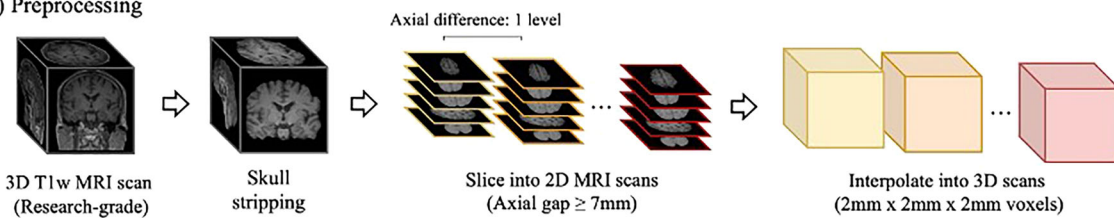
The PPMI dataset contains 3D research grade T1-weighted MRI scans for (1) subjects with no neurologic disorder and no first degree relative with PD (cognitively unimpaired cohort; CU), (2) participants who are at risk of Parkinson's based on clinical features, genetic variants, or other biomarkers (prodromal cohort; PRO), and (3) patients diagnosed with Parkinson's disease (PD cohort). Our brain age model predicted brain age MAE of 2.59 years after age bias correction for the CU cohort with mean corrected brain age gap of 1.03 years (Fig. 5b). Mean corrected brain age gap was 1.42 years for subjects in the PRO cohort and 2.23 years for patients in the PD cohort. The cohorts showed significantly different mean brain age gap ($p < 0.001$, Welch's ANOVA). Further post hoc analysis showed the brain age gap of subjects in the PRO or PD cohort were significantly greater than CU subjects ($p < 0.001$, Pairwise Games–Howell test). Brain age gap was also

(a) Illustration of the 3D DenseNet-169-based model architecture for age prediction
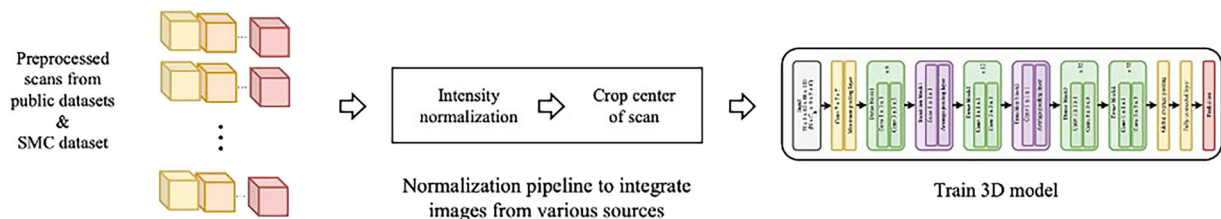
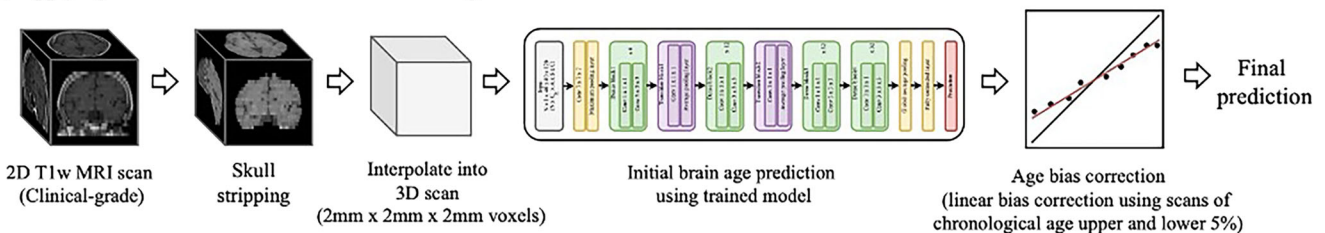(b) Creating training dataset from research-grade 3D T1w MRI scans

(1) Preprocessing

(2) Integrating various public datasets and SMC dataset

(c) Applying model trained with 3D T1w MRI images to 2D T1w MRI scans

Fig. 1 | Brain age prediction model architecture and overall research framework. a The age prediction model is based on the 3D version of DenseNet-169. **b** The process for creating the training dataset from research grade 3D T1-weighted MRI scans. The 3D scans are first sliced to generate synthetic 2D axial scans, then interpolated back into a common 3D voxel space. **c** Applying the trained model to inference brain age of actual clinical-grade 2D T1-weighted MRI scans. The 2D scans are interpolated into 3D scans for initial brain age prediction, then age bias correction is applied to produce the final prediction.

significantly greater in patients diagnosed with PD compared to subjects in the prodromal phase ($p < 0.01$, Pairwise Games–Howell test).

Discussion

In this study, we have demonstrated accurate brain age prediction on clinical grade 2D axial T1 weighted MRI scans using a 3D DenseNet based model trained with research grade 3D T1 weighted MRI scans. This training pipeline can potentially be applied to MRI scans of other modalities. Our model achieved MAE of 3.68 years in brain age prediction for axial-sliced 3D T1 weighted MRI scans, and MAE of 2.73 years after age bias correction for clinical grade 2D axial MRI scans. This is a great improvement compared to previous brain age prediction models targeted for clinical grade 2D T1 weighted MRI scans, and even comparable to the performance of state-of-the-art models that uses entire research grade 3D MRI scans[14]. We were able

