

Deep learning-derived age of hippocampus-centred regions is influenced by *APOE* genotype and modifiable risk factors

Chao Dong ^{a,d} *, Yuangang Pan ^b, Anbupalam Thalamuthu ^a, Jiyang Jiang ^a, Jing Du ^a, Karen A. Mather ^a, Perminder S. Sachdev ^{a,c}, Wei Wen ^{a,c} *

^a Centre for Healthy Brain Ageing (CHeBA), Discipline of Psychiatry and Mental Health, School of Clinical Medicine, UNSW, Sydney, Australia.

^b Centre for Frontier AI Research (CFAR), A*STAR, Singapore

^c Neuropsychiatric Institute (NPI), Prince of Wales Hospital, Randwick, New South Wales, 2031, Australia.

^d Department of Psychology, University of Cambridge, Cambridge, CB2 3EB, United Kingdom.

Corresponding author:

Chao Dong E-mail: cd888@cam.ac.uk

Wei Wen E-mail: w.wen@unsw.edu.au

Abstract

Brain age has been widely investigated by using the whole brain image. However, the age of some specific brain regions, such as those related to the hippocampus, remains underexplored. This study developed age prediction models for left and right hippocampus-centred regions of interest (hippocampus ROI) using three-dimensional convolutional neural networks (3D-CNN) based on MRI scans from 31,370 healthy participants in the UK Biobank. The hippocampus ROI age (HA) gap was calculated by subtracting chronological age from predicted HA. Additionally, the longitudinal change rate of the HA gap was estimated in 3,893 participants with imaging data at two time points over an average follow-up of 2.63 years. The models achieved state-of-the-art performance (mean absolute error (MAE): 2.47 – 2.84 years). Cross-sectional analysis revealed that *APOE* ϵ 4 homozygotes had a greater HA gap compared to *APOE* ϵ 4 non-carriers. Participants with hypertension, diabetes, heavy alcohol consumption, or smoking also exhibited larger HA gap. Transfer learning applied to an independent dataset confirmed similar trends in some variables, though findings were not statistically significant. Interestingly, longitudinal analysis showed that *APOE* ϵ 4 homozygotes had a higher annual change rate in the left HA gap compared to *APOE* ϵ 2 homozygotes. Occlusion analysis saliency maps indicated that regions around the hippocampus, including the thalamus, pallidum, nearby cerebral cortex, and white matter, significantly contributed to the age prediction. The left HA gap emerges as a potential biomarker linked to the *APOE* genotype and an indicator of health.

Keywords: hippocampus, brain age, deep learning, convolutional neural network, *APOE*.

1 Introduction

Medial temporal lobe (MTL), which includes the hippocampus, entorhinal, perirhinal, and parahippocampal cortices, is altered early and severely in Alzheimer's disease (AD) (Chauveau et al., 2021; de Flores et al., 2020) and is also one of risk factors for vascular dementia (Pendlebury and Rothwell, 2009). AD is characterised by initial regional brain atrophy in the hippocampus and medial temporal regions, followed by subsequent spread to other cortical areas (Chapleau et al., 2016). The temporal lobes and hippocampus are high-risk regions for AD classification (Qiu et al., 2020; Wang et al., 2023), or related dementia (Martin et al., 2023). Therefore, an in-depth exploration of hippocampus-centred brain regions could provide great neuroscientific value.

The Apolipoprotein E (*APOE*) gene, which is polymorphic with three common isoforms: $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$, is encoded by two single nucleotide polymorphisms. *APOE* $\epsilon 4$ is the strongest genetic risk factor for late-onset AD, *APOE* $\epsilon 3$ is neutral, and *APOE* $\epsilon 2$ is protective (Huang et al., 2017; Serrano-Pozo et al., 2021). Individuals with one copy of *APOE* $\epsilon 4$ (*APOE* $\epsilon 4$ heterozygotes) have an increased risk of developing AD, and those with two copies (*APOE* $\epsilon 4$ homozygous) have an even higher risk (Gharbi-Meliani et al., 2021). *APOE* $\epsilon 4$ has been associated with brain atrophy (Cacciaglia et al., 2018), white matter lesions, memory decline (Koizumi et al., 2018), and MTL subregions (de Flores et al., 2023). In addition to AD, *APOE* $\epsilon 4$ has also been associated with vascular dementia (Skrobot et al., 2016), Parkinson's disease and Lewy body dementia (Rongve et al., 2019; Saeed et al., 2021).

In longitudinal studies involving patients with AD, the rate of hippocampal atrophy is influenced by both *APOE* genotypes and the severity of the disease, with *APOE* $\epsilon 4$ carriers showing greater rates of atrophy (Mori et al., 2002). *APOE* $\epsilon 4/\epsilon 4$ homozygotes also exhibited accelerated atrophy in hippocampus (Abushakra et al., 2020). Previous study suggests that

APOE ϵ 4 carriers may exhibit a greater rate of hippocampal atrophy compared to their counterparts, even without an AD diagnosis (Moffat et al., 2000). Furthermore, *APOE* ϵ 4 carriers also showed faster rates of structural loss in MTL (Donix et al., 2010; Mishra et al., 2018). Additionally, accumulating evidence showed that hypertension, diabetes, excessive alcohol consumption, smoking, and physical inactivity are high risk factors for dementia or adverse brain health outcomes (Chudasama et al., 2020; Franz et al., 2021; Gottesman and Seshadri, 2022; Livingston et al., 2020). These modifiable risk factors can influence hippocampal atrophy (Fotuhi et al., 2012) and multiple body age (Tian et al., 2023).

Brain age, or brain-predicted age, can be estimated using neuroimaging data and may differ from chronological age, resulting in a brain age gap (Cole and Franke, 2017). By summarizing complex neuroimaging features into a simple, interpretable summary metric, brain age gap may reflect a comprehensive indicator of personalized brain health, providing prognostic and predictive values (Cole and Franke, 2017; Millar et al., 2023). For example, a greater brain age gap is associated with AD (Gaser et al., 2013) and mortality (Cole et al., 2018). Grey matter density around hippocampus and amygdala has been identified as a key factor influencing age predictions (Wang et al., 2019). A recent study which used the hippocampal-centred regions to predict brain age found that brain age gap might be a biomarker to help AD and mild cognitive impairment (MCI) diagnosis (Poloni and Ferrari, 2022). Cumulatively, this evidence highlights the critical importance of the hippocampus and its neighbouring structures. Additionally, a study demonstrated that the 3D bounding boxes of the ROI surrounding the hippocampus could act as a promising representation for dementia diagnosis and prediction in deep learning models, as it contains information around hippocampus structure and there is no requirement of the accurate hippocampus segmentation (Sarasua et al., 2022). However, most studies on brain age are cross-sectional, and longitudinal studies are needed to explore the change rate of the predicted age.

In this study, we conceptualized the age of “hippocampus-centred regions of interest (hippocampus ROI age)” as a measure representing region-specific biological age.

Hippocampus ROI age (HA) can be regarded as the proxy for the biological age of the hippocampus ROI which are 3D bounding boxes centred around hippocampus. They do not only contain structure information of hippocampus but also their nearby brain structures.

Here, (1) we aimed to estimate the deep learning-based HA from 31,370 healthy participants who aged 44 to 83 (five-fold cross-validation) and explore whether the presence of the *APOE* $\epsilon 4$ allele and modifiable risk factors are associated with HA gap. (2) We aimed to evaluate the longitudinal change rate of the HA gap in different *APOE* genotype groups. (3) We also aimed to identify which areas contribute more to the HA prediction by applying occlusion analysis. As the hippocampus ROI are the first to be affected by AD, and *APOE* $\epsilon 4$ is the major genetic risk factor for AD, we hypothesised that HA gap might be associated with *APOE* genotype, with *APOE* $\epsilon 4/\epsilon 4$ carriers exhibiting a greater HA gap.

2 Methods

2.1 Participants

UK Biobank. The UK Biobank is a large-scale dataset containing cross-modality neuroimaging data. It is a population-based study which consists of over 500,000 participants aged between 40-70 years at study entry (Sudlow et al., 2015). Written consent was acquired from all participants and ethics approval was provided by the National Health Service National Research Ethics Service (11/NW/0382). In the current study, after excluding participants with severe self-reported brain related disorders (Table S1), 31,370 healthy participants with one time point image were included in the final analysis. Additionally, there were 3,893 participants who have two-time points image. All the 31,370 healthy participants with one time point image were randomly split into five equally sized folds, with 60% for

training, 20% for validation, and 20% for test set. Model performance in the training set was assessed using 5-fold cross validation.

Sydney Memory and Ageing Study (MAS). MAS is a cohort of 1,037 community dwelling adults aged 70–90 years (Sachdev et al., 2010). Ethics approval was obtained from the Human Research Ethics Committees of the University of New South Wales and the South Eastern Sydney Local Health District. For the current study, 534 participants with MRI scans from baseline were used.

2.2 Non-imaging data

2.2.1 APOE genotype

For UK Biobank, DNA extracted from the blood samples of the initial cohort (~ 50,000 participants), underwent genotyping within the UK Biobank utilising the Affymetrix UK BiLEVE Axiom array. For MAS, DNA was extracted from peripheral blood leukocytes or saliva samples using standard procedures at Genetics Repositories Australia (www.powmri.edu.au/GRA.htm) (Sachdev et al., 2010). Two *APOE* coding single nucleotide polymorphisms (SNPs), namely rs7412 and rs429358, were retrieved from the genotyped data to determine the *APOE* genotype. There are six *APOE* genotypes: $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 3$, $\epsilon 2/\epsilon 4$, $\epsilon 3/\epsilon 3$, $\epsilon 3/\epsilon 4$, and $\epsilon 4/\epsilon 4$. *APOE* $\epsilon 4$ allele carriers include $\epsilon 2/\epsilon 4$, $\epsilon 3/\epsilon 4$, and $\epsilon 4/\epsilon 4$. Non-carriers include $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 3$, and $\epsilon 3/\epsilon 3$.

2.2.2 Cognitive measures

UKB: Cognitive assessments were administered on a fully automated touchscreen questionnaire (Fawns-Ritchie and Deary, 2020; Sudlow et al., 2015). Seven tests from the UK Biobank battery of cognitive tests were selected for the current study to represent three cognitive domains: “Reaction Time”, “Trail Making Test A”, and “Symbol Digit Substitution” formed the Processing Speed domain; “Numeric Memory” and “Pairs

Matching” contributed to the Memory domain; and “Trail Making Test B” and “Fluid Intelligence” formed the Executive Function domain. All test scores were first z-transformed and then averaged to form domain scores. Therefore, processing speed, executive function, and memory were used in the current study. Global cognition was estimated by averaging the domain scores and z-transform.

