Article

Prediction of Chronological Age in Healthy Elderly Subjects with Machine Learning from MRI Brain Segmentation and Cortical Parcellation

Jaime Gómez-Ramírez [1,*] , Miguel A. Fernández-Blázquez [2] and Javier J. González-Rosa [1]

1   Institute of Biomedical Research Cadiz (INiBICA), Universidad de Cádiz, 11003 Cádiz, Spain; javier.rosa@uca.es
2   Department of Biological and Health Psychology, Universidad Autónoma de Madrid, 28049 Madrid, Spain; miguelangel.fernandezb@uam.es
*   Correspondence: jd.gomezramirez@gmail.com

Abstract: Normal aging is associated with changes in volumetric indices of brain atrophy. A quantitative understanding of age-related brain changes can shed light on successful aging. To investigate the effect of age on global and regional brain volumes and cortical thickness, 3514 magnetic resonance imaging scans were analyzed using automated brain segmentation and parcellation methods in elderly healthy individuals (69–88 years of age). The machine learning algorithm extreme gradient boosting (XGBoost) achieved a mean absolute error of 2 years in predicting the age of new subjects. Feature importance analysis showed that the brain-to-intracranial-volume ratio is the most important feature in predicting age, followed by the hippocampi volumes. The cortical thickness in temporal and parietal lobes showed a superior predictive value than frontal and occipital lobes. Insights from this approach that integrate model prediction and interpretation may help to shorten the current explanatory gap between chronological age and biological brain age.

Keywords: aging; MRI; machine learning; XGBoost; feature importance; shapley values; brain segmentation; cortical parcellation; age prediction; biological aging

1. Introduction

Magnetic resonance imaging (MRI) has revolutionized clinical neuroscience, assisting in better diagnostics and playing a crucial role in helping to close the gap between basic and clinical research [1]. The contrast and detail that MRI achieves are unparalleled in detecting brain abnormalities, tumors, and micro-hemorrhages. Technological advancements—in particular, software development—are making it possible to assess conditions that went previously undetected [2]. Furthermore, faster imaging is helping to alleviate the clinical burden in patients and clinicians alike, with AI-based analysis drastically reducing the time required for image reconstruction [3,4]. In addition, the ever-increasing computational capacity and availability of AI techniques are accelerating the use of automated image post-processing for diagnostics and prognosis assessment in clinics and hospitals. Volumetric analysis aiming at quantifying the volume of brain structures and the thickness and gyrification of cortical areas can be particularly effective in flagging brain abnormalities in large datasets [5,6].

MRI is also helping to characterize the neuroanatomy of healthy brain aging across ages and conditions. Several studies show the intricacies of morphological changes visible for the whole brain [7,8], as well as for the cerebral cortex [9], subcortical grey matter structures, and white matter integrity differences [10].

Changes in functional abilities and brain structural alterations, such as atrophy, are expected during aging. However, there is not always a clear separation line that distinguishes the effects of brain aging from neurodegeneration. For example, white matter hyperintensities or microbleeds are present in neurological conditions such as Alzheimer's disease [11],

and they have also been found in aging asymptomatic individuals [12]. Both normal aging and neurodegenerative diseases are accompanied by brain morphological changes, notably atrophy or the loss of tissue volume resulting from cellular death. At advanced stages of the disease, the cellular loss or synaptic pruning associated with atrophy is more easily recognizable than at early stages, making the early detection of neurodegenerative pathologies particularly challenging.

Whole-brain atrophy intended as a diminution of brain volume normalized to the intracranial volume can be measured with T1-weighted structural MR images. Both grey matter and white matter decline with aging, with, at the same time, an increase in the size of the ventricles and cerebrospinal fluid (CSF) volume [13].

Among the subcortical structures in the limbic system and basal ganglia, the hippocampus is arguably the best characterized anatomic volume concerning aging. Hippocampal volume loss is also a well-recognized biomarker for the diagnosis of Alzheimer's disease in cross-sectional [14], longitudinal [15], and meta-analysis [16]. The variation in hippocampal volume as a function of age studied in the UK Biobank Imaging dataset (N = 19,793) shows an acceleration in hippocampal volume reduction in middle age (starting at around 50 years of age) [17]. Age-related atrophy in other subcortical structures varies, with a more rapid decline in the thalamus and putamen compared to caudate and amygdala [18]. However, the software libraries used for automated segmentation can introduce variability in the volume estimates [19].

The "last in first out" hypothesis states that the last brain regions to develop tend to be the first ones to decline. The hypothesis is rooted in the idea of the lifespan composed of two stages: first, a developmental phase, and then followed by an aging phase characterized by progressive cortical thinning, with the onset linked to late-maturing regions of the brain, such as the heteromodal association cortices [20]. Cortical thinning spans widespread cortical regions with unequal effects, e.g., being more prominent in the prefrontal cortex and less so in the parahippocampal cortex [9,21]. Longitudinal estimates of cortical thinning converge with cross-sectional studies that find significant thinning in the heteromodal association cortex with a later expected atrophy in the primary cortex, in line with the "last in first out" hypothesis [22].

Since, as already mentioned, brain volume decline is associated with age, it is possible, at least in principle, to use the volumetric measurement of brain atrophy to estimate its age. It is, however, useful to distinguish between biological and chronological age. The gap between both ages could indicate the aging pace of a given subject; that is to say, a subject with biological age larger than chronological age would imply a faster than expected aging decline. On the other hand, a person more chronologically than biologically aged could indicate that that person is aging more slowly than expected.

Neuroimaging-based studies for chronological age prediction use different features, acquisition techniques, and MRI sequences [23–25]. In [26], a supervised regression model used to predict age using the Destrieux atlas of cortical parcellation, obtained a mean absolute error of 4.05 years [27]. Cortical thickness combined with diffusion MRI has also been used as an input for predicting brain age in multiple regression models [28,29]. Studies comparing the performance between linear and nonlinear modeling approaches (e.g., neural networks) and combining a wide range of features have shown that performance gain from larger training can be limited if data are obtained with different acquisition protocols [30].

"Brain age" is used in the machine learning literature as the age estimate in a regression model [31]. Accordingly, the delta or difference between the brain age (estimated) and the chronological age (given) would be indicative of the health status of the subject. However, for this approach to be sensible, it requires a model producing accurate estimates of the biological age. In [32], a bias-adjustment technique was shown to shorten the delta or difference between the chronological and estimated brain age. However, the pertinence of any approach that tries to reduce the above-mentioned delta will need to be confronted with direct assessments of brain age. Recently, new cellular and molecular approaches for brain

senescence and decline have been put forward. For instance, transcriptome profiling [33], DNA methylation [34–36], or immune metrics such as the inflammatory clock of aging (iAge) [37], aim to capture cellular senescence, holding promise for understanding longevity and neurodegenerative processes.

Although the association between the thinning of cortical areas and aging is recognized, how cortical thinning and overall brain atrophy progress in elderly healthy subjects is an open problem. In this study, we address this issue by building a machine learning model to predict chronological age in a large set of subjects using brain segmentation and cortical parcellation data, collected in a time horizon of 6 years. Different from other studies, we tackled not only the model prediction capabilities but also the model interpretability to assess the importance of the different brain areas, including whole brain and subcortical volumes and the cortical thickness of sulcal and gyral areas for age prediction using SHAP (Shapley additive explanations) [38].

2. Methods

2.1. Study Participants

The dataset used in this study comes from a single-center, observational cohort study [19,39,40]. The participants are home-dwelling elderly volunteers, 69–88 years of age, and without relevant psychiatric, neurological, or systemic disorders. The participants signed informed consent and undertook a yearly systematic clinical assessment, including medical history, neurological and neuropsychological examinations, and brain MRI. Apolipoprotein E (APOE) genotype was also studied, with total DNA isolated from peripheral blood following standard procedures. Ethical approval (CEI PEI 46_2011-v2014) was granted by the Research Ethics Committee of Instituto de Salud Carlos III, and written informed consent was obtained from all of the participants. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation, and with the Helsinki Declaration of 1975 and its later amendments.