to achieve such performance with a dataset covering a wide range of age distribution with range from 16.22 to 95 years, and standard deviation of 21.74 years. While previous studies using 2D CNN architecture models or trained with in-house 2D axial T1 scans have failed to accurately predict brain age of clinical grade 2D axial T1 weight MRI scans[20,24], we overcame this problem by (1) employing a 3D CNN architecture and (2) developing a pipeline to leverage widely available research grade 3D MRI scans for training. Given common use of axial 2D MRI scans in routine clinical examinations and primary healthcare settings, our findings provide a reliable result for precise brain age prediction, potentially enabling early detection of neurological abnormalities and improving patient outcomes. Increase in brain age gap (i.e., older predicted brain age) of our model was associated with disease progression in both Alzheimer's disease and Parkinson's disease. The results of this study show the feasibility of utilizing
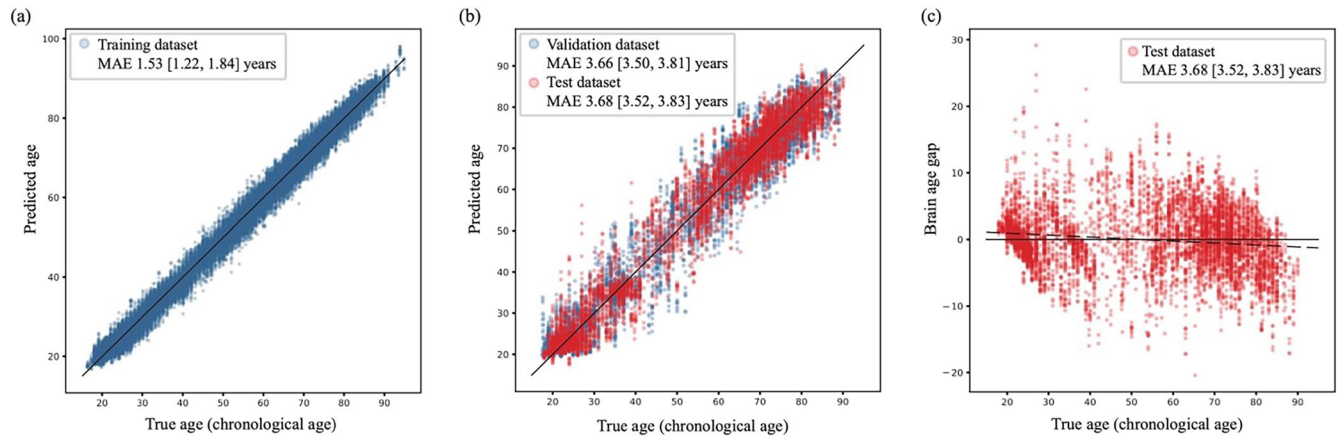
Fig. 2 | Training brain age prediction model using interpolated 2D axial T1-weighted MRI scans created from research grade 3D T1-weighted MRI scans. 95% confidence intervals are shown with MAE averaged over 5-fold cross-validation. **a** The 3D DenseNet-169 based model was successfully trained using 3D scans interpolated from the 2D axial scans (Training dataset: 55,602 scans created from 6903 research grade 3D scans, training MAE = 1.53 years with 95% CI [1.22, 1.84],

Pearson's r = 0.996 [0.994, 0.998]). **b** The validation dataset was used for early stopping, and no further adjustments were made to the model output (Validation dataset: 7035 scans created from 863 research grade 3D scans, validation MAE = 3.66 [3.50, 3.81] years, r = 0.974 [0.970, 0.978]; Test dataset: 6961 scans created from 863 research grade 3D scans, test MAE = 3.68 [3.52, 3.83] years, r = 0.973 [0.970, 0.977]). **c** Minimal brain age gap bias by chronological age was observed.
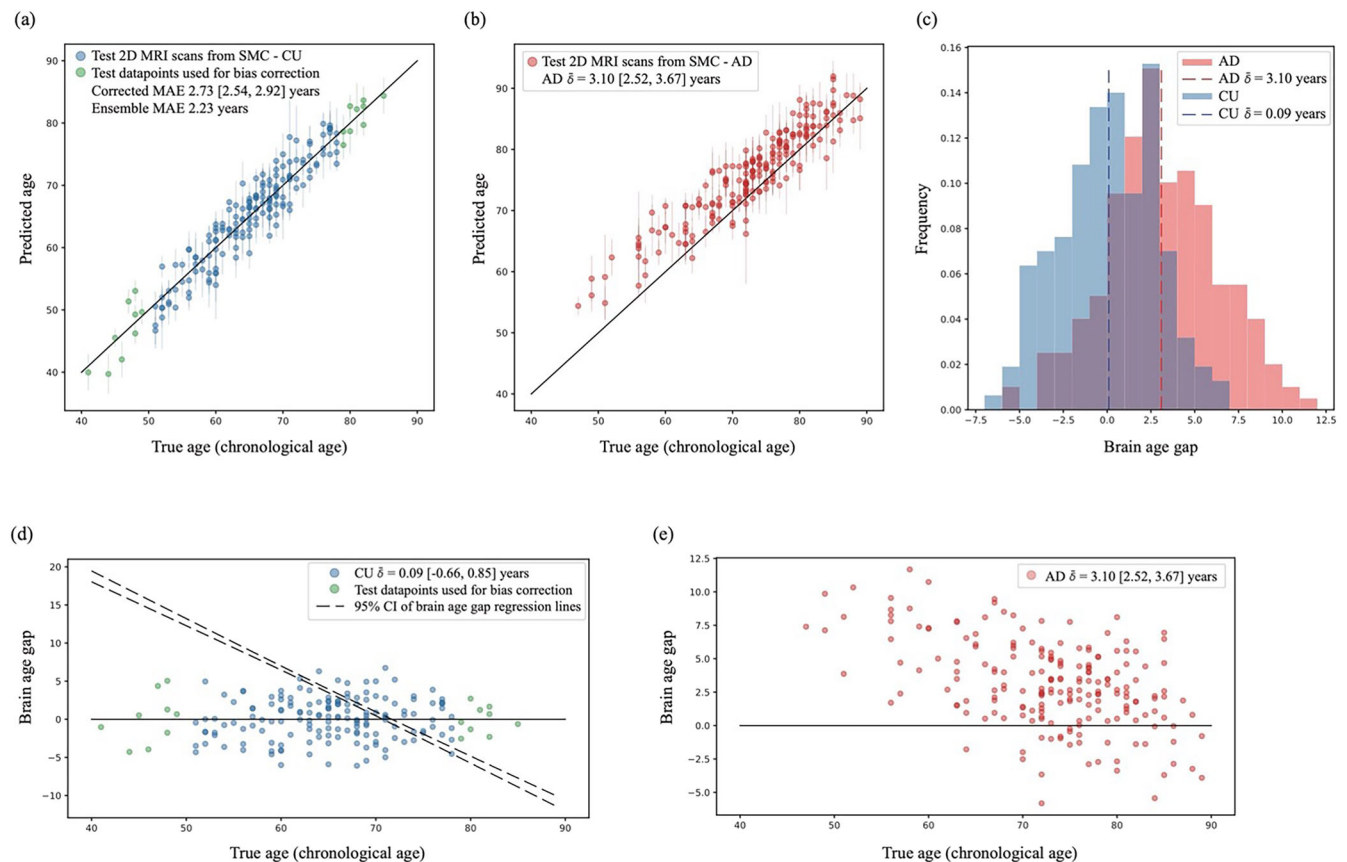