MAS: A battery of interview-based cognitive tests was administered by trained psychologists to examine cognitive domains (Sachdev et al., 2010). Cognition domains include: attention processing speed (including “Digit Symbol-Coding” and “Trail Making Test A”), memory (including “Logical Memory Story A delayed recall”, “Rey Auditory Verbal Learning Test”, and “Benton Visual Retention Test recognition”), verbal memory (including “Logical Memory Story A delayed recall” and “Rey Auditory Verbal Learning Test”), language (including “Boston Naming Test – 30 items” and “Semantic Fluency (Animals)”), visuospatial (“Block Design”), executive function (“Controlled Oral Word Association Test (FAS)” and “Trail Making Test B”). All the details can be found in previous studies (Sachdev et al., 2010).

2.2.3 Modifiable risk factors

In the current study, we focused on the following modifiable risk factors from UK Biobank: (1) alcohol intake frequency [Data-Field 1558], (2) smoking status [Data-Field 20116], (3) physical activity [Data-Field 22040], (4) hypertension [Data-Field 4079 and Data-Field 4080], and (5) diabetes [Data-Field 2443]. Specially, alcohol intake frequency was assessed based on questionnaire, and it was classified as: daily or almost daily, three or four times a week, once or twice a week, one to three times a month, special occasions only, or never. Smoking status was assessed by a questionnaire. Individuals were asked whether they were never, previous, or current smokers. Physical activity was measured by metabolic equivalent

of task (MET) calculation (Doherty et al., 2017; Le Goallec et al., 2023). Data was self-reported and classified as: no physical activity, medium ($>0, <7.5$ MET hours/week), and medium ($>0, <7.5$ MET hours/week). Participants with hypertension (high blood pressure) were defined with blood pressure over 140/90 mmHg (<https://www.who.int/news-room/fact-sheets/detail/hypertension>). Diabetes was defined according to the doctor's diagnosis. Additionally, we also considered body mass index (BMI) (kg/m^2), which was calculated based on the measured weight and height: normal (18.5-24.9 kg/m^2), underweight (<18.5 kg/m^2), overweight (25.0-29.9 kg/m^2), obese (> 30 kg/m^2). In MAS, smoking status, hypertension, and diabetes are defined in the same way as in UK Biobank. Alcohol consumption was classified as never, previous, or current drinkers.

2.3 Image data and processing

UK Biobank structural MRI scans were acquired on three 3T Siemens Skyra MRI scanners. The key parameters for MRI imaging were (a) T1-weighted MRI: TI = 880 ms, TR = 2 000 ms, resolution = $1.0 \times 1.0 \times 1.0$ mm, matrix size = $208 \times 256 \times 256$; (b) T2-weighted FLAIR: TI = 1 800 ms, TR = 5 000 ms, resolution = $1.05 \times 1.0 \times 1.0$ mm, matrix size = $192 \times 256 \times 256$. The full protocol is provided at <http://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=2367>. FreeSurfer (version 7.1.0) was used to model the cortical surface. For MAS cohort, we used T1-weighted and FLAIR scans and the key parameters for MRI imaging were: (a) T1-weighted MRI: TR = 6.39 ms, TE = 2.9 ms, flip angle = 8° , matrix size = 256×256 , FOV (field of view) = $256 \times 256 \times 190$, and slice thickness = 1 mm with no gap in between, yielding $1 \times 1 \times 1$ mm^3 isotropic voxels; (b) T2-weighted FLAIR: TR = 10000 ms, TE = 110 ms, TI = 2800 ms, matrix size = 512×512 , slice thickness = 3.5 mm without gap, and in plane resolution = 0.488×0.488 mm.

We used both the T1 and T2 FLAIR images as inputs to the FreeSurfer modelling (or only the T1 when the T2 was not available). The FreeSurfer output image “nu.mgz” was used for the extraction of hippocampus-centred regions in the current study, with a resolution of $1 \times 1 \times 1$ mm and inhomogeneity correction applied. The process for defining hippocampus-centred regions involved the following steps:

- (1) The gravity centre of each participant's left and right hippocampus was identified, and then the lower, upper, left, right, front, and back dimensions were extended by 32 voxels based on the gravity centre, resulting in hippocampus-centered regions of interest (ROIs) measuring $64 \times 64 \times 64$ voxels.
- (2) The reference brain in a standard space underwent the same method to obtain the standard space hippocampus ROI.
- (3) The hippocampus ROI in standard space served as a reference for registering each participant's hippocampus ROI, obtaining the registration transformation matrix.
- (4) The transformation matrix was applied to register the entire brain image, followed by repeating step (1) to acquire the final hippocampus ROI with a size of $64 \times 64 \times 64$ voxels.

By following these procedures, we achieved the consistency of each participant's left and right hippocampus ROI in the same space.

2.4 3D-CNN architecture

The architecture follows the Simple Fully Convolutional Network (SFCN) architectures (Peng et al., 2021). The architecture consisted of six 3D convolutional layers with varying channel numbers [32, 64, 128, 256, 256, 64]. Each convolutional layer has a kernel size of 3,

padding of 1, and a stride of 2. Following the convolutional layers is a two-layer Multilayer Perceptron (MLP). The input data size was $64 \times 64 \times 64$, and the output was the predicted hippocampus ROI age in years. Specifically, each of the first five 3D convolutional layer was followed by a batch normalisation layer, a max pooling layer, and a Rectified Linear Unit (ReLU) activation layer. The sixth 3D convolutional layer was followed solely by a ReLU activation layer, without a max pooling layer. The dimension of the output after all convolutional layers was 512. Subsequently, the output from the convolutional layers was flattened to be fed into two MLP layers. The first fully connected (linear) layer produced 100 output features, and the second produced a single scalar value, representing the final predicted age. During training, a learning rate of 1×10^{-3} was employed, and the batch size was set to 128.

Predicted ages often face challenges such as underfitting caused by regression dilution and a non-Gaussian age distribution. This implies that older participants might be estimated with a younger brain age, while younger participants might be estimated with an older brain age. Bias correction is an essential postprocessing technique in most brain-age prediction studies, and age correction was performed using method described in Smith et al. study (Smith et al., 2019). We fitted a linear regression $y_2 = a \times y_1 + b$ (y_1 : chronological age, y_2 : predicted age) to the validation set, and then applied the learned coefficients (a , b) to the test set. Therefore, the corrected brain age can be calculated by $(y_2 - b) / a$. All further analyses used the age-corrected gap. To avoid inflating estimates of prediction accuracy (Butler et al., 2021), the uncorrected predicted age values were applied for model performance evaluation. Model performance was evaluated by mean absolute error (MAE) and person's correlation between chronological age and predicted age.

Additionally, we also applied the whole brain image (FreeSurfer output image “nu.mgz”) to the 3D-CNN architecture to predict the whole brain age. Details can be found in supplementary materials.

2.5 Statistical analysis

Cross-sectional analysis

All statistical analyses were conducted in R 4.0.2. Participants’ predicted age was merged across the 5 folds, so that each participant can get a predicted hippocampus-centred regions age.

The HA gap in *APOE* genotype groups: We investigated the HA gap estimation in the participants with *APOE* $\epsilon 2/\epsilon 2$, *APOE* $\epsilon 2/\epsilon 3$, *APOE* $\epsilon 3/\epsilon 3$, *APOE* $\epsilon 3/\epsilon 4$, and *APOE* $\epsilon 4/\epsilon 4$ genotype. We also grouped participants as *APOE* $\epsilon 4$ homozygotes ($\epsilon 4/\epsilon 4$), *APOE* $\epsilon 4$ heterozygotes ($\epsilon 2/\epsilon 4$, $\epsilon 3/\epsilon 4$), and *APOE* $\epsilon 4$ non-carriers. HA gap was set as the dependent variable, and *APOE* $\epsilon 4$ carrier status (or *APOE* genotype), age, sex, scanner, and intracranial volume (ICV) were set as independent variables. Similarly, we also investigated the associations between hippocampus volume and *APOE* $\epsilon 4$ carrier status (or *APOE* genotype) using similar regression model by setting hippocampus volume (HV) as the dependent variable. For both analyses, *APOE* $\epsilon 4$ non-carriers served as the reference group, with other groups compared against this reference. Additionally, *APOE* genotypes $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 3$, and $\epsilon 3/\epsilon 3$ were individually set as reference groups for comparisons with the remaining groups.

Associations between HA gap and cognition: We also performed regression model to explore the associations between HA gap and cognition, by setting cognition as the dependent variable and HA gap as independent variable. All variables were z-transformed. We also included chronological age, sex, scanner, and years of education. as covariance variables in

the regression model. Additionally, associations between hippocampus volume and cognition were calculated using a similar method.

HA and modifiable risk factors: Associations between the left HA gap and modifiable risk factors (hypertension and diabetes, alcohol intake frequency, smoking status, physical activity) were explored by applying the regression model: left or right HA gap ~ hypertension + diabetes + alcohol intake frequency + smoking status + physical activity + BMI + *APOE* ε4 status + baseline age + sex + scanner + ICV. Here, BMI, *APOE* ε4 status, baseline age, sex, scanner, and ICV were used as covariance variables. Each risk factor had a specific reference group: participants who are non-hypertensive, or non-diabetic, drink daily or almost daily, those who have never smoked, engage in no physical activity. Other corresponding groups were compared against these reference groups individually.

Longitudinal analysis

The HA gap change rate in *APOE* genotype groups: The annual HA gap change rate was calculated by the gap difference between baseline and follow-up and then divided by the time interval between these two time points (years). The change rate was compared within different *APOE* genotype groups. We applied the similar regression model, which was used in the cross-sectional analysis, but set annual HA gap change as the dependent variable instead of HA gap. The annual hippocampus volume change was also estimated in a similar regression model. All variables were z-transformed, and these regression models were controlled by baseline age, sex, scanner, and ICV.

2.6 Occlusion analysis

To interpret the CNN model results and understand the regions of the input that contribute HA prediction, occlusion analysis (Lee et al., 2022; Zeiler and Fergus, 2014) was conducted

within the test dataset. Specifically, the input images were divided by $4 \times 4 \times 4$ grid, and for each square region, a corresponding occlusion mask with $16 \times 16 \times 16$ voxels was applied. The occlusion process involved systematically making the different occlusion mask with zero values. The age prediction on occluded images was performed using our original pretrained CNN model, and the model performance was evaluated by $MAE_{occlusion}$. The delta MAE was calculated as $MAE_{delta} = MAE_{occlusion} - MAE_{original}$ ($MAE_{original}$ represents the MAE of CNN model using original images). A delta MAE matrix ($4 \times 4 \times 4$) was generated by iterating occlusion for each square region. Finally, cubic interpolation was applied to the delta MAE matrix to obtain the original image size ($64 \times 64 \times 64$). This procedure was done in each test set in five-fold and the average of the 5 folds was finally calculated. Normalization was carried out by dividing the whole image by its maximum value.