Of the initial 1213 subjects, those diagnosed with mild cognitive impairment (MCI) or dementia were excluded, resulting in a cohort of 948 healthy elderly subjects. Cognitive status was determined with the Mini-Mental Status Examination (MMSE), free and cued selective reminding test (FCSRT), semantic fluency, 101 digit-symbol test and Functional Activities Questionnaire (FAQ).

The subjects were assessed yearly for six years, with the number of yearly visits per subject varying from 1 to 6 visits. The distribution of the number of visits with the corresponding neuroradiological and cognitive assessment is as follows: 19.2% of subjects with one initial assessment completed, 11.3% of subjects with 2 assessments completed, 8.5% with 3 assessments completed, 8.9% of subjects with 4 assessments completed, 23% of subjects with 5 assessments, and 29.1% of subjects with all 6 assessments completed.

2.2. MRI Data Acquisition and Preprocessing

For all of the subjects, brain MRIs were collected in 6 yearly visits: 948 in the first visit, 744 in the second, 711 in the third, 622 in the fourth, 529 in the fifth, and 364 in the sixth visit.

The imaging data were acquired on a 3T General Electric scanner (GE Milwaukee) utilizing the following configuration: T1-weighted inversion recovery, flip angle 12°, 3-D pulse sequence, echo time Min. full, time inversion 600 ms, receiver bandwidth 19.23 kHz, field of view = 24.0 cm, slice thickness 1 mm, Freq. × Phase (288 × 288).

The preprocessing of MRI 3 Tesla images consisted of generating an isotropic brain image with non-brain tissue removed [41]. We used the FreeSurfer cortical surface reconstruction pipeline as the initial preprocessing step. The postprocessing was performed with FreeSurfer [42], version freesurfer-darwin-OSX-ElCapitan-dev-20190328-6241d26 running under Mac OS X, product version 10.14.5 as described in [43]. Parcellation was performed using the Destrieux cortical atlas, which is based on the division of the cortex into gyri or sections of the cortex visible on the pial view, and sulci or the hidden parts of the cortex

according to the curvature value of the surface [27]. Figure 1 shows two views of the Destrieux atlas parcellation of the left hemisphere.
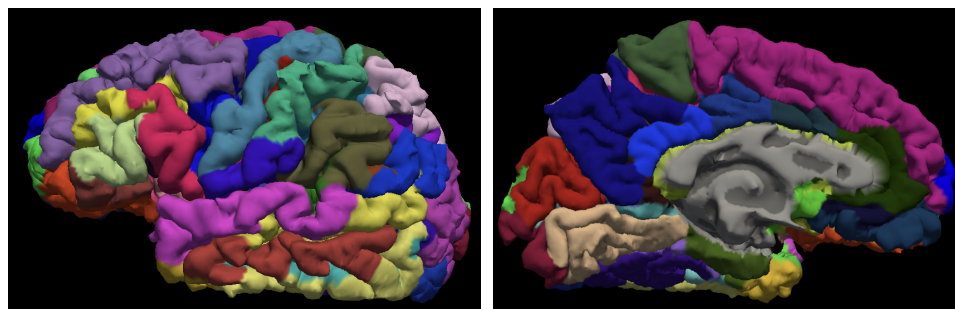
Figure 1. Cortical parcellation of the left hemisphere according to the Destriaux atlas.

2.3. Anomaly Detection with the Isolation Forest Algorithm

In this study we computed the brain volume, the mean subcortical volumes of seven structures—thalamus, putamen, hippocampus, caudate, pallidum, amygdala, accumbens—and the mean cortical thickness of the 148 regions of interest defined in the Destrieux atlas. In the first stage, images were manually assessed for quality, and scans considered not suitable for analysis due to artifacts were discarded. In the second step, we performed anomaly detection using the scikit-learn Isolation Forest implementation [44], resulting in a total of 3514 MRI examinations. The alogrithm's parameters—number of estimators, number of samples, and contamination level—were set to the default values.

Isolation Forest is an ensemble method [45] used as an outlier detector in a variety of datasets [43,46,47]. The algorithm works by isolating each point in the dataset to assess whether the point is an outlier. The idea behind the algorithm is that points that are outliers are easier to separate than those that they are not. Thus, for each point and each feature, a range between the minimum-maximum values is declared. The algorithm randomly selects a feature and a value and, depending on where the value falls in the range, the range is switched upwards to the maximum or downwards to the minimum value. The algorithm proceeds iteratively until the point is isolated; that is, the point is alone inside the range for all features. Finally, the algorithm considers a point as an outlier depending on the number of iterations required to isolate it.

2.4. Statistical Analysis

The sample size of the study (number of MRI scans) and demographic (age, sex, educational attainment level) and genetic information (presence of the APOE4 allele) of each subject in the study is summarized in Table 1. The MRI scans were performed in a time span of 6 years. The number of MRI examinations that passed visual inspection and automatic anomaly detection is as follows: year 1: 857 scans, year 2: 683 scans, year 3: 658 scans, year 4: 555 scans, year 5: 450 scans, year 6: 311 scans, for a total of 3514 MRI examinations.

Table 2 shows the results of the segmentation and cortical parcellation analysis. The brain-to-intracranial-volume ratio (Brain2ICV) in % or brain-to-intracranial-volume ratio and the estimated volume of seven subcortical structures for both hemispheres in mm^3 are followed by the average cortical thickness in mm of the regions defined in the surface-based atlas known as Destrieux atlas [27].

The total intracranial volume (eTIV) estimated by FreeSurfer has previously been reported to have linear correlations of 0.9 with manually estimated intracranial volume [48,49]. Depending on whether CSF is included or not, one can dissociate the total brain volume (TBV) from the total intracranial volume. The normalized TBV is widely used as an index for brain atrophy, as the head size remains stable across the life span and serves as a good measure to reduce between-subject differences with regard to maximum brain size. Whole-brain volume, on the other hand, changes throughout the life span of an individual [50,51]. Measurements of total brain volume (TBV) with FreeSurfer are robust across

field strength [52]. The total intracranial volume acts as a scaffolding of the brain and sets an upper bound for the brain's volume. Accordingly, it is possible to build a proxy of the brain atrophy that an elder person went through in their adult life by computing the ratio between the brain volume (TBV) and the total intracranial volume (eTIV), which represents the upper limit of brain volume [53]. Thus, $Brain2ICV = \frac{TBV}{eTIV}$.

Table 1. The subjects participating in the study had at least one MRI. After anomaly detection, a total of 3514 neuroradiological assessments were analyzed, with the resulting volumetric analysis shown in Table 2. The table shows the description of the demographic and genetic variables included in the study. M: male, F: female, APOE $\epsilon(23)\epsilon(23)$: lacking allele $\epsilon4$, $\epsilon(23)\epsilon4$ one allele $\epsilon4$ and $\epsilon4\epsilon4$ both alleles $\epsilon4$. Educational attainment level, None: no formal education, Primary: primary education degree, Secondary: secondary high school, University: university studies.

SAMPLE SIZE (MRI Scans) = 3514	$N_{SUBJECTS} = 948$
Age	$\mu = 74.68 \pm \sigma = 3.85$
Sex	F 519, M 280
APOE	$\epsilon(23)\epsilon(23) : 145, \epsilon(23)\epsilon4 : 647, \epsilon4\epsilon4 : 7$
Education	None 150, Primary 232, Secondary 203, University 214

Table 2. The table columns from left to right include the name of the subccortical or cortical area, the gray matter volume (mm^3) for subcortical structures, the cortical thickness (mm) defined for areas in the surface cortical atlas used, and the *p*-values of the *T*-test and the ANOVA test for Sex and APOE4 variables, respectively. The first row of data shows the brain-to-intracranial-volume (Brain2ICV) ratio, followed by the subcortical structures: thalamus, putamen, amygdala, pallidum, caudate, hippocampus, and accumbens. Next, the parcellation defined in [27], in which, the cortex is divided into gyral and sulcal regions. The brain areas names on the left column are self-descriptive, with LH and RH referring to each hemisphere: the first letter S, G refers to gyral or sulcal thickness. (<0.05 (*) <0.01 (**)).