Fig. 3 | Brain age prediction of clinical grade axial T1-weighted MRI scans (2D scans). 95% confidence intervals across 5-fold cross-validation are shown with error bars. **a** When the trained model was tested on cognitively unimpaired individuals from an independent dataset from SMC, the model successfully predicted brain age after brain age bias correction using oldest and youngest 5% individuals (Test dataset: 175 2D axial scans, test MAE = 2.73 years with 95% CI [2.54, 2.92] after bias correction, Pearson's r = 0.918 [0.901, 0.936]; when using ensemble of all 5 models from each CV runs - MAE = 2.23 years). **b** Individuals diagnosed with Alzheimer's

disease showed positive brain age gap and this trend was stronger in relatively younger individuals (CU: n = 175, mean corrected brain age gap = 0.09 [−0.66, 0.85] years; AD: n = 199, mean corrected brain age gap = 3.10 [2.52, 3.67] years). **c** Brain age gap was significantly higher in individuals diagnosed with AD compared to CU individuals ($p < 0.001$, Two sample Z-test). **d** Brain age gap in CU individuals after linear age bias correction. Dotted lines indicate 95% confidence bands of regression lines used in bias correction. **e** Brain age gap in individuals diagnosed with AD after age bias correction with coefficients learned from CU individuals.
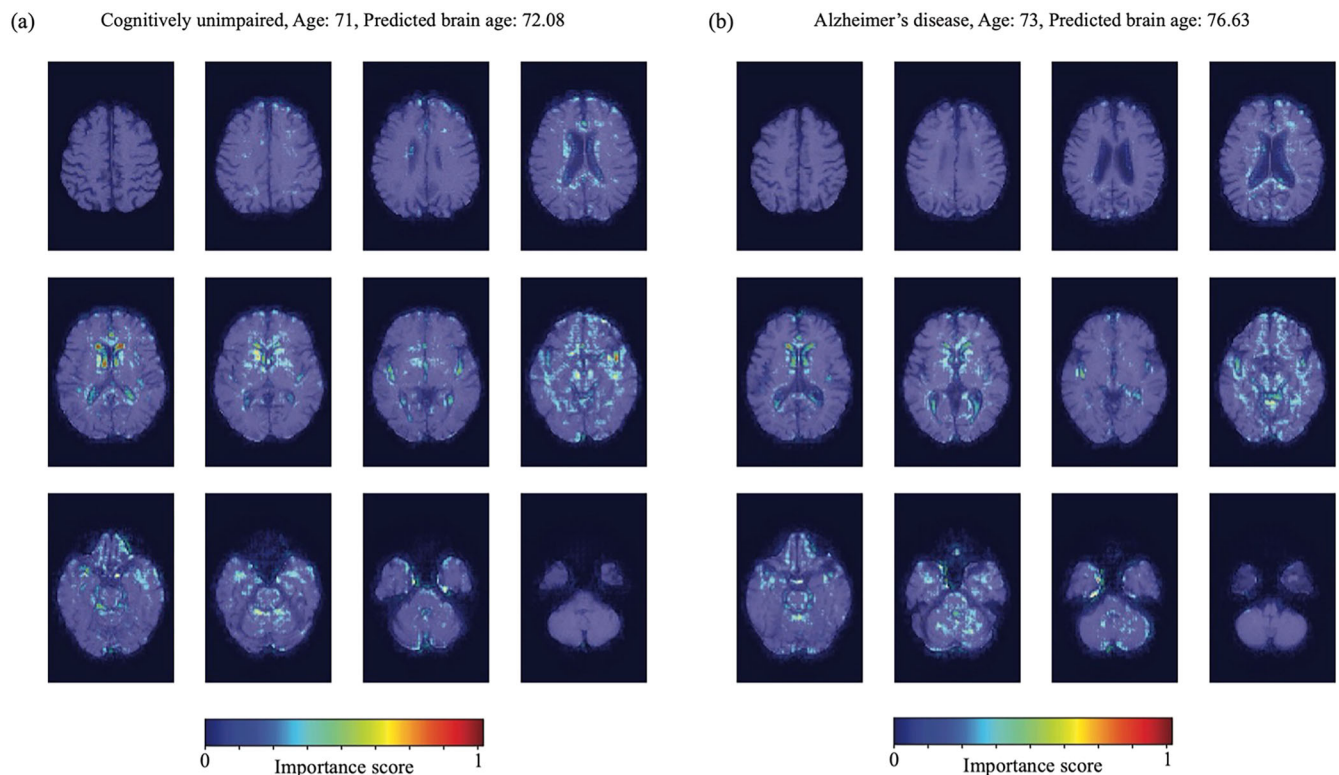
Fig. 4 | Visualization of brain age prediction model using guided back-propagation. Clinical grade axial T1-weighted MRI scans (2D scans) from (**a**) a cognitively unimpaired individual and (**b**) an individual diagnosed with Alzheimer's disease were used. The model seems to focus on cerebrospinal fluid spaces and their adjacent regions including the lateral ventricle, 3rd and 4th ventricles, and region involving the insula.

brain age gap as a screening tool to assess the risk of neurodegenerative diseases during routine clinical examinations. Negative correlation between brain age gap from clinical 2D MRI scans and MMSE was observed in the AD cohorts of the SMC dataset, suggesting additional clinical values of our brain age prediction model. However, this correlation was weaker and not statistically significant when calculated for the 3D MRI scans in the AD cohorts from the ADNI dataset. One possible explanation for this discrepancy could be due to the difference of the homogeneity of the datasets. The SMC dataset hold data obtained from a single center with the same scanner and protocol whereas the ADNI dataset is a collection of datasets from various centers, and this could introduce biases in both MRI imaging and MMSE scoring.