To examine how important the hippocampus is in the HA prediction, we also removed hippocampus from the original ROI (occluded using the hippocampus mask), and then used it as the input of the pretrained model. The MAE was estimated to assess the model's performance.

2.7 Apply the 3D-CNN on a new dataset using transfer learning

We continue to examine the 3D-CNN model performance on another dataset: MAS cohort. As the sample size in MAS dataset is relatively small ($N=523$), we used 3-fold cross validation within MAS dataset, with $2/3$ participants using as the training set and $1/3$ as test set. The imaging processing procedure was the same as the procedure in UKB dataset. Specifically, the MAS dataset was randomly split into 3 folds. The MAS training set was full-fined using the pre-trained 3D-CNN model in UKB. As we used 5-fold cross validation in UKB, these 5 pre-trained models were applied to transfer learning separately.

3 Results

3.1 Demographics

Table 1 summarizes the demographic characteristics of the participants from the UK Biobank and the MAS. In the current study, 31,370 participants in UK Biobank were included in the five-fold analysis, and 26,206 of them have both HA data and genetic data, with *APOE* ϵ 4 non-carriers (72.18%), *APOE* ϵ 4 heterozygotes (25.59%), and *APOE* ϵ 4 homozygotes (2.23%). In the longitudinal analysis, 3,276 participants were included, with *APOE* ϵ 4 non-carriers (72.92%), *APOE* ϵ 4 heterozygotes (24.97%), and *APOE* ϵ 4 homozygotes (2.11%). The MAS cohort included 523 participants with a mean age of 78.35 years (SD = 4.66). The study workflow can be found in Figure 1.

3.2 Cross-sectional analysis

The HA gap in *APOE* genotype groups: Our model achieved a competitive state-of-the-art metric (MAE: 2.52 – 2.84 years in left HA model; 2.47 – 2.72 years in right HA model) (Table 2), with the correlation between chronological and predicted age ranging from 0.901 to 0.907 (Figure S1). Compared with *APOE* $\epsilon 4$ non-carriers, *APOE* $\epsilon 4$ homozygotes ($\epsilon 4/\epsilon 4$) showed significantly greater HA gap (left: $\beta = 0.166$, 95% confidence interval [CI] = [0.085, 0.248], $p = 6.30e-05$; right: $\beta = 0.192$, CI = [0.111, 0.274], $p = 4.03e-06$) and smaller hippocampus volume in both left and right hemispheres. Specifically, the HA gap in *APOE* $\epsilon 4$ homozygotes group was also significantly greater than that in $\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 3$ and $\epsilon 3/\epsilon 3$ group (Figure 2 and Table 3).

Whole brain age gap in *APOE* genotype groups: When the whole brain image used as input, MAE ranged from 2.62 to 2.94 years (Table 2). *APOE* $\epsilon 4$ homozygotes ($\epsilon 4/\epsilon 4$) showed greater whole brain age gap ($\beta = 0.162$, CI = [0.081, 0.244], $p = 9.98e-05$). Details can be found in Table S2.

Associations between HA gap and cognition: In the cross-sectional analysis in UK

Biobank, 19,441 participants have both HA gap data and cognition data. The greater HA gap in both left and right hemispheres was associated with lower cognitive performance:

processing speed, executive function, memory, and global cognition. Detailed results can be found in Figure 3.

HA gap and modifiable risk factors: As both the left and right HA gaps were significantly associated with non-modifiable variable: *APOE* genotype, we continued to examine their associations with modifiable risk factors: (hypertension, diabetes, alcohol intake frequency, smoking status, and physical activity). In UK Biobank dataset, we found that participants diagnosed with hypertension or diabetes showed significantly greater HA gap compared with healthy counterparts. Additionally, participants who have alcohol 1-3 times a month or special occasion only showed a smaller HA gap compared with daily drinkers ($p < 0.001$). Previous and current smokers exhibited a significantly greater HA gap compared with non-smokers ($p < 0.001$). However, we did not find significant associations between HA gap and physical activities. All details can be found in Figure 4 and Table S3.

3.3 Longitudinal analysis

The HA gap change rate in *APOE* genotype groups: Compared with *APOE* $\epsilon 2$

homozygotes, *APOE* $\epsilon 4$ homozygotes showed a greater longitudinal change rate in the left HA gap ($\beta = 0.581$, $p = 0.025$) (Table 4 and Figure 2). However, this effect was not observed in the right hemisphere. The left HA gap trajectory in different *APOE* genotype groups can be found in Figure S2. In addition, no significant differences in annual HV changes were found among the various *APOE* genotype groups.

3.4 Occlusion analysis

As a greater longitudinal age gap change rate in *APOE* $\epsilon 4$ homozygotes was only found in the left hippocampus ROI (compared with *APOE* $\epsilon 2$ homozygotes), we subsequently performed occlusion analysis on this region. The saliency map of the prediction model revealed that beyond the hippocampus, regions in proximity, including peaks around the thalamus, pallidum, putamen, amygdala, nearby cerebral cortex, and cerebral white matter, also play an important role in age prediction (Figure 5). The detailed saliency map can be found in Figure S3. Additionally, compared with the original model performance, the model using ROI without hippocampus exhibited higher MAE (2.96 – 3.59 years in left HA model; 2.79 – 3.50 years in right HA model). The specific MAE values for each of the five folds can be found in Table 2.

3.5 Transfer learning result in MAS dataset

The pretrained models from the UK Biobank were applied to the MAS dataset using transfer learning, and the performance is summarized in Table 5. The models were evaluated using three-fold cross-validation in MAS, with the MAE calculated for each test fold. For the left hippocampus ROI, the MAE ranged from 3.12 to 3.86 years. For the right hippocampus ROI, the MAE ranged from 3.11 to 4.77 years. These results indicated that the pretrained models from the UK Biobank can be effectively transferred to the MAS dataset, though the performance is generally lower than in the original UK Biobank dataset. The associations between HA gap, *APOE* genotypes, cognition, and modifiable risk factors can be found in Figure 2, Figure 3, and Figure S4.

4 Discussion

In the current study, we developed a 3D-CNN model to predict the age of left and right hippocampus ROI in a large population sample. Subsequently, we evaluated the *APOE* genotype influences on the HA gap, the change rate of HA gap, cognition, and various modifiable risk factors. Our findings indicate that *APOE* $\epsilon 4$ influences the HA gap in both hemispheres, and individuals with greater HA gap tend to exhibit poorer cognitive performance. Specially, participants with hypertension, diabetes, those who engage in heavy daily alcohol consumption, or are previous/current smokers are more likely to have a greater HA gap, which serves as an indicator of poorer health. In the longitudinal analysis, a significant difference in the change rate of the left HA gap was observed between *APOE* $\epsilon 4$ homozygotes and *APOE* $\epsilon 2$ homozygotes.

In the cross-sectional analysis, the HA gap varied across different *APOE* genotype groups, with *APOE* $\epsilon 4$ homozygotes (*APOE* $\epsilon 4/\epsilon 4$ carriers) exhibiting the greatest average HA gap in both hemispheres compared to others ($\epsilon 2/\epsilon 2$, $\epsilon 2/\epsilon 3$, $\epsilon 3/\epsilon 3$). This suggests that *APOE* $\epsilon 4/\epsilon 4$ accelerates the ageing process in the hippocampus ROI. In contrast, *APOE* $\epsilon 4$ heterozygotes showed a significantly greater HA gap only in right hemisphere compared to *APOE* $\epsilon 4$ non-carriers. Additionally, the differences in hippocampal volume (HV) across *APOE* genotypes were less pronounced than the differences in HA gap, indicating that *APOE* genotype may have a stronger influence on HA gap than on HV. Our findings indicated that *APOE* $\epsilon 4$ carriers exhibit not only smaller hippocampus volume but also a greater HA gap, even in the absence of an AD diagnosis.

Advanced age is usually accompanied by cognitive decline, and studies have consistently indicated a negative association between brain age gap and cognitive scores (Cole et al., 2018; Elliott et al., 2021). The brain age gap may act as the mediator between modifiable risk

factors and cognitive functioning (Chen et al., 2022). In early onset Alzheimer's disease, *APOE* $\epsilon 4$ carriers tend to experience a more rapid decline in cognitive functions such as processing speed, executive function, and memory (Polsinelli et al., 2023). Our finding that a greater HA gap is associated with poorer cognitive performance adds to the body of evidence supporting the link between brain age gap and cognitive decline.

Interestingly, in the longitudinal analysis, a notable divergence was observed: *APOE* $\epsilon 4/\epsilon 4$ carriers showed a rapid increase in left HA gap, while $\epsilon 2/\epsilon 2$ carriers demonstrated a decreased HA gap over the years. These effects, however, were not mirrored in the change rate of the right HA gap. Given the established higher risk associated with *APOE* $\epsilon 4/\epsilon 4$ for AD and the protective nature of *APOE* $\epsilon 2/\epsilon 2$ (Serrano-Pozo et al., 2021), it is plausible that the increased change rate of HA gap contributes to the heightened risk in $\epsilon 4/\epsilon 4$ carriers, particularly in the left hippocampus ROI. Supporting this, a recent study demonstrated that *APOE* $\epsilon 4$ contributed to faster-accelerated atrophy in the left hippocampus in persistent cognitive normal and the transition from normal cognitive stages to dementia (Huang et al., 2023). Furthermore, verbal memory decline was also correlated with left hippocampal atrophy (Baxter et al., 2023). In addition to longitudinal volume atrophy, *APOE* $\epsilon 4$ was also related to longitudinal decrease in hippocampal activation in normal ageing (Håglin et al., 2023). Our observation of longitudinal change rate in HA gap provides additional evidence that *APOE* $\epsilon 4$ can influence both functional and structural aspects of hippocampus-related regions.

The investigation into associations between HA gaps and both non-modifiable and modifiable variables has yielded insightful findings. Participants diagnosed with hypertension or diabetes exhibited significantly greater HA gaps compared to their healthier counterparts. Additionally, individuals who engage in heavy daily alcohol consumption or smoking also

exhibited greater HA gaps. This aligns with previous findings that link increased brain age to frequent tobacco and alcohol use (Ning et al., 2020). The association between unhealthy lifestyle, such as high alcohol use, and lower hippocampal volume (Binnewies et al., 2023) may further explain why heavy daily alcohol users tend to have a larger HA gap. Together, these findings enhance our understanding of the complex factors contributing to aging in hippocampus-related regions.