	Volume (mm^3)	Sex p-val	APOE PR (>F)
Brain2ICV	$0.698 \pm 0.029(\%)$	**	>0.05
LH Th	6000 ± 653	**	>0.05
RH Th	5848 ± 582	**	>0.05
LH Pu	3551 ± 387	**	>0.05
RH Pu	3635 ± 415	**	>0.05
LH Am	1414 ± 215	**	**
RH Am	1263 ± 193	**	*
LH Pa	1721 ± 252	**	>0.05
RH Pa	1653 ± 244	**	>0.05
LH Ca	3249 ± 490	**	>0.05
RH Ca	3430 ± 252	**	>0.05
LH Hp	3551 ± 387	**	>0.05
RH Hp	3634 ± 415	**	**
LH Ac	462 ± 91	**	>0.05
RH Ac	465 ± 90	**	>0.05
	Avg.Thickness (mm)	Sex (p-val)	APOE Pr (>F)
RH STempSuperior	2.3 ± 0.2	**	>0.05
LH STempSuperior	2.4 ± 0.1	**	>0.05
RH STempSuperior	2.4 ± 0.1	**	>0.05
LH STempInferior	2.5 ± 0.1	**	>0.05
RH STempInferior	2.4 ± 0.1	**	>0.05
LH SOccTempMedandLingual	2.4 ± 0.2	**	>0.05
RH SOccTempMedandLingual	2.4 ± 0.2	**	>0.05
LH SOccTempLat	2.5 ± 0.2	**	>0.05
RH SOccTempLat	2.5 ± 0.2	**	>0.05

Table 2. *Cont.*

LH GTempMid	2.9 ± 0.1	**	>0.05
RH GTempMid	2.9 ± 0.1	>0.05	>0.05
LH GTempInf	2.9 ± 1.6	>0.05	>0.05
RH GTempInf	2.8 ± 1.7	>0.05	>0.05
LH GTempSup	2.5 ± 0.1	**	>0.05
RH GTempSup	2.5 ± 0.1	**	>0.05
LH GTempSupPlanPolar	3.1 ± 0.2	*	>0.05
RH GTempSupPlanPolar	3.0 ± 0.72	**	>0.05
LH GTempSupLateral	3.0 ± 0.1	**	>0.05
RH GTempSupLateral	2.4 ± 1.7	>0.05	>0.05
LH GTempSupTransv	2.4 ± 1.7	**	>0.05
RH GTempSupTransv	2.5 ± 1.8	**	>0.05
RH SIn	2.1 ± 0.1	**	>0.05
LH SFrontSup	2.1 ± 0.1	**	>0.05
RH SFrontSup	2.1 ± 0.1	**	>0.05
LH SFrontMid	2.1 ± 0.1	**	>0.05
RH SFrontMid	2.1 ± 0.1	>0.05	>0.05
LH SFrontInf	2.2 ± 0.1	>0.05	>0.05
RH SFrontInf	2.1 ± 0.1	>0.05	**
LH SFrontSup	2.2 ± 0.1	**	>0.05
RH SFrontSup	2.1 ± 0.1	**	>0.05
LH GFrontSupp	2.6 ± 1.1	**	>0.05
RH GFrontSupp	2.6 ± 1.1	**	>0.05
LH GFrontMid	2.5 ± 1.1	**	>0.05
RH GFrontMid	2.5 ± 1.1	**	>0.05
LH GFrontInfTriangul	2.5 ± 1.6	**	
RH GFrontInfTriangul	2.5 ± 1.4	**	
LH GFrontInfOrbital	2.7 ± 0.2	**	>0.05
RH GFrontInfOrbital	2.6 ± 0.2	**	>0.05
LH GFrontInfOpercular	2.65 ± 0.13	>0.05	>0.05
RH GFrontInfOpercular	2.65 ± 0.13	*	>0.05
LH GCingPostV	2.4 ± 0.3	*	>0.05
RH GCingPostV	2.5 ± 0.3	>0.05	>0.05
LH SCingMarginalis	2.1 ± 0.1	>0.05	>0.05
RH SCingMarginalis	2.1 ± 0.1	**	>0.05
LH SSubParietal	2.2 ± 0.1	**	>0.05
RH SSubParietal	2.3 ± 0.1	**	>0.05
LH SSubOrbital	2.3 ± 0.2	>0.05	>0.05
RH SSubOrbital	2.3 ± 0.3	*	**
LH SPreCentralSuperior	2.3 ± 0.1	**	>0.05
RH SPreCentralSuperior	2.3 ± 0.1	**	>0.05
LH SPreCentralInferior	2.3 ± 0.1	**	>0.05
RH SPreCentralInferior	2.3 ± 0.1	>0.05	>0.05
LH SPostCentral	2.1 ± 0.1	**	>0.05
RH SPostCentral	2.1 ± 0.1	**	>0.05
RH SPeriCallosal	1.8 ± 0.3	**	>0.05
LH SParietoOcc	2.2 ± 0.1	**	>0.05
RH SParietoOcc	2.2 ± 0.1	**	>0.05
LH SOrbMedOlfact	2.1 ± 1.6	**	>0.05
RH SOrbMedOlfact	2.1 ± 1.5	**	>0.05
LH SOrbitalLat	2.1 ± 1.1	**	>0.05
RH SOrbitalLat	2.1 ± 1.1	**	*
LH SOrbitalHShaped	2.6 ± 1.2	**	>0.05
RH SOrbitalHShaped	2.5 ± 1.2	**	>0.05
LH SOccMideandLunatus	2.3 ± 1.2	**	>0.05
RH SOccMideandLunatus	2.3 ± 1.1	**	>0.05
LH SIntraParietandPariettrans	2.1 ± 0.1	**	
RH SIntraParietandPariettrans	2.1 ± 0.1	**	

Table 2. *Cont.*

LH GParietalSup	2.3 ± 1.2	**	>0.05
RH GParietalSup	2.3 ± 1.2	**	>0.05
LH GParietInfSupramar	2.6 ± 1.3	**	>0.05
RH GParietInfSupramar	2.6 ± 1.2	**	>0.05
LH GParietInfAngular	2.5 ± 1.2	**	>0.05
RH GParietInfAngular	2.5 ± 1.3	**	**
LH SCollatTransvPost	2.1 ± 0.1	**	>0.05
RH SCollTatransvPost	2.1 ± 0.1	**	>0.05
LH SCollTransvAnt	2.6 ± 0.2	**	>0.05
RH SCollTransvAnt	2.5 ± 0.2	**	>0.05
LH PoleOcc	3.3 ± 0.2	>0.05	>0.05
RH PoleOcc	3.3 ± 0.2	**	>0.05
LH GOccSup	2.1 ± 1.1	**	>0.05
RH GOccSup	2.4 ± 1.3	**	>0.05
LH GOccMid	2.5 ± 0.1	**	>0.05
RH GOccMid	2.5 ± 0.1	**	>0.05
LH GOccTempMedParahip	3.1 ± 0.6	**	>0.05
RH GOccTempMedParahip	3.2 ± 0.2	**	>0.05
LH GOccTempMedLingual	2.1 ± 0.1	>0.05	>0.05
RH GOccTempMedLingual	2.1 ± 0.1	>0.05	>0.05
LH GOccTempLatFusi	2.8 ± 1.5	>0.05	>0.05
RH GOccTempLatFusi	2.8 ± 1.5	**	>0.05
LH SInsSup	2.4 ± 0.1	>0.05	>0.05
RH SInsSup	2.4 ± 0.1	**	**
LH SInsInf	2.6 ± 0.2	**	>0.05
RH SInsInf	2.5 ± 0.1	>0.05	>0.05
LH SCircInsAnt	2.7 ± 0.2	>0.05	>0.05
RH SCircInsAnt	2.7 ± 0.2	**	>0.05
LH GInsularShort	3.4 ± 0.3	**	>0.05
RH GInsularShort	3.4 ± 0.3	**	>0.05
LH GCentInsula	3.2 ± 0.3	**	>0.05
RH GCentInsula	3.3 ± 0.3	**	>0.05
LH SCentral	2.0 ± 0.1	**	>0.05
RH SCentral	1.9 ± 0.1	**	>0.05
LH GPreCentral	2.6 ± 1.7	**	>0.05
RH GPreCentral	2.6 ± 1.6	**	>0.05
LH GPostCentral	2.1 ± 1.5	**	>0.05
RH GPostCentral	2.1 ± 1.5	**	*
LH SCalcarine	1.9 ± 0.1	>0.05	>0.05
RH SCalcarine	1.9 ± 0.1	**	>0.05
LH GRectus	2.5 ± 0.2	>0.05	>0.05
RH GRectus	2.5 ± 0.2	>0.05	>0.05
LH GOrbital	2.7 ± 1.7	**	>0.05
RH GOrbital	2.7 ± 1.7	**	>0.05
LH GCuneus	1.9 ± 0.1	>0.05	>0.05
RH GCuneus	1.9 ± 0.1	**	>0.05
LH LatFisPost	2.3 ± 0.1	**	>0.05
RH LatFisPost	2.3 ± 0.1	**	*
LH LatFisAntHoriz	2.1 ± 0.1	>0.05	>0.05
RH LatFisAntHoriz	2.4 ± 0.1	**	>0.05