Another important consideration when developing models for routine clinical use is inference speed and computation resource usage. Although 2D CNN models have cheaper computational cost, our trained 3D DenseNet based model takes only 1.3 seconds to (1) load the model and a clinical 2D axial scan saved in NIfTI format, then (2) predict brain age. This experiment was performed using a NVIDIA L40S GPU and required less than 1GB of GPU memory. Therefore, using a pretrained 3D CNN model only for brain age inference would not be a challenge in clinical settings.

Guided backpropagation reveled CSF and their adjacent regions are important in predicting brain age for our model. This involves multiple major CSF regions, indicating that overall CSF volume is important in predicting brain age. Previous study using clinical grade 2D T2 weighted MRI scans to predict brain age also showed that CSF regions are associated with brain age prediction[19]. One possible explanation for this is that overall brain atrophy from aging leads to larger CSF volumes.
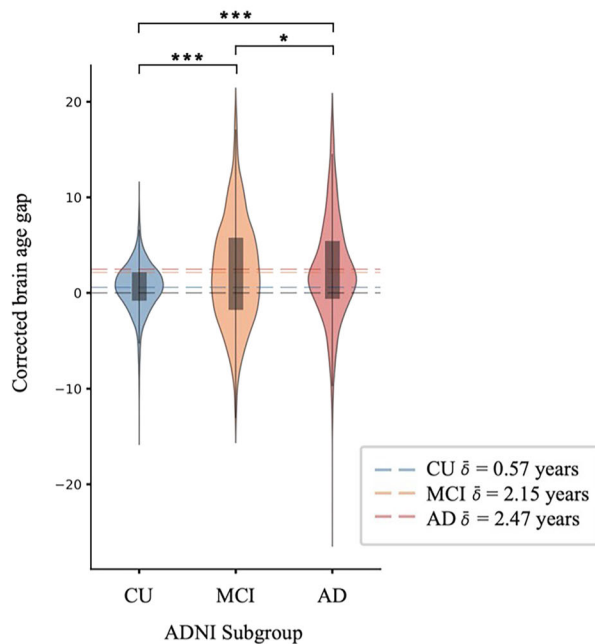
Generalizability of a model is a critical factor in developing brain age prediction models. Scans from different scanner vendors or hospitals contain inherent heterogeneity, which leads to suboptimal predictions. Thus, accurately predicting the brain age of unrepresentative data has been a major challenge for brain age prediction models. Our model showed great generalizability during training since the training dataset

was collected from various sources. The model was able to accurately predict brain age in the validation and test datasets, and minimal brain age bias was observed. However, when the model was applied to clinical 2D axial datasets, age bias correction had to be performed. This may be due to the innate differences between research grade 3D MRI scans and clinical 2D MRI scans. Although our preprocessing pipeline to convert research grade 3D MRI scans into the training dataset tried to imitate clinical 2D MRI scans, these innate differences may have led to regression dilution or attenuation bias. Without age bias correction, the model will overestimate the brain age of MRI scans from younger individuals and underestimate the brain age of MRI scans from older individuals. Therefore, an accurate age bias correction step is required for clinical application on 2D MRI scans. We showed that this is possible using only a few scans from cognitively unimpaired individuals.

There are some limitations of this study. First, further validation of the model using clinical 2D scans from various sources is required. Since significant brain age bias correction was required when applying the model to the SMC clinical axial 2D MRI scan dataset, slight adjustments to the learned linear bias correction coefficients may be necessary for scans from different sources. Furthermore, to use the model as a screening tool for measuring risk of neurodegenerative diseases, training the model to classify healthy subjects and patients could enhance its performance. When trained sorely to predict brain age, it does not necessarily learn the features of neurodegenerative diseases. While our model successfully predicted greater brain age gap in neurodegenerative disease cohorts compared to healthy subjects, training a model targeted specifically for a particular neurodegenerative disease could lead to the development of a specialized model for risk assessment for that disease.

In conclusion, we have demonstrated that it is possible to accurately predict brain age of clinical grade 2D axial T1 weighted MRI scans by training a 3D DenseNet based model with publicly available research grade 3D T1 weighted MRI scans. The same training pipeline can possibly be used MRI scans of other modalities. The high performance and fast inference

(a)   Corrected brain age gap by ADNI dataset subgroup

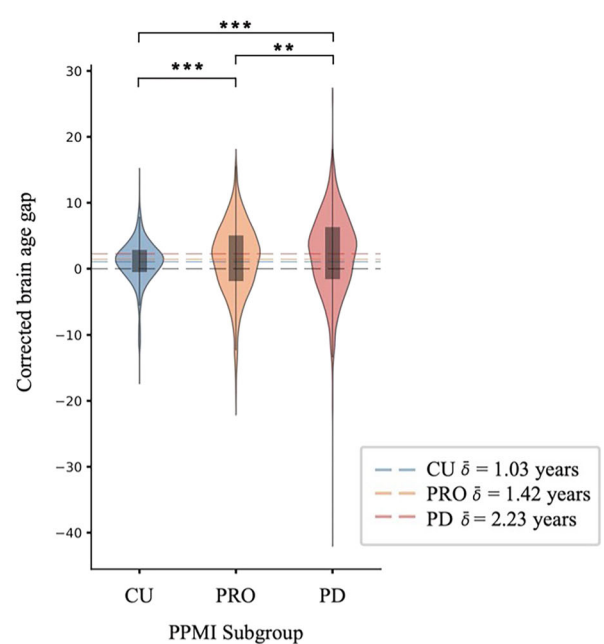(b)   Corrected brain age gap by PPMI dataset subgroup