The “black box” nature of deep learning models usually makes the results hard to explain, but occlusion analysis in the current study helped the interpretability of the HA prediction model. Interestingly, occlusion analysis demonstrated that the hippocampus nearby regions have great contribution to HA prediction. Specifically, thalamus, pallidum, nearby cerebral cortex, and cerebral white matter are critical regions in HA prediction. Many of these regions belong to the basal ganglia and cerebellum, which are essential for motor control and emotional processing and influence limbic function (Arber and Costa, 2022; Pierce and Péron, 2020). Age-related differences in thalamocortical connectivity can contribute to changes in attention, working memory, and episodic memory processes associated with ageing (Fama and Sullivan, 2015; Hughes et al., 2012; Ystad et al., 2010). Additionally, the HA prediction model using $64 \times 64 \times 64$ ROI removing hippocampus showed a slightly larger MAE than the original model which included hippocampus. This slight MAE differences between models indicate that while the hippocampus contributes to HA prediction, the surrounding regions also play a vital role in age estimation. Collectively, these analyses demonstrate that the 3D bounding boxes of the ROI surrounding the hippocampus could serve as a valuable representation for brain age prediction.

Our study has some strengths and limitations. Firstly, while the sample size for both training and test sets is relatively large, generalizability to broader populations, including diverse

demographics and clinical conditions, may require further exploration. Secondly, although our findings offer valuable insights into the associations between HA gap and factors such as *APOE* genotype, chronic diseases, and lifestyle factors, it is important to note that our study does not establish HA gap as a definitive biomarker for AD. Future research is needed to thoroughly investigate and validate the potential of HA gap as a biomarker for AD diagnosis, particularly in cohorts with confirmed AD diagnoses. Lastly, while our study focused on associations with *APOE* genotype, chronic diseases, and lifestyle factors, incorporating additional variables such as other genetic markers, socioeconomic factors, and broader health metrics could offer a more comprehensive view of the determinants influencing HA gap.

Conclusion

In conclusion, our study has shed light on the intricate relationship between HA gap and various determinants, including both non-modifiable factors like *APOE* genotype and modifiable factors. *APOE* $\epsilon 4$ homozygotes exhibited greater left and right HA gaps, along with a faster change rate in left HA gap, indicating that left HA gap might be a potential *APOE*-related biomarker or a health indicator.

Acknowledgments

C.D. is supported by Scientia Scholarship of University of New South Wales, Sydney. We extend our gratitude to the UK Biobank and the Sydney Memory and Ageing Study (MAS) databases for providing valuable data.

References

- Abushakra, S., Porsteinsson, A.P., Sabbagh, M., Bracoud, L., Schaerer, J., Power, A., Hey, J.A., Scott, D., Suhy, J., Tolar, M., for the Alzheimer's Disease Neuroimaging, I., 2020. APOE $\epsilon 4/\epsilon 4$ homozygotes with early Alzheimer's disease show accelerated hippocampal atrophy and cortical thinning that correlates with cognitive decline. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 6, e12117.
- Arber, S., Costa, R.M., 2022. Networking brainstem and basal ganglia circuits for movement. *Nature Reviews Neuroscience* 23, 342-360.
- Baxter, L.C., Limback-Stokin, M., Patten, K.J., Arreola, A.C., Locke, D.E.C., Hu, L., Zhou, Y., Caselli, R.J., 2023. Hippocampal connectivity and memory decline in cognitively intact APOE $\epsilon 4$ carriers. *Alzheimer's & Dementia* 19, 3806-3814.
- Binnewies, J., Nawijn, L., Brandmaier, A.M., Baaré, W.F.C., Boraxbekk, C.-J., Demnitz, N., Drevon, C.A., Fjell, A.M., Lindenberger, U., Madsen, K.S., Nyberg, L., Topiwala, A., Walhovd, K.B., Ebmeier, K.P., Penninx, B.W.J.H., 2023. Lifestyle-related risk factors and their cumulative associations with hippocampal and total grey matter volume across the adult lifespan: A pooled analysis in the European Lifebrain consortium. *Brain Research Bulletin* 200, 110692.
- Butler, E.R., Chen, A., Ramadan, R., Le, T.T., Ruparel, K., Moore, T.M., Satterthwaite, T.D., Zhang, F., Shou, H., Gur, R.C., Nichols, T.E., Shinohara, R.T., 2021. Pitfalls in brain age analyses. *Human Brain Mapping* 42, 4092-4101.
- Cacciaglia, R., Molinuevo, J.L., Falcón, C., Brugulat-Serrat, A., Sánchez-Benavides, G., Gramunt, N., Esteller, M., Morán, S., Minguillón, C., Fauria, K., Gispert, J.D., for the, A.s., 2018. Effects of APOE- $\epsilon 4$ allele load on brain morphology in a cohort of middle-aged healthy individuals with enriched genetic risk for Alzheimer's disease. *Alzheimer's & Dementia* 14, 902-912.
- Chapleau, M., Aldebert, J., Montembeault, M., Brambati, S.M., 2016. Atrophy in Alzheimer's Disease and Semantic Dementia: An ALE Meta-Analysis of Voxel-Based Morphometry Studies. *Journal of Alzheimer's Disease* 54, 941-955.
- Chauveau, L., Kuhn, E., Palix, C., Felisatti, F., Ourry, V., de La Sayette, V., Chételat, G., de Flores, R., 2021. Medial Temporal Lobe Subregional Atrophy in Aging and Alzheimer's Disease: A Longitudinal Study. 13.
- Chen, C.-L., Kuo, M.-C., Chen, P.-Y., Tung, Y.-H., Hsu, Y.-C., Huang, C.-W.C., Chan, W.P., Tseng, W.-Y.I., 2022. Validation of neuroimaging-based brain age gap as a mediator between modifiable risk factors and cognition. *Neurobiology of Aging* 114, 61-72.
- Chudasama, Y.V., Khunti, K., Gillies, C.L., Dhalwani, N.N., Davies, M.J., Yates, T., Zaccardi, F., 2020. Healthy lifestyle and life expectancy in people with multimorbidity in the UK Biobank: A longitudinal cohort study. *PLOS Medicine* 17, e1003332.
- Cole, J.H., Franke, K., 2017. Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. *Trends in Neurosciences* 40, 681-690.

Cole, J.H., Ritchie, S.J., Bastin, M.E., Valdés Hernández, M.C., Muñoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S.E., Zhang, Q., Wray, N.R., Redmond, P., Marioni, R.E., Starr, J.M., Cox, S.R., Wardlaw, J.M., Sharp, D.J., Deary, I.J., 2018. Brain age predicts mortality. *Molecular Psychiatry* 23, 1385-1392.

de Flores, R., Demeilliez-Servouin, S., Kuhn, E., Chauveau, L., Landeau, B., Delcroix, N., Gonneaud, J., Vivien, D., Chételat, G., 2023. Respective influence of beta-amyloid and APOE ϵ 4 genotype on medial temporal lobe subregions in cognitively unimpaired older adults. *Neurobiology of Disease* 181, 106127.

de Flores, R., Wisse, L.E.M., Das, S.R., Xie, L., McMillan, C.T., Trojanowski, J.Q., Robinson, J.L., Grossman, M., Lee, E., Irwin, D.J., Yushkevich, P.A., Wolck, D.A., 2020. Contribution of mixed pathology to medial temporal lobe atrophy in Alzheimer's disease. *Alzheimer's & Dementia* 16, 843-852.

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M.H., White, T., van Hees, V.T., Trenell, M.I., Owen, C.G., Preece, S.J., Gillions, R., Sheard, S., Peakman, T., Brage, S., Wareham, N.J., 2017. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLOS ONE* 12, e0169649.

Donix, M., Burggren, A.C., Suthana, N.A., Siddarth, P., Ekstrom, A.D., Krupa, A.K., Jones, M., Rao, A., Martin-Harris, L., Ercoli, L.M., Miller, K.J., Small, G.W., Bookheimer, S.Y., 2010. Longitudinal changes in medial temporal cortical thickness in normal subjects with the APOE-4 polymorphism. *NeuroImage* 53, 37-43.

Elliott, M.L., Belsky, D.W., Knodt, A.R., Ireland, D., Melzer, T.R., Poulton, R., Ramrakha, S., Caspi, A., Moffitt, T.E., Hariri, A.R., 2021. Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort. *Molecular Psychiatry* 26, 3829-3838.

Fama, R., Sullivan, E.V., 2015. Thalamic structures and associated cognitive functions: Relations with age and aging. *Neuroscience & Biobehavioral Reviews* 54, 29-37.

Fawns-Ritchie, C., Deary, I.J., 2020. Reliability and validity of the UK Biobank cognitive tests. *PLOS ONE* 15, e0231627.

Fotuhi, M., Do, D., Jack, C., 2012. Modifiable factors that alter the size of the hippocampus with ageing. *Nature Reviews Neurology* 8, 189-202.

Franz, C.E., Hatton, S.N., Elman, J.A., Warren, T., Gillespie, N.A., Whitsel, N.A., Puckett, O.K., Dale, A.M., Eyler, L.T., Fennema-Notestine, C., Hagler, D.J., Hauger, R.L., McKenzie, R., Neale, M.C., Panizzon, M.S., Pearce, R.C., Reynolds, C.A., Sanderson-Cimino, M., Toomey, R., Tu, X.M., Williams, M., Xian, H., Lyons, M.J., Kremen, W.S., 2021. Lifestyle and the aging brain: interactive effects of modifiable lifestyle behaviors and cognitive ability in men from midlife to old age. *Neurobiology of Aging* 108, 80-89.

Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H., Alzheimer's Disease Neuroimaging, I., 2013. BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLOS ONE* 8, e67346.

Gharbi-Meliani, A., Dugravot, A., Sabia, S., Regy, M., Fayosse, A., Schnitzler, A., Kivimäki, M., Singh-Manoux, A., Dumurgier, J., 2021. The association of APOE ϵ 4 with cognitive function over the adult life course and incidence of dementia: 20 years follow-up of the Whitehall II study. *Alzheimers Res Ther* 13, 5.

Gottesman, R.F., Seshadri, S., 2022. Risk Factors, Lifestyle Behaviors, and Vascular Brain Health. *Stroke* 53, 394-403.

Håglin, S., Koch, E., Schäfer Hackenhaar, F., Nyberg, L., Kauppi, K., 2023. APOE ϵ 4, but not polygenic Alzheimer's disease risk, is related to longitudinal decrease in hippocampal brain activity in non-demented individuals. *Scientific Reports* 13, 8433.

Huang, Y.-W.A., Zhou, B., Wernig, M., Südhof, T.C., 2017. ApoE2, ApoE3, and ApoE4 Differentially Stimulate APP Transcription and A β Secretion. *Cell* 168, 427-441.e421.