The relationship between volumetric and thickness estimates and the variables Sex and APOE4 is investigated using regression analysis. We conduct hypothesis testing on the regression coefficients obtained in the regression model $Y = \beta_0 + \beta_1 X_{Sex}$, where the target variable Y is the chronological age and X_{Sex} is the sex of the participants.

The effect of the APOE4 allele on volumetric or cortical thickness estimates in asymptomatic individuals is approached using ANOVA analysis with the three-valued variable APOE4.

For the sake of illustration, Figure 2 shows the intracranial volume segmentation and cortical parcellation analysis obtained for four subjects in the study. The summary of the automated segmentation and cortical parcellation of the 3514 scans included in the study are shown in Table 2.

(a)

(b)

(c)

(d)

(e)

(f)
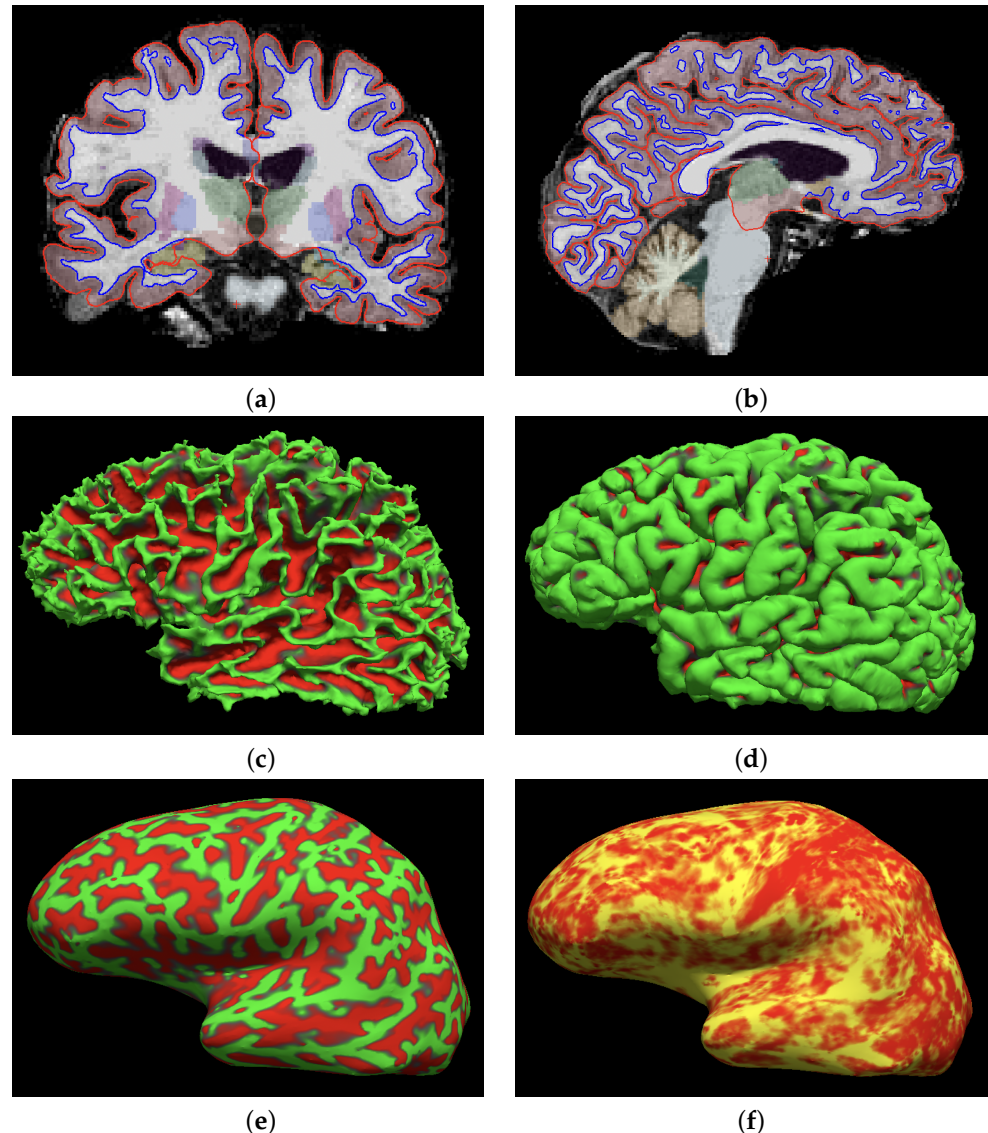
Figure 2. (**a**,**b**) show the coronal and sagittal segmentation results. The edge color blue indicates the demarcation of the white matter surface, and the red edge the pial surface. Plots (**c**–**f**) show the three-dimensional view of the surface analysis for the same subject. (**a**) Coronal view. (**b**) Sagittal view. (**c**) 3D view of white matter surface. (**d**) 3D view of pial surface. (**e**) 3D view of the inflated surface: giry (green) and sulci (red). (**f**) 3D view of the thickness map of the inflated surface.

2.5. Age Prediction Analysis

We aim to predict the chronological age of elderly healthy adults using structural MRI volumetric analysis and demographic features of interest, such as sex, educational attainment level, and the genetic risk factor in dementia (APOE4). We built two types of predicting models, linear and nonlinear, and we assessed their performance by testing the model predictions on a held-out dataset of points not used for training.

Schematically, the prediction problem we aim to resolve can be succinctly described as the supervised learning model shown in Equation (1)

$$\Gamma(X) \rightarrow Y \tag{1}$$

such that the function Γ maps the input space X = *(sex, educational attainment level, APOE4), (subcortical volume estimates), (cortical thickness estimates)* into the output space Y = *(chronological age)*.

The dimensionality of the input space X of the model is 153, including the sex of the individuals (female or male), education level, and the apolipoprotein E gene called APOE4, together with 150 brain imaging features, namely the brain-to-intracranial-volume ratio (Brain2ICV), the volume estimates of seven subcortical structures (caudate, pallidum, putamen, thalamus, hippocampus, amygdala, and accumbens), and the cortical sulci and gyri thickness based on the parcellation defined in the Destrieux cortical atlas (Figure 1).

We built two supervised learning models: partial least squares (PLS) and extreme gradient boosting (XGBoost).

The PLS regression model extracts the components of the input X and the output Y that explain the most shared variance among X and Y [54]. Thus, the two components maximally correlated correspond to the first component in feature space X and the target Y. PLS is based on principal component analysis, so it deals with multi-colinearity in the input space by finding a linear transformation W such that the new input space is linearly independent while, at the same time, maximizing the covariance between Y and X. Thus, X is transformed into X' by a linear transformation W as in $X' = XW$.

The nonlinear supervised learning model of choice is the decision-tree-based ensemble machine learning algorithm called extreme gradient boosting, or XGBoost for short. XGBoost [55] is an implementation of a gradient-boosted trees algorithm that uses a regularized gradient-boosting algorithm to accurately predict the target variable: in our case, the chronological age. XGBoost combines weaker models with stronger ones, which are added to improve the overall performance. During training, the gradient descent algorithm minimizes the loss by adding new trees to predict more accurately than weaker decision trees. The models here are all regression trees mapping the input data points (X) to one of the tree's leaves, which contains the age (Y) that we want to predict. The objective function that the algorithm tries to minimize combines the difference between the predicted and target outputs and a penalty term for model complexity.