Fig. 5 | Brain age prediction on neurodegenerative disease cohorts. a The ADNI dataset contains brain MRI scans of cognitively unimpaired (CU; $n = 6921$), mild cognitive impairment (MCI), and Alzheimer's disease (AD) cohorts. The cohorts showed significantly different mean brain age gap ($p < 0.001$, Welch's ANOVA; brain age gap: 0.57 vs 2.15 vs 2.47 years in CU vs MCI vs AD cohorts). Brain age gap of patients diagnosed with MCI or AD were significantly greater than CU subjects ($p < 0.001$, Pairwise Games–Howell test). Brain age gap was significantly greater in patients with AD compared to patients with MCI ($p < 0.05$, Pairwise Games–Howell test). **b** The

PPMI dataset contains brain MRI scans of cognitively unimpaired (CU), prodromal (PRO), and Parkinson's disease (PD) cohorts. The cohorts showed significantly different mean brain age gap ($p < 0.001$, Welch's ANOVA; brain age gap: 1.03 vs 1.42 vs 2.23 years in CU vs PRO vs PD cohorts). Brain age gap of patients in PRO or PD cohorts were significantly greater than CU subjects ($p < 0.001$, Pairwise Games–Howell test). Brain age gap was significantly greater in patients diagnosed with PD compared to patients in the PRO cohort ($p < 0.01$, Pairwise Games–Howell test). Dotted lines are drawn at the mean brain age gap values of each cohort.

speed of the model show potential for application in clinical situations, pending further validation with additional clinical 2D MRI scans.

Methods
Datasets
A total of 8681 research grade 3D T1-weighted MRI scans were acquired from 8495 cognitively unimpaired subjects from the Samsung Medical Center (SMC) and 24 publicly available datasets: Alzheimer's Disease Neuroimaging Initiative[25] (ADNI; http://adni.loni.usc.edu/), Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository[26,27], Calgary-Campinas-359 (CC-359) dataset[28], Dallas Lifespan Brain Study (DLBS), International Consortium for Brain Mapping[29] (ICBM; www.loni.usc.edu/ICBM), Information eXtraction from Images (IXI; http://brain-development.org/ixi-dataset/), Open Access Series of Imaging Studies-3 (OASIS-3)[30], Parkinson's Progression Markers Initiative (PPMI)[31], Southwest University Adult Lifespan Dataset[32] (SALD), and 15 datasets[33–47] from OpenNeuro[48] (Table 1). The description of each cohort and information about image protocols and scanners are described in Supplementary Information 1. Scans taken from a same subject were collected if the scans were taken more than 5 years apart. Otherwise, only the scan taken from the earliest timepoint was selected. Of the 8681 3D scans, 4821 scans were from female subjects and the mean age of the scans was 51.76 ($\pm$21.74) years with a range from 16.22 to 95.

The study protocol received approval from the Institutional Review Board of SMC, and all procedures were conducted in accordance with the approved guidelines. Written consent was obtained from each participant prior to their involvement in the study. To test the model's performance on clinical grade axial 2D T1-weighted MRI scans, 175 scans from cognitively unimpaired subjects and 199 scans from patients diagnosed with AD were also collected from SMC (Table 2). The mean age of the SMC 2D cognitively

unimpaired cohort was 65.06 ($\pm$8.67) with a range of 41 to 85, and 111 out of the 175 subjects were female. The mean age of the SMC 2D AD cohort was 72.83 ($\pm$8.64) with a range of 47 to 89, and 133 out of the 199 subjects were female. The cognitively unimpaired subjects in the test dataset do not overlap with the subjects in the SMC cohort in the training dataset. Further details of the SMC 2D dataset are described in Supplementary Information 1.

Image processing
All scans were converted into NIfTI format then skull stripped[49]. Since the model was targeted for clinical grade axial 2D MRI scans, the 3D scans used for training were first sliced to generate synthetic axial 2D scans. The axial gap between these slices was set to the smallest integer multiple of the axial gap of the original 3D scan that was greater than or equal to 7 mm. The synthetic 2D scans were then interpolated back into common voxel space of 2 mm x 2 mm x 2 mm. To normalize for the different MR signal ranges from different devices, the intensity of each scan was normalized by subtracting the mean and dividing by standard deviation. Then the background values were subtracted to every voxel to set background intensity to 0. The mean and standard deviation were computed using voxels between the 1st and 99th percentiles after excluding background voxels. To remove outliers, all voxels with intensity above the 99.9th percentile was replaced with the 99.9th percentile value. Finally, the scans were cropped in the center or padded into a common size (sagittal x coronal x axial 160 mm x 240 mm x 80 mm). No spatial registration or bias field correction was applied. To filter out corrupted scans containing only parts of the brain, scans with more than 92% background voxels were removed. The final training dataset composed of 69,598 3D scans interpolated from 2D axial T1-weighted MRI scans created from slicing 8629 research grade 3D T1-weighted MRI scans.

Table 1 | Training dataset composed of research grade 3D T1-weighted MRI scans from cognitively unimpaired individuals