Huang, Y., Shan, Y., Qin, W., Zhao, G., 2023. Apolipoprotein E ϵ 4 accelerates the longitudinal cerebral atrophy in open access series of imaging studies-3 elders without dementia at enrollment. 15.

Hughes, E.J., Bond, J., Svrckova, P., Makropoulos, A., Ball, G., Sharp, D.J., Edwards, A.D., Hajnal, J.V., Counsell, S.J., 2012. Regional changes in thalamic shape and volume with increasing age. *Neuroimage* 63, 1134-1142.

Koizumi, K., Hattori, Y., Ahn, S.J., Buendia, I., Ciacciarelli, A., Uekawa, K., Wang, G., Hiller, A., Zhao, L., Voss, H.U., Paul, S.M., Schaffer, C., Park, L., Iadecola, C., 2018. Apo ϵ 4 disrupts neurovascular regulation and undermines white matter integrity and cognitive function. *Nature Communications* 9, 3816.

Le Goallec, A., Collin, S., Jabri, M.H., Diai, S., Vincent, T., Patel, C.J., 2023. Machine learning approaches to predict age from accelerometer records of physical activity at biobank scale. *PLOS Digital Health* 2, e0000176.

Lee, J., Burkett, B.J., Min, H.-K., Senjem, M.L., Lundt, E.S., Botha, H., Graff-Radford, J., Barnard, L.R., Gunter, J.L., Schwarz, C.G., Kantarci, K., Knopman, D.S., Boeve, B.F., Lowe, V.J., Petersen, R.C., Jack, C.R., Jones, D.T., 2022. Deep learning-based brain age prediction in normal aging and dementia. *Nature Aging* 2, 412-424.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S.G., Dias, A., Fox, N., Gitlin, L.N., Howard, R., Kales, H.C., Kivimäki, M., Larson, E.B., Ogunniyi, A., Orgeta, V., Ritchie, K., Rockwood, K., Sampson, E.L., Samus, Q., Schneider, L.S., Selbæk, G., Teri, L., Mukadam, N., 2020. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet* 396, 413-446.

Martin, S.A., Townend, F.J., Barkhof, F., Cole, J.H., 2023. Interpretable machine learning for dementia: A systematic review. *Alzheimer's & Dementia* 19, 2135-2149.

Millar, P.R., Gordon, B.A., Lockett, P.H., Benzinger, T.L.S., Cruchaga, C., Fagan, A.M., Hassenstab, J.J., Perrin, R.J., Schindler, S.E., Allegri, R.F., Day, G.S., Farlow, M.R., Mori, H., Nübling, G., The Dominantly Inherited Alzheimer, N., Bateman, R.J., Morris, J.C., Ances,

- B.M., 2023. Multimodal brain age estimates relate to Alzheimer disease biomarkers and cognition in early stages: a cross-sectional observational study. *Elife* 12, e81869.
- Mishra, S., Blazey, T.M., Holtzman, D.M., Cruchaga, C., Su, Y., Morris, J.C., Benzinger, T.L.S., Gordon, B.A., 2018. Longitudinal brain imaging in preclinical Alzheimer disease: impact of APOE $\epsilon 4$ genotype. *Brain* 141, 1828-1839.
- Moffat, S.D., Szekely, C.A., Zonderman, A.B., Kabani, N.J., Resnick, S.M., 2000. Longitudinal change in hippocampal volume as a function of apolipoprotein E genotype. *Neurology* 55, 134-136.
- Mori, E., Lee, K., Yasuda, M., Hashimoto, M., Kazui, H., Hirono, N., Matsui, M., 2002. Accelerated hippocampal atrophy in Alzheimer's disease with apolipoprotein E $\epsilon 4$ allele. *Annals of Neurology* 51, 209-214.
- Ning, K., Zhao, L., Matloff, W., Sun, F., Toga, A.W., 2020. Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants. *Scientific Reports* 10, 10.
- Pendlebury, S.T., Rothwell, P.M., 2009. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: a systematic review and meta-analysis. *The Lancet Neurology* 8, 1006-1018.
- Peng, H., Gong, W., Beckmann, C.F., Vedaldi, A., Smith, S.M., 2021. Accurate brain age prediction with lightweight deep neural networks. *Medical Image Analysis* 68, 101871.
- Pierce, J.E., Péron, J., 2020. The basal ganglia and the cerebellum in human emotion. *Social Cognitive and Affective Neuroscience* 15, 599-613.
- Poloni, K.M., Ferrari, R.J., 2022. A deep ensemble hippocampal CNN model for brain age estimation applied to Alzheimer's diagnosis. *Expert Systems with Applications* 195, 116622.
- Polsinelli, A.J., Logan, P.E., Lane, K.A., Manchella, M.K., Nemes, S., Sanjay, A.B., Gao, S., Apostolova, L.G., 2023. APOE $\epsilon 4$ carrier status and sex differentiate rates of cognitive decline in early- and late-onset Alzheimer's disease. *Alzheimer's & Dementia* 19, 1983-1993.
- Qiu, S., Joshi, P.S., Miller, M.I., Xue, C., Zhou, X., Karjadi, C., Chang, G.H., Joshi, A.S., Dwyer, B., Zhu, S., Kaku, M., Zhou, Y., Alderazi, Y.J., Swaminathan, A., Kedar, S., Saint-Hilaire, M.-H., Auerbach, S.H., Yuan, J., Sartor, E.A., Au, R., Kolachalama, V.B., 2020. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. *Brain* 143, 1920-1933.
- Rongve, A., Witoelar, A., Ruiz, A., Athanasiu, L., Abdelnour, C., Clarimon, J., Heilmann-Heimbach, S., Hernández, I., Moreno-Grau, S., de Rojas, I., Morenas-Rodríguez, E., Fladby, T., Sando, S.B., Bråthen, G., Blanc, F., Bousiges, O., Lemstra, A.W., van Steenoven, I., Londos, E., Almdahl, I.S., Pålhaugen, L., Eriksen, J.A., Djurovic, S., Stordal, E., Saltvedt, I., Ulstein, I.D., Bettella, F., Desikan, R.S., Idland, A.-V., Toft, M., Pihlstrøm, L., Snaedal, J., Tárrega, L., Boada, M., Lleó, A., Stefánsson, H., Stefánsson, K., Ramírez, A., Aarsland, D., Andreassen, O.A., 2019. GBA and APOE $\epsilon 4$ associate with sporadic dementia with Lewy bodies in European genome wide association study. *Scientific Reports* 9, 7013.
- Sachdev, P.S., Brodaty, H., Reppermund, S., Kochan, N.A., Trollor, J.N., Draper, B., Slavin, M.J., Crawford, J., Kang, K., Broe, G.A., Mather, K.A., 2010. The Sydney Memory and

Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70-90 years. *International psychogeriatrics* 22, 1248.

Saeed, U., Desmarais, P., Masellis, M., 2021. The APOE ϵ 4 variant and hippocampal atrophy in Alzheimer's disease and Lewy body dementia: a systematic review of magnetic resonance imaging studies and therapeutic relevance. *Expert Review of Neurotherapeutics* 21, 851-870.

Sarasua, I., Pölsterl, S., Wachinger, C., 2022. Hippocampal representations for deep learning on Alzheimer's disease. *Scientific Reports* 12, 8619.

Serrano-Pozo, A., Das, S., Hyman, B.T., 2021. APOE and Alzheimer's disease: advances in genetics, pathophysiology, and therapeutic approaches. *The Lancet Neurology* 20, 68-80.

Skrobot, O.A., McKnight, A.J., Passmore, P.A., Seripa, D., Mecocci, P., Panza, F., Kalaria, R., Wilcock, G., Munafò, M., Erkinjuntti, T., Karhunen, P., Pessi, T., Martiskainen, M., Love, S., the, G., Environmental Risk for Alzheimer's disease, C., Kehoe, P.G., 2016. A Validation Study of Vascular Cognitive Impairment Genetics Meta-Analysis Findings in an Independent Collaborative Cohort. *Journal of Alzheimer's Disease* 53, 981-989.

Smith, S.M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T.E., Miller, K.L., 2019. Estimation of brain age delta from brain imaging. *NeuroImage* 200, 528-539.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* 12, e1001779.

Tian, Y.E., Cropley, V., Maier, A.B., Lautenschlager, N.T., Breakspear, M., Zalesky, A., 2023. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nature Medicine* 29, 1221-1231.

Wang, D., Honnorat, N., Fox, P.T., Ritter, K., Eickhoff, S.B., Seshadri, S., Habes, M., 2023. Deep neural network heatmaps capture Alzheimer's disease patterns reported in a large meta-analysis of neuroimaging studies. *NeuroImage* 269, 119929.

Wang, J., Knol, M.J., Tiulpin, A., Dubost, F., de Bruijne, M., Vernooij, M.W., Adams, H.H.H., Ikram, M.A., Niessen, W.J., Roshchupkin, G.V., 2019. Gray Matter Age Prediction as a Biomarker for Risk of Dementia. *Proceedings of the National Academy of Sciences* 116, 21213-21218.

Ystad, M., Eichele, T., Lundervold, A.J., Lundervold, A., 2010. Subcortical functional connectivity and verbal episodic memory in healthy elderly—A resting state fMRI study. *Neuroimage* 52, 379-388.

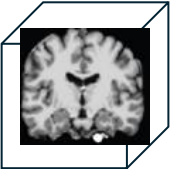
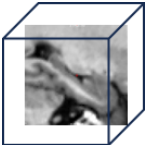
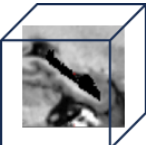
Zeiler, M.D., Fergus, R., 2014. Visualizing and Understanding Convolutional Networks. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), *Computer Vision – ECCV 2014*. Springer International Publishing, Cham, pp. 818-833.

Table 1. Demographic characteristics.

Participants (number)	UKB: Only have one time point imaging data. (N=31,370)	UKB: Have two-time points imaging data. (N=3,893)	MAS: (N=523)
Age, years, mean (SD)	64.50 (7.52)	Baseline: 62.07 (7.40) Follow-up: 64.70 (7.20) Gap years: 2.63 (1.05)	78.35 (4.66)
Male, number (%)	14915 (47.55%)	1816 (46.65%)	233 (44.55%)
Genetic Data			
APOE genotype, number (%)	Total: N=26,206	Total: N=3,276	Total: N=517
$\epsilon 2\epsilon 2$	135 (0.52%)	19 (0.5%)	3 (0.58%)
$\epsilon 2\epsilon 3$	3235 (12.34%)	393 (12%)	74 (14.31%)
$\epsilon 2\epsilon 4$	605 (2.31%)	88 (2.69%)	6 (1.16%)
$\epsilon 3\epsilon 3$	15546 (59.32%)	1977 (60.35%)	316 (61.12%)
$\epsilon 3\epsilon 4$	6101 (23.28%)	730 (22.28%)	109 (21.08%)
$\epsilon 4\epsilon 4$	584 (2.23%)	69 (2.11%)	9 (1.74%)
$\epsilon 4$ non-carriers	18916 (72.18%)	2389 (72.92%)	393 (76.02%)
$\epsilon 4$ -heterozygotes	6706 (25.59%)	818 (24.97%)	115 (22.24%)
$\epsilon 4$ -homozygotes	584 (2.23%)	69 (2.11%)	9 (1.74)

Notes: UKB: UK Biobank. MAS: Sydney Memory and Ageing Study. SD: standard deviation.