Formally, given the dataset tuple (X,Y), the gradient tree boosting algorithm tries to minimize the differentiable loss function L (Equation (2))

$$L(Y, \Gamma(X)) \tag{2}$$

where $\Gamma(X)$ represents an ensemble of n regression trees that are sequentially added to incrementally better predict the residuals of previous trees. Formally,

$$\Gamma_n(X) = \Gamma_{n-1}(X) + \alpha_n g_n(X, \rho_{n-1}) \tag{3}$$

with α_i and ρ_i $i = 1..n$ denoting the regularization and the residual parameter, respectively, for the i_{th} tree. The function g_n is trained to predict the residuals of the precedent tree in the forest (ρ_{i-1}).

Feature importance analysis provides insight into the workings of the predictive model used, allowing for the interpretability of the results. Shapley additive explanations, or SHAP for short, was originally developed in cooperative or coalitional game theory and in contrast to non-cooperative Nash equilibrium models [56]. Shapley values show how much a given feature changes the prediction compared to the prediction at the baseline value of that feature. The Shapley value [38]-based approach is being increasingly used by the machine learning community to deal with the interpretable feature subset selection problem [57].

Formally, the Shapley value Φ is defined via a value function v of all features in a set S. Specifically, the Shapley value of a feature value is its contribution to the payout (e.g., if the

average prediction for all instances is 0.9 and the actual prediction is 0.8, the payout is 0.1) weighted and summed over all possible value combinations.

$$\Phi_j(v) = \sum_{S \in \{x_1,\dots,x_p\}\setminus\{x_j\}} \frac{|S|!(p-|S|-1)!}{p!}(v(S \cup \{x_j\}) - v(S)) \tag{4}$$

where p is the number of features, S is the subset of features, x is the vector of feature values of the particular instance to be explained, and $v(S)$ is the prediction for feature values in S, marginalized over features that are not included in the set S.

The Shapley value is arguably the best permutation-based method for explaining the effects of feature values in the average prediction. An important drawback of Shapley values is that they provide additive contributions (attributions) of explanatory variables. If the model is not additive, then the Shapley values may be misleading. For a more in detail description of SHAP values, see [39] and the references within.

3. Results

We conducted hypothesis testing to investigate the relationship between volumetric and thickness estimates and the variables Sex and APOE4.

In the first case, Sex~Brain volumetric and thickness, the null hypothesis is $H0 : \beta_1 = 0$, and the alternative hypothesis, $HA : \beta_1 \neq 0$. Using the p-value method, H0 is rejected when the p-value of the test statistic is small (e.g., less than 0.05(*) or 0.01(**)). The null hypothesis is rejected for Sex in the Brain2ICV variable, all of the subcortical structures, and in the large majority of cortical thickness areas (Table 2).

Carriage of the APOE4 genotype is the main genetic risk factor for developing late-onset Alzheimer's disease. The relationship between APOE4 and brain volumetric and thickness estimates is investigated using ANOVA, the hypotheses of interest being: $H0 : \mu_1 = \mu_2 = \mu_3$ and $H1$: means are not all equal. The results of the ANOVA analysis with the three-valued variable APOE4 is indicated in Table 2, column "APOE PR(>F)". ANOVA analysis of the APOE4 allele fails to reject the null hypothesis in the large majority of subcortical and cortical regions. These results are not surprising since the volume and thickness were not adjusted for cerebral volume, and effect sizes for sex differences are to be expected in either volume estimates and cortical thicknesses [58,59].

Our results are in agreement with recent studies showing no overall risk effects associated with APOE4 in healthy adults for cortical thickness and subcortical volume [60]. Nevertheless, we find that the APOE4 genotype may have a deleterious effect on hippocampal and amygdala volumes, which is in agreement with the literature on atrophy of the hippocampus and amygdala in healthy elderly and impaired memory individuals [61,62].

Table 3 shows the results for the age predictors built, both linear and nonlinear. The performance is evaluated for the holdout set. The dataset is split into train and test sets. We use 75% of data for training and the remaining 25% for testing the model performance on unseen data.

Table 3. Performance metric of PLS and XGBoost models in the test set (unseen subjects). The performance measures shown in the table are the maximum absolute error (MAE), the maximum residual error (MXE), the mean absolute percentage error (MAPE), and the median absolute error (MEDAE).

| Model | Test Performance Measure | | | |
	MAE	MXE	MAPE	MEDAE
PLS	2.570177	10.2293404	0.03348523	2.15809710
XGBoost	2.0301	8.7138485	0.0265	1.74578

For the partial least squares regression model (PLS), the age of the subjects can be estimated with a maximum residual error (MAE) of 2.57 years. Other metrics, such as

the maximum residual error (MXE), the mean absolute percentage error (MAPE), and the median absolute error (MEDAE), are also shown in Table 3 (first row).

Next, we build, train, and tune using cross validation the nonlinear regressor using the extreme gradient boosting algorithm (XGBoost). The optimization or tuning of the hyperparameters -η, γ, *subsample, colsample by tree, max depth,* and *min child weight* is performed using the grid search method [63].

The hyperparameters η and γ correspond to the learning rate and the minimum loss reduction, respectively. The learning rate, also called the shrinkage rate, η, is used to prevent overfitting. The minimum loss reduction γ acts as a pseudo-regularization hyperparameter in gradient boosting: the higher the γ, the higher the regularization. The hyperparameters *subsample* and *colsample by tree* control the sampling of the dataset at each boosting round: *subsample* is the fraction of observations (rows) and *colsample by tree* is the fraction of features (columns) used to train each tree. The hyperparameters *max depth, min child weight* add constraints on the architecture of the trees: *max depth* is the maximum number of nodes allowed from the root to the farthest leaf (a very large max depth can cause overfitting) and *min child weight* is the minimum weight required in order to create a new node in the tree.

The XGBoost model achieves a mean absolute error (MAE) equals to 2.03, which is a 21% improvement relative to the partial least squares (PLS) model ($MAE = 2.57$) as shown in Table 3 (last row).

For the sake of comparison, the performance metrics of the two models built—PLS and XGBoost—are compared with two dummy models: one always tries to predict the mean and the other the median age of the subjects. The mean absolute error (MAE) achieved by the former is 3.258 (PLS MAE = 2.57, XGboost MAE = 2.030) and the median absolute error (MEDAE) achieved by the latter is 2.850 (PLS MEDAE = 2.158, XGboost MEDAE = 1.745) (Figure 3). The R^2 by the dummy models is, as expected, 0.0, whereas $R^2 = 0.353$ for PLS and $R^2 = 0.5915$ for XGBoost.
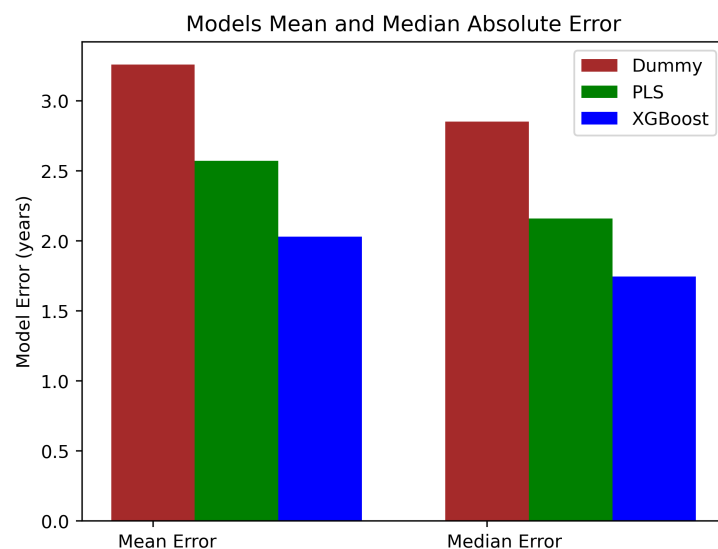
Figure 3. Mean and median absolute error for the dummy, PLS, and XGBoost models. The XGBoost model shows superior results to the PLS.