Cohort	*N* (scans)	Female/ Male	Age (μ ± σ)	Age range	Reference
Scans from SMC					
SMC 3D scans	1037	630/407	69.87 ± 10.34	21–94	
OpenNeuro[a] Datasets (Accession numbers)					
ds000222	79	41/38	44.43 ± 19.99	21–73	[Reference links]
ds000240	63	35/28	48.98 ± 24.21	19 - 84	
ds002382	60	36/24	48.44 ± 23.76	19–81	
ds003097	928	483/445	22.85 ± 1.71	19–26	
ds003592	299	168/131	40.66 ± 22.86	18–83	
ds003763	62	50/12	45.81 ± 13.84	18–63	
ds003826	136	87/49	24.33 ± 3.72	18–35	
ds003831	77	52/25	38.21 ± 14.10	18–64	
ds004173	148	95/53	30.01 ± 12.72	18.15 - 74.8	
ds004215	154	101/53	33.93 ± 12.67	18–72	
ds004285	78	62/16	42.03 ± 24.02	19–81	
ds004636	109	71/38	23.85 ± 5.54	18–45	
ds004648	88	0/88	29.63 ± 5.35	21–40	
ds004711	189	99/90	45.05 ± 19.22	16.22 - 81.48	
ds004746	397	234/163	36.07 ± 3.42	27–77	
Other public datasets					
ADNI	983	551/432	73.32 ± 7.43	51–95	[Reference links]
Cam-CAN	653	330/323	54.31 ± 18.57	18–88	
CC359	359	183/176	53.47 ± 7.83	29–80	
DLBS	315	198/117	54.62 ± 20.06	20.57–89.11	
ICBM	640	291/349	30.63 ± 12.16	18–80	
IXI	229	123/106	74.31 ± 6.41	56–89	
OASIS-3	910	521/389	67.62 ± 9.11	42.5–93.63	
PPMI	196	73/123	61.17 ± 11.64	30.6–82.7	
SALD	492	307/185	45.18 ± 17.43	19–80	
Total	**8681**	**4821/ 3860**	**51.76 ± 21.74**	**16.22–95**	

[a]https://openneuro.org/.

SMC Samsung Medical Center, *ADNI* Alzheimer's Disease Neuroimaging Initiative, *Cam-CAN* Cambridge Centre for Ageing and Neuroscience, *CC359* Calgary-Campinas dataset, *DLBS* Dallas Lifespan Brain Study, *ICBM* International Consortium for Brain Mapping, *IXI* Information eXtraction from Images, *OASIS* Open Access Series of Imaging Studies, *PPMI* Parkinson's Progression Markers Initiative, *SALD* Southwest University Adult Lifespan Dataset.

For the clinical grade axial 2D T1-weighted MRI scans, the same preprocessing was applied excluding the axial slicing procedure. Open access tools were used for preprocessing: dcm2niix[50] was used to convert scans from DICOM to NIfTI format; HD-BET[49], a brain extraction tool built with open access python libraries was used for skull stripping; MONAI[51], an open-source project built on PyTorch was used to interpolate 2D scans into 3D scans; Python libraries NiBabel, and Numpy were used for all other preprocessing procedures.

Brain age prediction model
A 3D DenseNet-169 based model[52] was used for brain age prediction. DenseNet concatenates learned features from all preceding layers to each layer to prevent the vanishing gradient problem and improve information flow. It was previously used to successfully predict brain age from research grade 3D T1-weighted MRI scans[53], and clinical grade axial T2-weighted or diffusion-weighted scans. For application in this study, the last transitional layer was removed due to smaller input size (80x120x40 compared to 224 x 224 x 224 of the original model). The softmax layer within the classification layer was replaced with a fully connected layer where the output is a single scalar for the predicted brain age. LeakyReLU with negative slope 0.01 was used for the activation layers instead of ReLU. All other configurations including convolution kernel sizes, stride, dense block configurations, and the growth rate was set to values in the original paper. An illustration of the model architecture is shown in Fig. 1a. Our final model contained 20.4 million parameters. 5-fold cross validation with training 80%, validation 10%, test 10% splits were conducted where the validation dataset was used to select the best model. To prevent data leakage, the dataset was split at the research grade 3D T1-weighted MRI scan level.

He initialization using the Kaiming normal distribution was used for weight initialization[54]. The Adam optimizer with β1 = 0.9 and β2 = 0.9 and weight decay λ = 1e−4 was used to minimize the mean squared error loss between predicted and chronological age. Initial learning rate was set to 1e−4 and was scheduled to be reduced by a factor of 2 when no improvement in validation loss is observed for 5 epochs. To prevent overfitting, three data augmentation methods were applied during model training. For every training epoch, the training batch is (1) randomly shifted by up to 2 voxels in the axial direction, and up to 10 voxels in the sagittal and coronal directions, (2) mirrored with respect to the sagittal plane with 50% probability, and (3) gamma correction with $\gamma \sim U(0.7, 1.3)$ and intensity was multiplied with by factor of $\alpha \sim U(0.7, 1.3)$ for data augmentation. Mini batch size was set to 32, and total epoch number was set to 70. We observed stability in the model after an iteration of 70 epochs. Training time per epoch was approximately 200 s, and the entire training time including validation and testing was approximately 4.2 h when using one NVIDIA L40S GPU for training. Model implementation, training, and testing were done using PyTorch[55]. Python scripts for model implementation and inference will be made available under GNU General Public License version 3 upon request to the corresponding author.

Applying model on clinical grade axial 2D T1-weighted MRI scans
The model was applied on clinical grade axial 2D T1-weighted MRI scans collected from SMC (Table 2). Brain age gap, defined as the difference between predicted age (*y*) and chronological age (x), is known to be overestimated in younger individuals and underestimated in older individuals due to regression dilution[56]. Although our model showed minimal brain age bias during training, brain age bias was apparent when applied on clinical grade axial 2D scans. This may be due to the innate differences of clinical grade axial 2D images and research grade 3D images. Therefore, linear bias correction using scans from oldest and youngest 5% cognitively unimpaired subjects was added after model prediction. A linear regression $y = ax + b$ is first fit using these subjects, then the corrected prediction y' was calculated as $y' = y + [x - (ax + b)]$ for all subjects. The coefficients *a* and *b* learned from cognitively unimpaired subjects are used for age bias correction for age prediction results of the MRI scans from patients with Alzheimer's disease. Brain age gap of cognitively unimpaired subjects was compared to patients with Alzheimer's disease using the two sample Z-test.

Guided backpropagation is a gradient based method to visualize how deep networks predict. It gives a detailed, pixel-space heatmap of importance on model prediction[57]. Guided backpropagation was used to visualize important regions of the brain on predicting brain age of clinical grade 2D MRI scans from cognitively unimpaired subjects and patients with Alzheimer's disease.