Table 2. Performance of 3D Convolutional Neural Network (3D-CNN) model with five-fold cross validation (evaluated by MAE) in the UK Biobank.

Input image	Measures	CV1	CV2	CV3	CV4	CV5
	Whole brain (MAE)	2.62	2.70	2.66	2.62	2.94
	Left ROI (MAE)	2.52	2.61	2.61	2.60	2.84
	Right ROI (MAE)	2.47	2.63	2.69	2.53	2.72
	Left ROI without hippocampus (MAE)	2.96	3.48	3.59	2.80	3.15
	Right ROI without hippocampus (MAE)	2.79	3.14	3.15	3.50	2.99

Notes: MAE: mean absolute error. CV1-CV5: five-fold cross-validation. ROI: regions of interests.

Table 3: Cross-sectional analysis in UK Biobank: The HA gap and HV in different *APOE* genotype groups.

Cross-sectional regression model: HA gap or HV ~ <i>APOE</i> 4 status + real age + sex + scanner + ICV													
Reference	Phenotypes	<i>APOE</i> ε4 heterozygotes				<i>APOE</i> ε4 homozygotes							
		β	95% CI	SE	P	β	95% CI	SE	P				
<i>APOE</i> ε4 non-carriers	Left HA gap	0.023	[-0.005, 0.051]	0.014	0.101	0.166	[0.085, 0.248]	0.042	6.30e-05				
	Right HA gap	0.045	[0.017, 0.073]	0.014	0.001	0.192	[0.111, 0.274]	0.042	4.03e-06				
	Left HV	-0.019	[-0.041, 0.003]	0.011	0.086	-0.105	[-0.170, -0.041]	0.033	0.001				
	Right HV	-4.68e-04	[-0.022, 0.021]	0.011	0.967	-0.094	[-0.159, -0.029]	0.033	0.004				
Cross-sectional regression model: HA gap or HV ~ <i>APOE</i> genotype + real age + sex + scanner + ICV													
Reference	Phenotypes	e2e4 (N=605)				e3e4 (N=6101)				e4e4 (N=584)			
		β	95% CI	SE	P	β	95% CI	SE	P	β	95% CI	SE	P
e2e2 (N=135)	Left HA gap	0.095	[-0.090, 0.280]	0.094	0.313	0.107	[-0.061, 0.276]	0.086	0.213	0.249	[0.064, 0.435]	0.094	0.008
	Right HA gap	0.029	[-0.156, 0.215]	0.095	0.756	0.077	[-0.093, 0.246]	0.086	0.376	0.220	[0.034, 0.406]	0.095	0.020
	Left HV	-0.057	[-0.202, 0.089]	0.074	0.447	-0.030	[-0.163, 0.104]	0.068	0.664	-0.118	[-0.265, 0.028]	0.075	0.113
	Right HV	0.025	[-0.121, 0.172]	0.075	0.734	0.030	[-0.105, 0.164]	0.068	0.666	-0.064	[-0.212, 0.083]	0.075	0.392
e2e3 (N=3235)	Left HA gap	-0.008	[-0.094, 0.078]	0.044	0.859	0.005	[-0.038, 0.047]	0.022	0.833	0.147	[0.060, 0.234]	0.044	0.001
	Right HA gap	-0.016	[-0.102, 0.070]	0.044	0.718	0.031	[-0.011, 0.074]	0.022	0.148	0.175	[0.087, 0.262]	0.045	9.34e-05
	Left HV	-0.045	[-0.113, 0.023]	0.035	0.191	-0.018	[-0.052, 0.015]	0.017	0.282	-0.107	[-0.176, -0.038]	0.035	0.002
	Right HV	-0.007	[-0.075, 0.061]	0.035	0.842	-0.003	[-0.036, 0.031]	0.017	0.870	-0.097	[-0.166, -0.027]	0.035	0.006
e3e3 (N=15546)	Left HA gap	0.015	[-0.065, 0.096]	0.041	0.712	0.028	[-0.002, 0.057]	0.015	0.066	0.170	[0.088, 0.251]	0.042	4.73e-05
	Right HA gap	0.006	[-0.075, 0.086]	0.041	0.894	0.053	[0.023, 0.082]	0.015	4.56e-04	0.196	[0.114, 0.278]	0.042	2.90e-06
	Left HV	-0.043	[-0.107, 0.020]	0.032	0.183	-0.016	[-0.039, 0.007]	0.012	0.171	-0.105	[-0.170, -0.040]	0.033	0.001
	Right HV	-0.004	[-0.068, 0.060]	0.033	0.904	0.000	[-0.023, 0.024]	0.012	0.986	-0.094	[-0.159, -0.029]	0.033	0.005

Notes: HA: hippocampus ROI age; HV: hippocampus volume; ICV: intracranial volume.

Table 4: Longitudinal analysis in UK Biobank: The change rate of HA gap and HV in different *APOE* genotype groups.

Longitudinal regression model: HA gap or HV change rate ~ <i>APOE</i> 4 status + real age + sex + scanner + ICV													
Reference	Phenotypes	<i>APOE</i> ε4 heterozygotes				<i>APOE</i> ε4 homozygotes							
		β	95% CI	SE	P	β	95% CI	SE	P				
<i>APOE</i> ε4 non-carriers	Left HA gap	0.054	[-0.025, 0.134]	0.040	0.181	0.223	[-0.016, 0.462]	0.122	0.067				
	Right HA gap	0.043	[-0.036, 0.123]	0.041	0.284	0.049	[-0.190, 0.289]	0.122	0.685				
	Left HV	0.022	[-0.054, 0.097]	0.039	0.573	0.095	[-0.133, 0.323]	0.116	0.414				
	Right HV	-0.049	[-0.127, 0.029]	0.040	0.221	0.055	[-0.180, 0.289]	0.120	0.646				
Longitudinal regression model: HA gap or HV change rate ~ <i>APOE</i> genotype + real age + sex + scanner + ICV													
Reference	Phenotypes	e2e4 (N=88)				e3e4 (N=730)				e4e4 (N=69)			
		β	95% CI	SE	P	β	95% CI	SE	P	β	95% CI	SE	P
e2e2 (N=19)	Left HA gap	0.485	[-0.010, 0.980]	0.253	0.055	0.403	[-0.052, 0.857]	0.232	0.083	0.581	[0.074, 1.087]	0.258	0.025
	Right HA gap	-0.125	[-0.620, 0.371]	0.253	0.622	-0.094	[-0.550, 0.361]	0.232	0.685	-0.091	[-0.599, 0.416]	0.259	0.724
	Left HV	0.112	[-0.361, 0.584]	0.241	0.643	0.193	[-0.241, 0.626]	0.221	0.383	0.257	[-0.226, 0.740]	0.246	0.297
	Right HV	-0.209	[-0.695, 0.277]	0.248	0.398	-0.079	[-0.525, 0.367]	0.228	0.728	0.010	[-0.487, 0.508]	0.254	0.968
e2e3 (N=393)	Left HA gap	0.095	[-0.136, 0.325]	0.118	0.422	0.012	[-0.110, 0.135]	0.063	0.843	0.190	[-0.065, 0.446]	0.130	0.144
	Right HA gap	0.007	[-0.224, 0.238]	0.118	0.954	0.037	[-0.086, 0.160]	0.063	0.554	0.040	[-0.216, 0.296]	0.130	0.760
	Left HV	-0.096	[-0.316, 0.124]	0.112	0.395	-0.014	[-0.131, 0.103]	0.060	0.810	0.050	[-0.193, 0.294]	0.124	0.687
	Right HV	-0.105	[-0.332, 0.121]	0.115	0.361	0.025	[-0.096, 0.145]	0.061	0.687	0.114	[-0.136, 0.365]	0.128	0.372
e3e3 (N=1977)	Left HA gap	0.131	[-0.083, 0.344]	0.109	0.230	0.048	[-0.036, 0.133]	0.043	0.263	0.226	[-0.013, 0.466]	0.122	0.064
	Right HA gap	0.020	[-0.194, 0.233]	0.109	0.857	0.050	[-0.035, 0.135]	0.043	0.250	0.053	[-0.187, 0.293]	0.122	0.666
	Left HV	-0.043	[-0.247, 0.160]	0.104	0.676	0.038	[-0.043, 0.119]	0.041	0.359	0.102	[-0.126, 0.331]	0.117	0.380
	Right HV	-0.176	[-0.385, 0.033]	0.107	0.099	-0.046	[-0.129, 0.037]	0.042	0.280	0.044	[-0.192, 0.279]	0.120	0.717

Notes: HA: hippocampus ROI age; HV: hippocampus volume; ICV: intracranial volume.

Table 5. Performance of pretrained model applied to MAS using transfer learning.

MAS measures	Pre-trained model	MAS: CV1	MAS: CV2	MAS: CV3
Left ROI (MAE)	UKB-CV1	3.31	3.13	3.47
	UKB-CV2	3.45	3.15	3.86
	UKB-CV3	3.37	3.12	3.56
	UKB-CV4	3.46	3.20	3.48
	UKB-CV5	3.38	3.14	3.62
Right ROI (MAE)	UKB-CV1	3.50	3.23	4.30
	UKB-CV2	3.64	3.11	4.55
	UKB-CV3	3.56	3.14	4.59
	UKB-CV4	3.53	3.13	4.77
	UKB-CV5	3.50	3.11	4.41

Notes: MAE: mean absolute error. CV: cross-validation. UKB-CV1-CV5: five-fold cross-validation in UK Biobank. MAS:CV1-CV3: three-fold cross-validation in MAS. ROI: regions of interest. MAS: Sydney Memory and Ageing Study.