Our results show that both linear and nonlinear models achieved a maximum mean absolute error of fewer than three years, which is below previous studies. For example, in [26], the mean absolute error achieved using a Gaussian process regression (GPR) algorithm in a sample of 2911 cognitively normal subjects (age 45–91 years) was 4.05. Relevance vector machine algorithms have been used as well to predict brain age, as in [28,64,65], all with a mean absolute error of around 4.5 years. Neuroimaging techniques other than T1-weighted MRI for brain age prediction—for example, diffusion tensor imaging in a cohort of 188

subjects aged 4–85 years—obtained a mean absolute error, depending on the age group and sex, between 6 and 10 years [66]. Neuroanatomical prediction studies of biological maturity using resting-state fMRI [67] in a cohort of 238 young subjects 7–30 years reported 55% of the sample variance, and, in multimodal studies combining MRI and diffusion-weighted imaging to predict child age, nonlinear modeling was able to account for more than 92% of the variance in age [68].

3.1. Feature Importance

In addition to build predictive models of chronological age from age, sex, APOE4 gene, and brain volumetric estimates, it is of particular interest to understand the relative feature importance among predictors. The identification of brain regions and structures may help us to characterize the localized effects of normal brain aging and to use the age predictors as potential biomarkers for neurodegenerative diseases [69,70].

The approach used here for addressing the feature selection problem at hand is Shapley values. Shapley or SHAP values can be used to decompose predictions into the sum of the effect of each feature.

Figure 4 shows the twenty most important features for predicting the chronological age of the subjects in the test set according to the Shapley value method. The most important features are the brain-to-intracranial-volume rate (Brain2ICV) followed by the volume of the hippocampi and the sulcal thickness of the occipito-temporal medial and lingual cortical regions.

Figure 4 shows the relative importance of the features for predicting the chronological age. On the left, Figure 4b, are the absolute Shapley values of the most important features, and on the right, Figure 4b, are the SHAP values for each subject. Thus, Figure 4a is the aggregate plot of Figure 4b. Every dot in Figure 4b represents a different subject in the test set, colored by red or blue if the feature for that subject tends to push towards the right (more age) or to the left (less age) in the model predictions. For example, large (red) values in the top feature (Brain2ICV) decrease the prediction of age; that is, subjects with a larger brain-to-intracranial-volume ratio will tend to be less aged than subjects with small (blue) values. The same can be said for the rest of the features according to the figure, with, however, a few exceptions, such as the thickness of the rectus in the left hemisphere, the thickness of the sulci in the circular anterior insula, and the left caudate volume, which seem to have the opposite effect (big values seem to push towards more age).

The brain-to-total-intracranial-volume ratio or $Brain2ICV = \frac{TBV}{ICV}$ and the hippocampi are, according to SHAP values, the most important variables in predicting chronological age. The key role of the hippocampus in learning and memory makes it a structure of particular interest to study the effects of normal aging in the brain [71,72]. A recent study with a large normative database confirms that hippocampal volume loss accelerates in middle age [17]. Interestingly, the brain-to-intracranial-volume ratio and the hippocampi volume are more important (SHAP value) in predicting chronological age than the cortical thickness of individual regions. We argue that the thickness of cortical areas could be more sensitive to the biological age than it is to the chronological age. How the current results are subject to change when predicting the biological age of the brain using, for example, DNA methylation directly from the brain [35] or the inflammatory aging clock [37], is a matter of future studies [73].

(**a**)

(**b**)
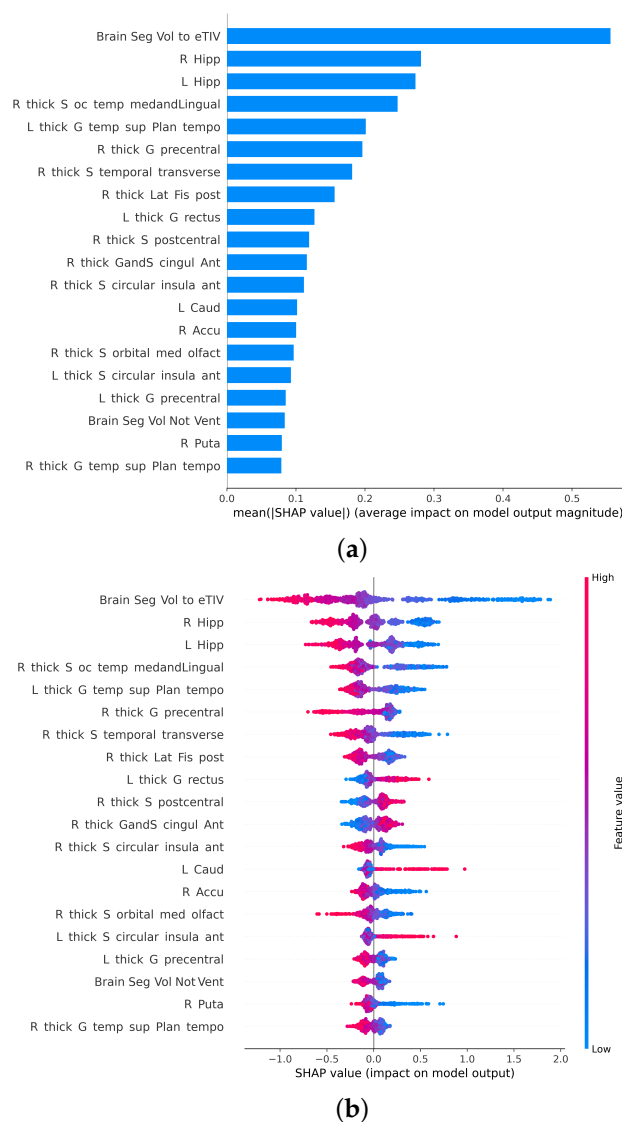
Figure 4. Study of the most important features based on the computation of the SHAP values for each feature and sample of the total of 804 subjects in the test set, both in aggregate (**a**) and for each data point (**b**). The vertical axis of each figure represents the features ranked by importance (top to bottom) calculated as the sum of the SHAP value magnitudes over all samples (horizontal axis). (**a**) Shapley values averaged for all subjects. The most important feature according to the SHAP values is the brain-to-intracranial-volume rate, followed by the volume of the hippocampi. (**b**) Shapley values for each subject. When the point distribution is clustered around 0, it indicates that the feature is unimportant; the more spread the distribution is, the more important the SHAP value is for predicting age.

3.2. Feature Importance of Cortical Gyri and Sulci

The evaluation of the importance of features for predicting chronological age can also be conducted by grouping the cortical areas based on their location (hemisphere and lobe) and type (sulci or gyri). Since the Destriaux atlas used for automated cortical parcellation identifies the location of sulci and gyri in the different brain lobes and the insula, it is possible to plot the relative importance of the cortical areas for predicting chronological age.

Figure 5 depicts the aggregate importance of sulci and gyri cortical areas according to hemispheres Figure 5a and lobes Figure 5b. The aggregate importance is computed as the mean average of the SHAP values normalized. According to the SHAP values, the total sulcus thickness in the right hemisphere contains, on average, more information regarding the chronological age of the subject than gyrus thickness in either hemisphere (0.238, 0.248) and sulcus thickness in the left hemisphere (0.169). Our results indicate that the thickness

of temporal lobe areas is a better predictor of chronological age than the frontal lobe, and, particularly, in cortical sulcal areas. This is in agreement with neuroanatomical evidence that indicates that between one-half to two-thirds of the cortical surface lies in the sulci and the lateral fossa of the brain [27].

Figure 5b shows the relative importance according to the SHAP values of sulci and gyri in the different brain lobes and the insula. The sulci areas in the temporal lobe are more important for predicting age than the rest of the parcellations. It is important to realize that the estimates depicted in the figure must be interpreted only in relative terms and not in absolute terms; that is to say, the sulci in the temporal lobe are more important for predicting chronological age than sulci and gyri in the frontal lobe. Likewise, the total thickness of sulci and gyri in both frontal and parietal lobes contains, according to this analysis, less information regarding the chronological age than the thickness of sulci and gyri in the temporal lobe.