Brain age prediction on neurodegenerative disease cohorts
Multisite large-scale studies often provide datasets including MRI scans from different stages of a neurodegenerative disease (Table 2). The Alzheimer's Disease Neuroimaging Initiative (ADNI) provides various types of MRI scans from cognitively unimpaired subjects, mild cognitive

Table 2 | Test dataset composed of clinical grade axial 2D T1-weighted MRI scans from SMC and 3D MRI scans of neurodegenerative disease cohorts from public datasets

Cohort	*N* (scans)	Female/Male	Age (μ ± σ)	Age range
SMC 2D scans				
Cognitively unimpaired	175	111/64	65.06 ± 8.67	41–85
Alzheimer's disease	199	133/66	72.83 ± 8.64	47–89
ADNI				
Alzheimer's disease	428	194/234	75.15 ± 7.89	55–91
Mild cognitive impairment	660	271/389	74.15 ± 7.77	55–91
PPMI				
Parkinson's disease	661	254/407	62.91 ± 9.53	34–84
Prodromal phase	514	254/260	65.39 ± 6.96	42–86

SMC Samsung Medical Center, *ADNI* Alzheimer's Disease Neuroimaging Initiative, *PPMI* Parkinson's Progression Markers Initiative.

impairment, to Alzheimer's disease. The Parkinson's Progression Markers Initiative (PPMI) dataset includes scans across all stages of Parkinson's disease. The same linear bias correction method as previously described was applied for each study to ensure correct prediction of brain age gap. The linear bias coefficients learned from cognitively unimpaired subjects are used for neurodegenerative disease cohorts from the same study. Given the model was originally tasked to predict brain age of cognitively unimpaired subjects and thus learn features of normal brain aging, the model does not necessarily learn features of neurodegenerative diseases. However, if greater brain age gap is observed in subjects with neurodegenerative disease, it could potentially highlight the role of the brain age prediction model in accessing risk for neurodegenerative diseases.

The brain age gap of subjects from different stages of the disease was compared with the brain age gap of cognitively unimpaired subjects using Welch's analysis of variance (ANOVA) with pairwise Games–Howell post hoc test. All statistical analysis were performed using the open-source Pingouin Python package[58].

Data availability
Data from publicly available research grade 3D brain MRI datasets are available from their respective sources. The datasets from SMC used in this study are not publicly available.

Code availability
The codes used for this study are available upon reasonable request to the authors.

References
1. Yankner, B. A., Lu, T. & Loerch, P. The aging brain. *Annu. Rev. Pathol.: Mechanisms Dis.* **3**, 41–66 (2008).
2. Thal, D. R., Del Tredici, K. & Braak, H. Neurodegeneration in normal brain aging and disease. *Science of Aging Knowledge Environment* **2004**, pe26-pe26. https://doi.org/10.1126/sageke.2004.23.pe26. (2004).
3. Mungas, D. et al. Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. *Neurology* **59**, 867–873 (2002).
4. Yan, S. et al. The influence of accelerated brain aging on coactivation pattern dynamics in Parkinson's disease. *J. Neurosci. Res.* **102**, e25357 (2024).
5. Horvath, S. et al. Huntington's disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. *Aging* **8**, 1485–1512 (2016).
6. Dugger, B. N. & Dickson, D. W. Pathology of neurodegenerative diseases. *Cold Spring Harb. Perspect. Biol.* **9**, a028035 (2017).
7. Gonzalez-Ortiz, F. et al. Plasma phospho-tau in Alzheimer's disease: towards diagnostic and therapeutic trial applications. *Mol. Neurodegener.* **18**, 18 (2023).
8. Simuni, T. et al. A biological definition of neuronal α-synuclein disease: towards an integrated staging system for research. *Lancet Neurol.* **23**, 178–190 (2024).
9. Sperling, R., Mormino, E. & Johnson, K. The evolution of preclinical Alzheimer's disease: implications for prevention trials. *Neuron* **84**, 608–622 (2014).
10. Hardy, J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci.* **20**, 154–159 (1997).
11. Rusinek, H. et al. Regional brain atrophy rate predicts future cognitive decline: 6-year longitudinal MR imaging study of normal aging. *Radiology* **229**, 691–696 (2003).
12. Lockhart, S. N. & DeCarli, C. Structural imaging measures of brain aging. *Neuropsychol. Rev.* **24**, 271–289 (2014).
13. Courchesne, E. et al. Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* **216**, 672–682 (2000).
14. Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A. & Smith, S. M. Accurate brain age prediction with lightweight deep neural networks. *Med. Image Anal.* **68**, 101871 (2021).
15. Franke, K. & Gaser, C. Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained?. *Front. Neurol.* **10**, 454252 (2019).
16. Soumya Kumari, L. K. & Sundarrajan, R. A review on brain age prediction models. *Brain Res.* **1823**, 148668 (2024).
17. Franke, K. & Gaser, C. Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's disease. *GeroPsych* **25**, 235–245 (2012).
18. Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N. & Sauer, H. BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer's disease. *PLoS ONE* **8**, e67346 (2013).
19. Wood, D. A. et al. Accurate brain-age models for routine clinical MRI examinations. *NeuroImage* **249**, 118871 (2022).
20. Valdes-Hernandez, P. A., Laffitte Nodarse, C., Cole, J. H. & Cruz-Almeida, Y. Feasibility of brain age predictions from clinical T1-weighted MRIs. *Brain Res. Bull.* **205**, 110811 (2023).
21. Hong, J. et al. Brain age prediction of children using routine brain MR images via deep learning. *Front. Neurol.* **11**. https://doi.org/10.3389/fneur.2020.584682 (2020).
22. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep inside convolutional networks: visualising image classification models and saliency maps. In *2nd International Conference on Learning Representations*. (ed. {Bengio, Y. & LeCun, Y.) http://arxiv.org/abs/1312.6034 (2014).
23. Bashyam, V. M. et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain* **143**, 2312–2324 (2020).

24. Valdes-Hernandez, P. A., Laffitte Nodarse, C., Peraza, J. A., Cole, J. H. & Cruz-Almeida, Y. Toward MR protocol-agnostic, unbiased brain age predicted from clinical-grade MRIs. *Sci. Rep.* **13**, 19570 (2023).