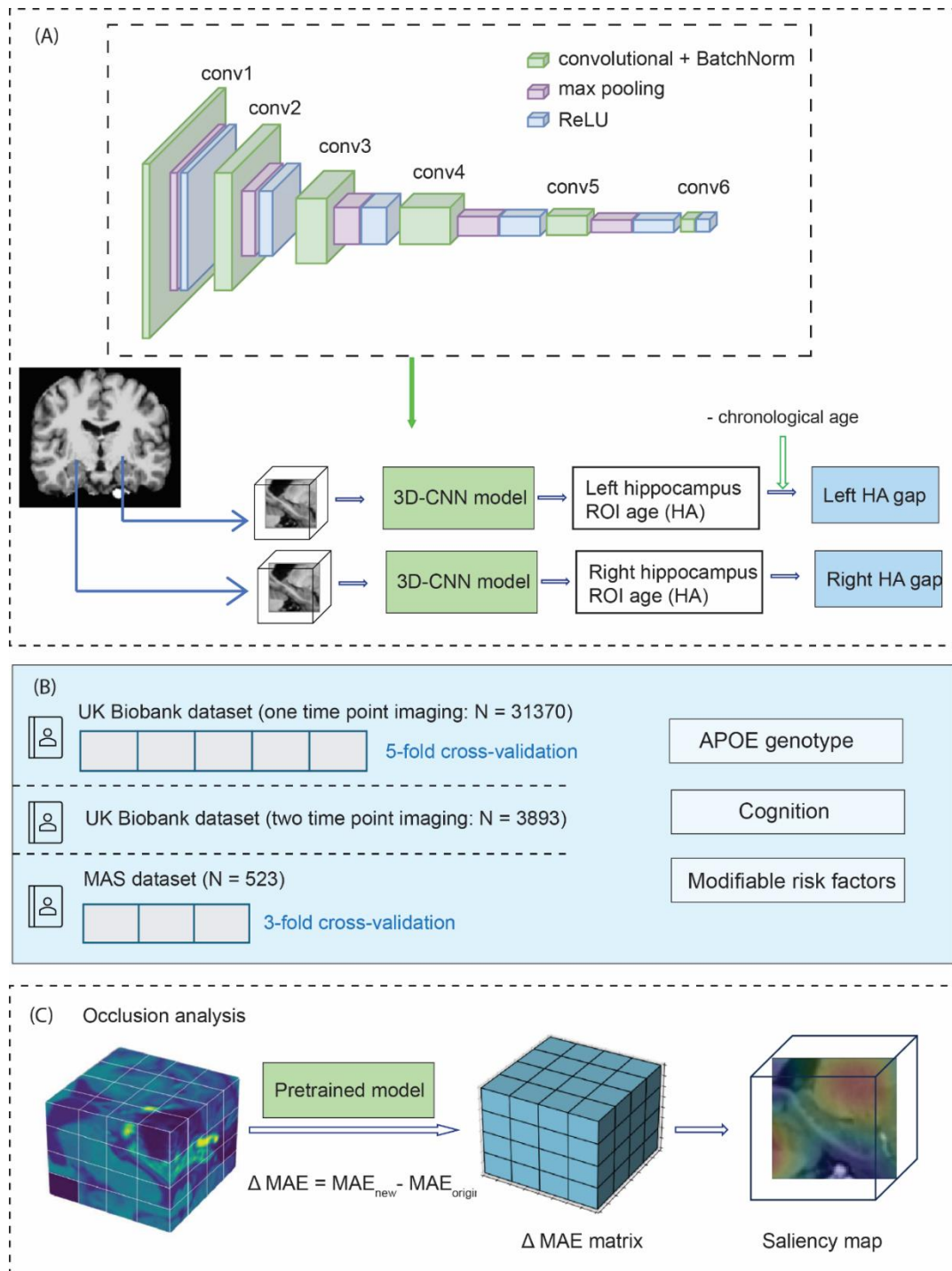
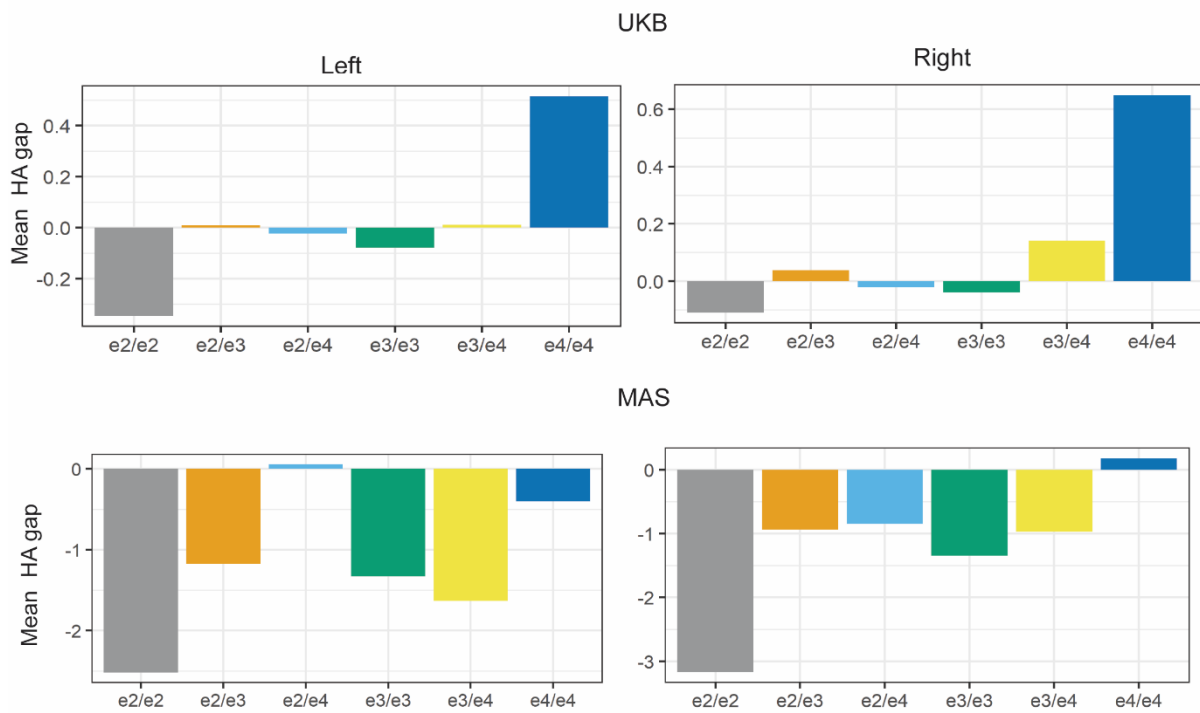


Figure 1. Study workflow. The architecture of the age prediction model in left and right hippocampus-centred regions of interest (hippocampus ROI) based on three-dimensional convolutional neural network (3D-CNN). Model input was 3D $64 \times 64 \times 64$ voxels hippocampus ROI image, and output was the predicted hippocampus ROI age (HA). The HA gap was calculated by subtracting the chronological age from the HA. Associations between HA gap and APOE genotype, cognition, and modifiable risk factors were further explored. The pre-trained model from the UK Biobank dataset was applied to the MAS dataset using transfer learning. (B) Dataset in the current study. (C) Occlusion analysis. MAS: Sydney Memory and Ageing Study.

(A) Cross-sectional analysis



(B) Longitudinal analysis

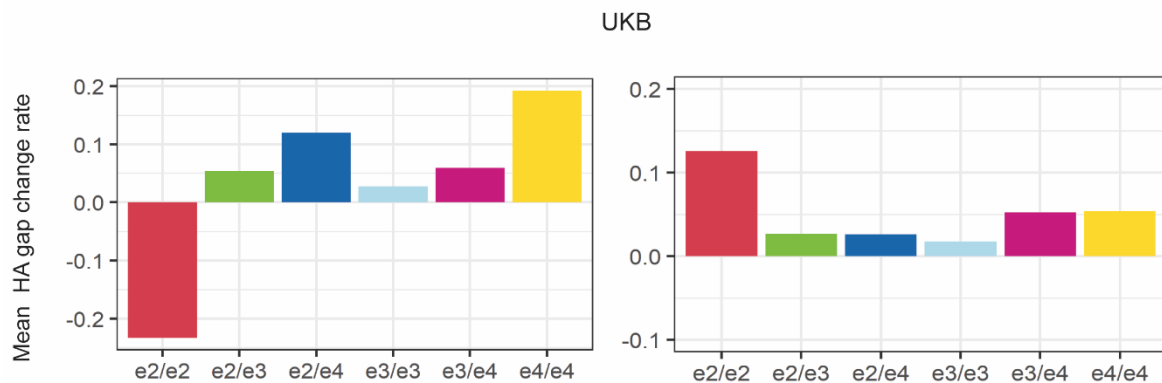


Figure 2. Left and right mean HA gap in cross-sectional analysis and mean HA gap change rate in longitudinal analysis based on APOE genotype. (A) Cross-sectional analysis in UKB and MAS. (B) Longitudinal analysis in UKB. X axis represents six APOE genotypes, and Y axis represents mean HA gap value (cross-sectional) or mean HA gap change rate (longitudinal). UKB: UK Biobank. MAS: Sydney Memory and Ageing Study. HA: hippocampus ROI age.