(**a**)
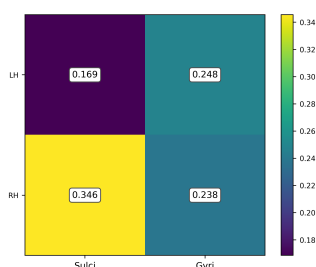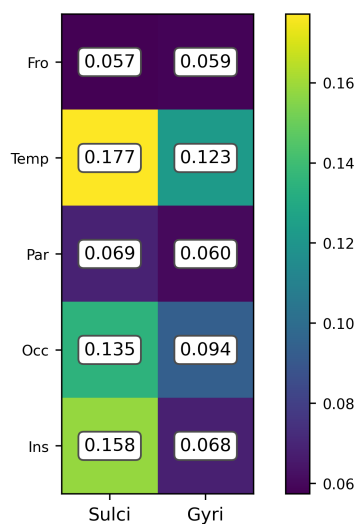
(**b**)

Figure 5. SHAP importance grouping cortical areas by hemisphere, lobe, and type of fold. (**a**) SHAP feature importance in relative terms for brain cortical areas depending on the hemisphere and the cortical surface type (sulci, gyri). As shown in the figure, the aggregate importance of sulci in the right hemisphere for predicting chronological age is 0.346 and is computed as the mean of the SHAP values of right sulci areas normalized by the total of sulci and gyri areas in both hemispheres. (**b**) SHAP feature importance in aggregate for brain cortical areas falling in brain lobes—frontal, occipital, parietal, temporal—and the insula. According to the SHAP values calculated, the temporal lobe contains more information for predicting age than the sulci and gyri-located regions in the other brain lobes.

4. Discussion

It has been suggested that the estimates of brain age from neuroanatomical data may suffer from systematic bias manifested as underestimated brain age for older subjects and overestimated for younger ones [74]. A plausible explanation for this bias may rely on the age distribution with non-normal or skew-normal distributions playing a putative role in introducing a confounding effect [75]. Nevertheless, characteristics of the sample data used to train the models can lead to different estimates of the mean absolute error in age predictions [76], with some studies having a wide range of ages [66,77,78] while others concentrating on a limited range [68].

In a healthy adult population, we may expect both maturation and aging effects in the brain, with an opposing rate of incidence; that is, being maturation prevalent during youth and aging during old age. The cortical thickness declines due to normal aging, with a more visible cortical thinning effect in areas responsible for executive processing tasks and episodic memory retrieval, which are also known to be associated with age-related cognitive decline [79].

The last in, first out hypothesis implies that late-maturing regions of the brain, such as the heteromodal association cortices in the frontal lobe, could be particularly vulnerable to the age-related loss of structural integrity [20]. However, the idea of brain aging following a simple pattern of neural decline with the prefrontal cortex as the region most disrupted is being challenged by studies that suggest a more intricate picture. For example, the prefrontal cortex could play a compensatory role [80] or, as suggested in [81], normal aging could reflect non-specific neural responses rather than the predictable decline of target areas or compensation.

While this study benefits from being single-center using an identical protocol for image acquisition, the generalization of prediction models coming from multi-site datasets is becoming a crucial problem in need of a solution[82]. A promising approach for achieving acceptable harmonization and coherence prediction results is transfer learning [83]. A recent study [84] used a transfer learning approach to build brain age prediction models from diffusion MRI datasets extracted from the CamCAN repository [85] achieved MAE within 4 to 5 years. Although not directly comparable, it is worth noting that the range of chronological ages in our study is within 69-88 years, while in [84] prediction is from an adult population (18 and older) with a smaller dataset (616 samples).

For a comprehensive evaluation of machine learning algorithms for age prediction, see [86]. Support vector regression (SVR) achieved the best accuracy in predicting age in healthy subjects: MAE~3 and MAE~5 years for training and test sets (N=788+88). The performance of the SVR in our dataset is shown in the Supplementary Material. The reduction in the prediction error indicated in our study may respond to several factors, such as the large single-center dataset, the scalability of the tree boosting algorithm, and the mapping from input features to output data. The input space in this study is the product of segmentation and parcellation methods used to extract subcortical volumes and cortical thickness, which is admittedly more computationally demanding than grey matter intensities extracted from the whole brain [87,88].

Lastly, it should be remarked that our model predicts chronological age and not biological brain age, which cannot be directly measured but could, however, be estimated via proxies such as DNA damage.

5. Conclusions

We trained machine learning models to predict chronological age. We find that the non-linear regression modeling extreme gradient boosting (XGBoost) obtains better results than the partial least squares (PLS) model; in particular, XGBoost achieves a mean absolute error of 2 years. Secondly, we find that the best predictor of chronological age is the brain-to-intracranial-volume ratio, followed by the hippocampi volume. Thirdly, the thickness of sulci is more important in predicting age than the thickness of gyri, and this is particularly the case for sulci in the temporal lobe. Our results show that simple volumetric

features, such as the brain-to-intracranial-volume ratio and hippocampal volume, are no less important in predicting chronological age than the cortical thickness of any specific area in the Destriaux atlas. Ultimately, these results enable future research in the gap between the brain's biological and chronological age. The operationalization of this gap using the methodology proposed here may derive into a frailty index for healthy individuals or a potential biomarker for neurodegenerative disorders.

References

1. Granziera, C.; Wuerfel, J.; Barkhof, F.; Calabrese, M.; De Stefano, N.; Enzinger, C.; Evangelou, N.; Filippi, M.; Geurts, J.J.; Reich, D.S.; et al. Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain* **2021**, *144*, 1296–1311. [CrossRef] [PubMed]
2. Harisinghani, M.G.; O'Shea, A.; Weissleder, R. Advances in clinical MRI technology. *Sci. Transl. Med.* **2019**, *11*, eaba2591. [CrossRef] [PubMed]
3. Prakkamakul, S.; Witzel, T.; Huang, S.; Boulter, D.; Borja, M.J.; Schaefer, P.; Rosen, B.; Heberlein, K.; Ratai, E.; Gonzalez, G.; et al. Ultrafast brain MRI: Clinical deployment and comparison to conventional brain MRI at 3T. *J. Neuroimaging* **2016**, *26*, 503–510. [CrossRef] [PubMed]
4. Zhao, X.; Duan, G.; Wu, K.; Anderson, S.W.; Zhang, X. Intelligent metamaterials based on nonlinearity for magnetic resonance imaging. *Adv. Mater.* **2019**, *31*, 1905461. [CrossRef]
5. Gudigar, A.; Raghavendra, U.; Hegde, A.; Kalyani, M.; Ciaccio, E.J.; Acharya, U.R. Brain pathology identification using computer aided diagnostic tool: A systematic review. *Comput. Methods Programs Biomed.* **2020**, *187*, 105205. [CrossRef]
6. Gauriau, R.; Bizzo, B.C.; Kitamura, F.C.; Landi Junior, O.; Ferraciolli, S.F.; Macruz, F.B.; Sanchez, T.A.; Garcia, M.R.; Vedolin, L.M.; Domingues, R.C.; et al. A Deep Learning-Based Model for Detecting Abnormalities on Brain MRI for Triaging: Preliminary Results from a Multi-Site Experience. *Radiol. Artif. Intell.* **2021**, *3*, e200184. [CrossRef]
7. MacDonald, M.E.; Pike, G.B. MRI of healthy brain aging: A review. *NMR Biomed.* **2021**, *3*, e4564. [CrossRef]
8. Royle, N.A.; Booth, T.; Hernández, M.C.V.; Penke, L.; Murray, C.; Gow, A.J.; Maniega, S.M.; Starr, J.; Bastin, M.E.; Deary, I.J.; et al. Estimated maximal and current brain volume predict cognitive ability in old age. *Neurobiol. Aging* **2013**, *34*, 2726–2733. [CrossRef]
9. Salat, D.H.; Buckner, R.L.; Snyder, A.Z.; Greve, D.N.; Desikan, R.S.; Busa, E.; Morris, J.C.; Dale, A.M.; Fischl, B. Thinning of the cerebral cortex in aging. *Cereb. Cortex* **2004**, *14*, 721–730. [CrossRef]