25. Petersen, R. C. et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology* **74**, 201–209 (2010).

26. Taylor, J. R. et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage* **144**, 262–269 (2017).

27. Shafto, M. A. et al. The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurol.* **14**, 204 (2014).

28. Souza, R. et al. An open, multi-vendor, multi-field-strength brain MR dataset and analysis of publicly available skull stripping methods agreement. *NeuroImage* **170**, 482–494 (2018).

29. Mazziotta, J. et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* **356**, 1293–1322 (2001).

30. LaMontagne, P. J. et al. OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. *medRxiv* 2019.2012.2013.19014902. https://doi.org/10.1101/2019.12.13.19014902 (2019).

31. Marek, K. et al. The Parkinson's progression markers initiative (PPMI) —establishing a PD biomarker cohort. *Ann. Clin. Transl. Neurol.* **5**, 1460–1477 (2018).

32. Wei, D. et al. Structural and functional brain scans from the cross-sectional Southwest University adult lifespan dataset. *Sci. Data* **5**, 180134 (2018).

33. Miranda-Angulo, A. L. et al. OpenNeuro Dataset ds004648 (Sympathovagal quotient and functional connectivity of control networks are related to gut Ruminococcaceae abundance in healthy men). OpenNeuro (2023).

34. Zareba, M. R. et al. OpenNeuro Dataset ds003826 (Structural (t1) images of 136 young healthy adults; study of effects of chronotype, sleep quality and daytime sleepiness on brain structure). OpenNeuro (2022).

35. FitzGerald, T. H. B., Haemmerer, D., Friston, K. J., Li, S.-C. & Dolan, R. J. OpenNeuro Dataset ds000222 (Sequential Inference VBM). OpenNeuro (2019).

36. Galiano, A. et al. OpenNeuro Dataset ds000240 (Resting State Perfusion in Healthy Aging). OpenNeuro, (2021).

37. Botvinik-Nezer, R., Petre, B., Ceko, M., Friedman, N. & Wager, T. OpenNeuro Dataset ds004746 (Paingen_placebo). OpenNeuro (2023).

38. Nugent, A. C. et al. OpenNeuro Dataset ds004215 (The NIMH Healthy Research Volunteer Dataset). OpenNeuro (2023).

39. Spreng, R. N. et al. OpenNeuro Dataset ds003592 (Neurocognitive aging data release with behavioral, structural, and multi-echo functional MRI measures). OpenNeuro (2022).

40. Nárai, Á. et al. OpenNeuro Dataset ds004173 (Movement-related artefacts (MR-ART) dataset). OpenNeuro (2022).

41. Rogers, C. S., Jones, M. S., McConkey, S. & Peelle, J. E. OpenNeuro Dataset ds004285 (Listening task). OpenNeuro (2022).

42. Dobrushina, O. R. et al. OpenNeuro Dataset ds003763 (Interoception during aging: The heartbeat detection task). OpenNeuro (2022).

43. Bissett, P. G. et al. OpenNeuro Dataset ds004636 (Cognitive tasks, anatomical MRI, and functional MRI data evaluating the construct of self-regulation). OpenNeuro (2024).

44. Fialkowski, K., Messias, I. & Bush, K. A. OpenNeuro Dataset ds003831 (Cognitive Control Theoretic Mechanisms of Real-time fMRI-Guided Neuromodulation (CTM)). OpenNeuro (2021).

45. Snoek, L. et al. OpenNeuro Dataset ds003097 (AOMIC-ID1000). OpenNeuro (2021).

46. Tisdall, L. & Mata, R. OpenNeuro Dataset ds004711 (AgeRisk). OpenNeuro (2023).

47. Rogers, C. S. et al. OpenNeuro Dataset ds002382 (Age-related differences in auditory cortex activity during spoken word recognition). OpenNeuro (2022).

48. Markiewicz, C. J. et al. The OpenNeuro resource for sharing of neuroscience data. *eLife* **10**, e71774 (2021).

49. Isensee, F. et al. Automated brain extraction of multisequence MRI using artificial neural networks. *Hum. Brain Mapp.* **40**, 4952–4964 (2019).

50. Li, X., Morgan, P. S., Ashburner, J., Smith, J. & Rorden, C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J. Neurosci. Methods* **264**, 47–56 (2016).

51. Cardoso, M. J. et al. MONAI: An open-source framework for deep learning in healthcare. Preprint at https://arxiv.org/abs/2211.02701 (2022).

52. Huang, G., Liu, Z., van der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017).

53. Lee, J. et al. Deep learning-based brain age prediction in normal aging and dementia. *Nat. Aging* **2**, 412–424 (2022).

54. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: surpassing human-level performance on Image Net Classification. In *IEEE International Conference on Computer Vision.* 1026–1034, https://doi.org/10.1109/ICCV.2015.123 (IEEE, 2015)

55. Paszke, A. et al. PyTorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems*. 8024–8035, https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html (2019)

56. de Lange, A. G. et al. Mind the gap: performance metric evaluation in brain-age prediction. *Hum. Brain Mapp.* **43**, 3113–3129 (2022).

57. Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. A. Striving for simplicity: the all convolutional net. In 3*rd International Conference on Learning Representations*. (ed. {Bengio, Y. & LeCun, Y.). http://arxiv.org/abs/1412.6806 (2015)

58. Vallat, R. Pingouin: statistics in Python. *J. Open Source Softw.* **3**, 1026 (2018).

Acknowledgements

Author contributions

Conceptualization, H.K. and K.K.; methodology, H.K.; investigation, H.K., S.P., S.W.S., and K.K., writing—original draft, H.K. and K.K.; writing—review & editing, H.K., S.P., S.W.S., D.L.N., H.J., J.P.K.,H.J.K.,S.H.K., and K.K.

Competing interests

H. Kim was funded by BeauBrain Healthcare Inc. S. Park and K. Kwak were employed by BeauBrain Healthcare Inc. The remaining authors state that the research was done without any commercial or financial conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41514-025-00260-x.

Correspondence and requests for materials should be addressed to Kichang Kwak.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.