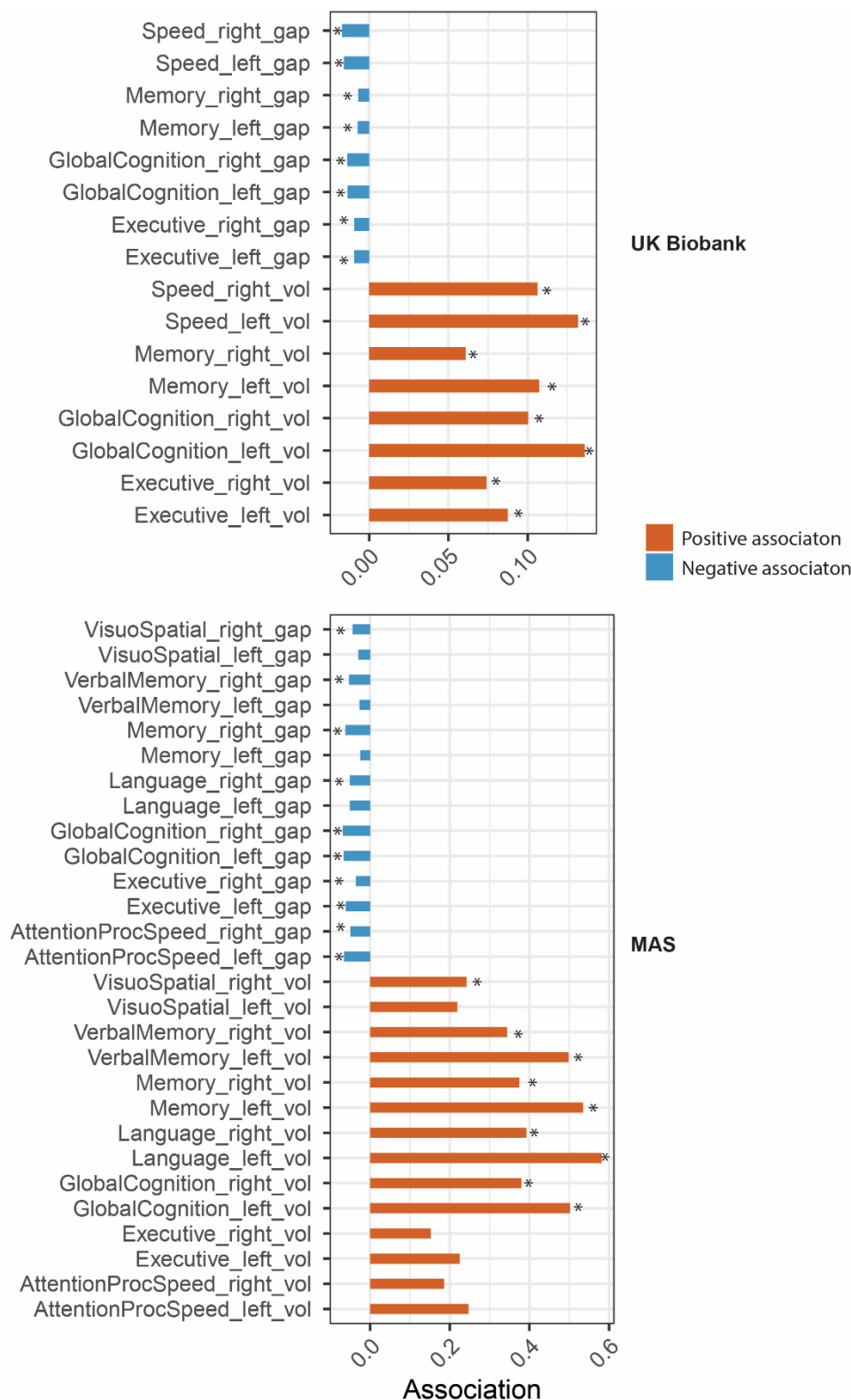


Figure 3. Associations between HA gap (or hippocampus volume) and cognition. Here, "gap" refers to the HA gap, while "vol" denotes hippocampus volume. Asterisks indicate statistically significant associations after false discovery rate (FDR) correction ($p < 0.05$). MAS: Sydney Memory and Ageing Study.

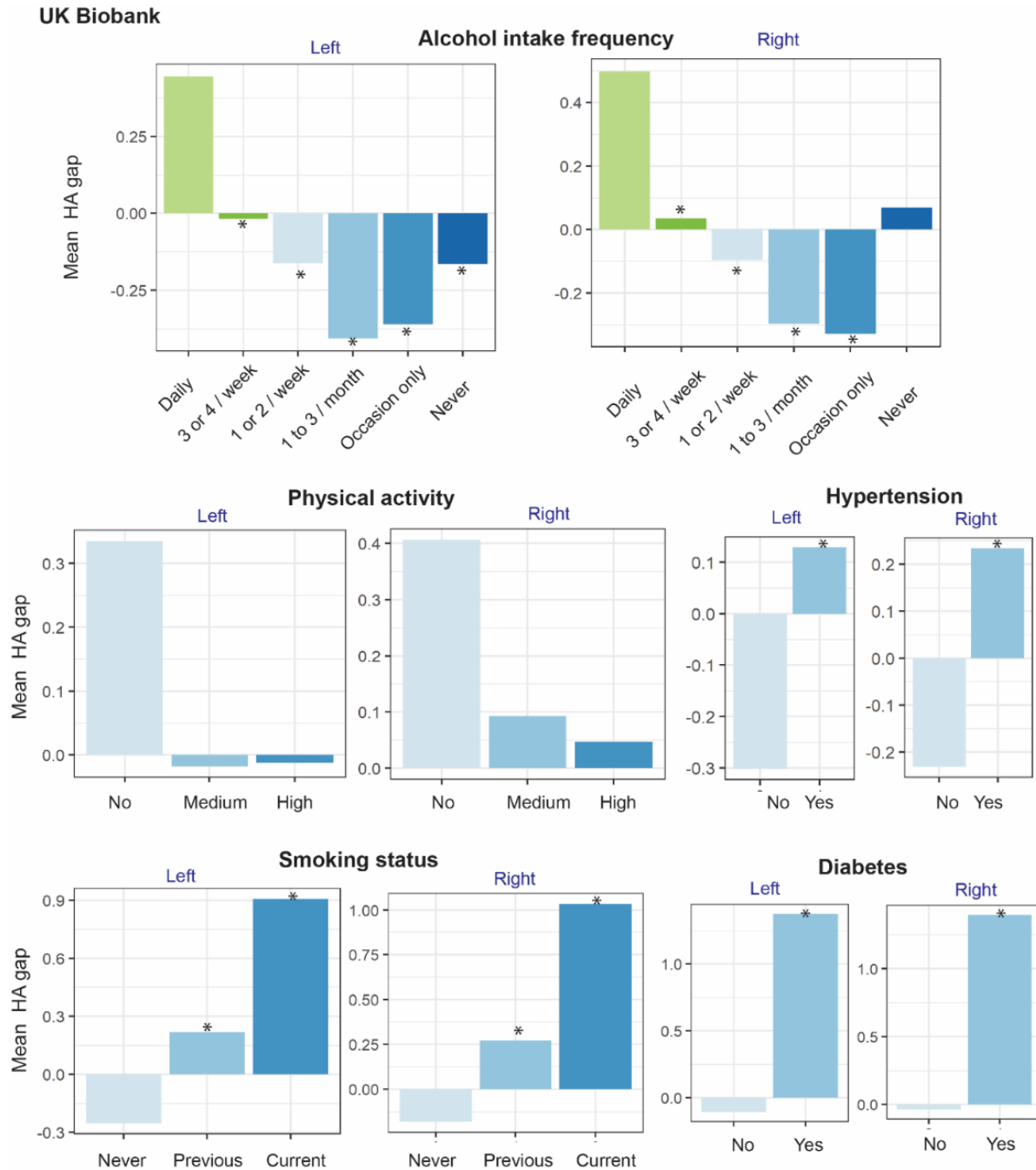


Figure 4. Associations between mean HA gap and other variables in UK Biobank. The X axis represents different groups based on different variables. The Y axis represents the mean left and right HA gap within each group. In each variable category, the first group served as the reference, and subsequent groups were compared with their corresponding reference group. Groups displaying statistically significant differences compared to the reference group were marked with an asterisk (*). Details can be found in Table S3. HA: hippocampus ROI age.

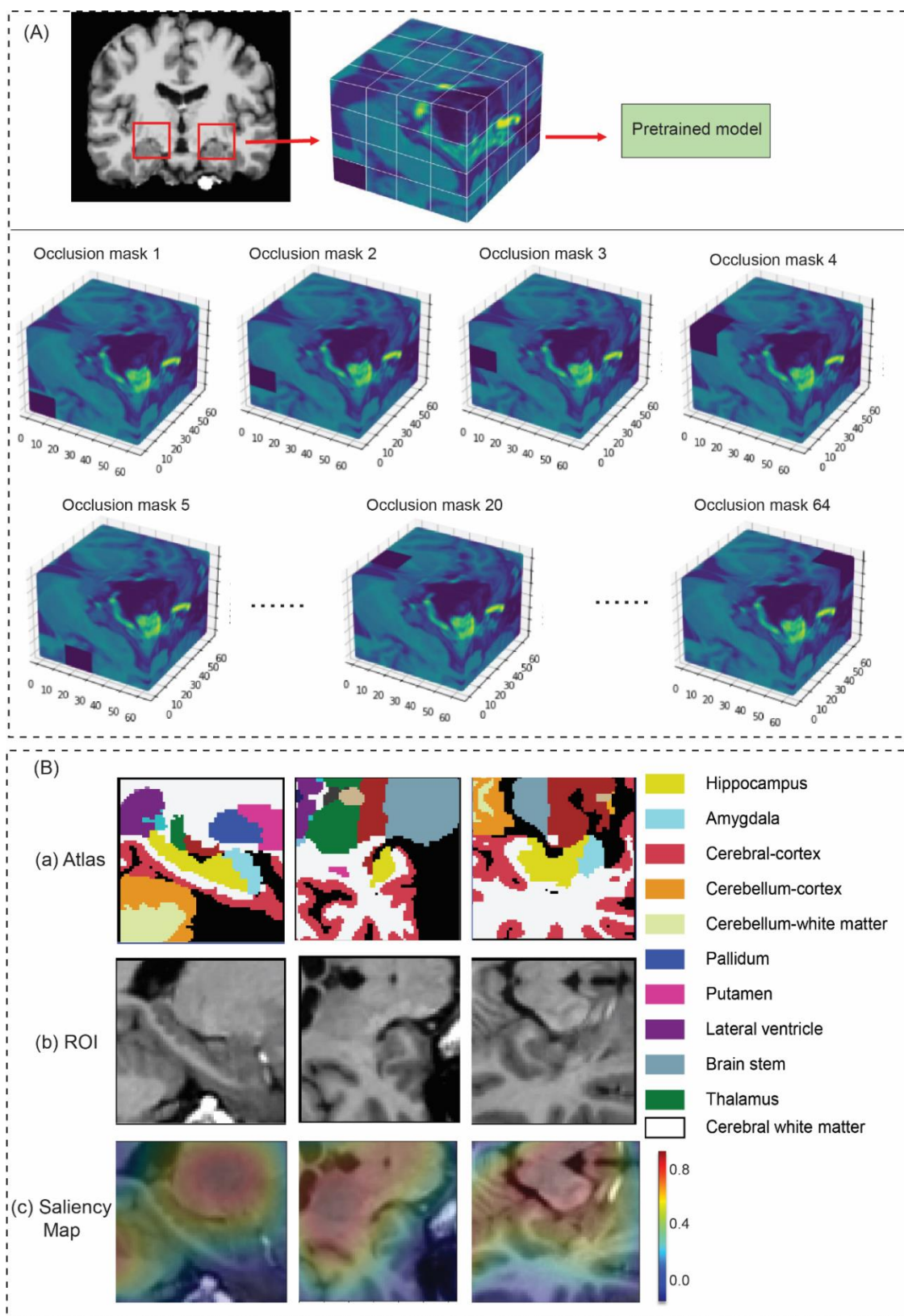


Figure 5. The occlusion analysis. (A) The occlusion analysis framework and the illustration of the different occlusion mask. (B) visualization of hippocampus ROI. The saliency map was produced by occlusion analysis. The warmer colours represent the higher importance of a region in HA prediction. HA: hippocampus ROI age.