10. Bennett, I.J.; Greenia, D.E.; Maillard, P.; Sajjadi, S.A.; DeCarli, C.; Corrada, M.M.; Kawas, C.H. Age-related white matter integrity differences in oldest-old without dementia. *Neurobiol. Aging* **2017**, *56*, 108–114. [CrossRef]

11. Shams, S.; Martola, J.; Cavallin, L.; Granberg, T.; Shams, M.; Aspelin, P.; Wahlund, L.; Kristoffersen-Wiberg, M. SWI or T2*: Which MRI sequence to use in the detection of cerebral microbleeds? The Karolinska Imaging Dementia Study. *Am. J. Neuroradiol.* **2015**, *36*, 1089–1095. [CrossRef] [PubMed]

12. Prins, N.D.; Scheltens, P. White matter hyperintensities, cognitive impairment and dementia: An update. *Nat. Rev. Neurol.* **2015**, *11*, 157–165. [CrossRef] [PubMed]

13. Grant, R.; Condon, B.; Lawrence, A.; Hadley, D.; Patterson, J.; Bone, I.; Teasdale, G. Human cranial CSF volumes measured by MRI: Sex and age influences. *Magn. Reson. Imaging* **1987**, *5*, 465–468. [CrossRef]

14. Maruszak, A.; Thuret, S. Why looking at the whole hippocampus is not enough—a critical role for anteroposterior axis, subfield and activation analyses to enhance predictive value of hippocampal changes for Alzheimer's disease diagnosis. *Front. Cell. Neurosci.* **2014**, *8*, 95. [CrossRef] [PubMed]

15. McRae-McKee, K.; Evans, S.; Hadjichrysanthou, C.; Wong, M.; De Wolf, F.; Anderson, R. Combining hippocampal volume metrics to better understand Alzheimer's disease progression in at-risk individuals. *Sci. Rep.* **2019**, *9*, 1–9. [CrossRef]

16. Whitwell, J.L.; Wiste, H.J.; Weigand, S.D.; Rocca, W.A.; Knopman, D.S.; Roberts, R.O.; Boeve, B.F.; Petersen, R.C.; Jack, C.R.; Initiative, A.D.N.; et al. Comparison of imaging biomarkers in the Alzheimer disease neuroimaging initiative and the Mayo Clinic Study of Aging. *Arch. Neurol.* **2012**, *69*, 614–622. [CrossRef]

17. Nobis, L.; Manohar, S.G.; Smith, S.M.; Alfaro-Almagro, F.; Jenkinson, M.; Mackay, C.E.; Husain, M. Hippocampal volume across age: Nomograms derived from over 19,700 people in UK Biobank. *NeuroImage Clin.* **2019**, *23*, 101904. [CrossRef]

18. Serbruyns, L.; Leunissen, I.; Huysmans, T.; Cuypers, K.; Meesen, R.L.; van Ruitenbeek, P.; Sijbers, J.; Swinnen, S.P. Subcortical volumetric changes across the adult lifespan: subregional thalamic atrophy accounts for age-related sensorimotor performance declines. *Cortex* **2015**, *65*, 128–138. [CrossRef]

19. Gomez-Ramirez, J.; Quilis-Sancho, J.; Fernandez-Blazquez, M.A. A Comparative Analysis of MRI Automated Segmentation of Subcortical Brain Volumes in a Large Dataset of Elderly Subjects. *Neuroinformatics* **2021**, 1–10. [CrossRef]

20. McGinnis, S.M.; Brickhouse, M.; Pascual, B.; Dickerson, B.C. Age-related changes in the thickness of cortical zones in humans. *Brain Topogr.* **2011**, *24*, 279–291. [CrossRef]

21. Sowell, E.R.; Peterson, B.S.; Thompson, P.M.; Welcome, S.E.; Henkenius, A.L.; Toga, A.W. Mapping cortical change across the human life span. *Nat. Neurosci.* **2003**, *6*, 309. [CrossRef] [PubMed]

22. Shaw, M.E.; Sachdev, P.S.; Anstey, K.J.; Cherbuin, N. Age-related cortical thinning in cognitively healthy individuals in their 60s: The PATH Through Life study. *Neurobiol. Aging* **2016**, *39*, 202–209. [CrossRef] [PubMed]

23. Li, W.; Wang, M.; Li, Y.; Huang, Y.; Chen, X. A novel brain network construction method for exploring age-related functional reorganization. *Comput. Intell. Neurosci.* **2016**, *2016*, 2429691. [CrossRef] [PubMed]

24. Jónsson, B.A.; Bjornsdottir, G.; Thorgeirsson, T.; Ellingsen, L.M.; Walters, G.B.; Gudbjartsson, D.; Stefansson, H.; Stefansson, K.; Ulfarsson, M. Brain age prediction using deep learning uncovers associated sequence variants. *Nat. Commun.* **2019**, *10*, 1–10. [CrossRef]

25. Mendes, S.L.; Pinaya, W.H.L.; Pan, P.; Sato, J.R. Estimating Gender and Age from Brain Structural MRI of Children and Adolescents: A 3D Convolutional Neural Network Multitask Learning Model. *Comput. Intell. Neurosci.* **2021**, *2021*, 5550914. [CrossRef]

26. Aycheh, H.M.; Seong, J.K.; Shin, J.H.; Na, D.L.; Kang, B.; Seo, S.W.; Sohn, K.A. Biological brain age prediction using cortical thickness data: A large scale cohort study. *Front. Aging Neurosci.* **2018**, *10*, 252. [CrossRef]

27. Destrieux, C.; Fischl, B.; Dale, A.; Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* **2010**, *53*, 1–15. [CrossRef]

28. Kondo, C.; Ito, K.; Wu, K.; Sato, K.; Taki, Y.; Fukuda, H.; Aoki, T. An age estimation method using brain local features for T1-weighted images. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 666–669.

29. Cherubini, A.; Caligiuri, M.E.; Péran, P.; Sabatini, U.; Cosentino, C.; Amato, F. Importance of multimodal MRI in characterizing brain tissue and its potential application for individual age prediction. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 1232–1239. [CrossRef]

30. MacDonald, M.E.; Williams, R.J.; Forkert, N.D.; Berman, A.J.; McCreary, C.R.; Frayne, R.; Pike, G.B. Interdatabase variability in cortical thickness measurements. *Cereb. Cortex* **2019**, *29*, 3282–3293. [CrossRef]

31. Cole, J.H.; Ritchie, S.J.; Bastin, M.E.; Hernández, V.; Muñoz Maniega, S.; Royle, N.; Corley, J.; Pattie, A.; Harris, S.E.; Zhang, Q.; et al. Brain age predicts mortality. *Mol. Psychiatry* **2018**, *23*, 1385–1392. [CrossRef]

32. Beheshti, I.; Nugent, S.; Potvin, O.; Duchesne, S. Bias-adjustment in neuroimaging-based brain age frameworks: A robust scheme. *NeuroImage Clin.* **2019**, *24*, 102063. [CrossRef] [PubMed]

33. Ham, S.; Lee, S.J.V. Advances in transcriptome analysis of human brain aging. *Exp. Mol. Med.* **2020**, *52*, 1787–1797. [CrossRef] [PubMed]

34. Hernandez, D.G.; Nalls, M.A.; Gibbs, J.R.; Arepalli, S.; van der Brug, M.; Chong, S.; Moore, M.; Longo, D.L.; Cookson, M.R.; Traynor, B.J.; et al. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum. Mol. Genet.* **2011**, *20*, 1164–1172. [CrossRef] [PubMed]

35. Horvath, S.; Zhang, Y.; Langfelder, P.; Kahn, R.S.; Boks, M.P.; van Eijk, K.; van den Berg, L.H.; Ophoff, R.A. Aging effects on DNA methylation modules in human brain and blood tissue. *Genome Biol.* **2012**, *13*, 1–18. [CrossRef] [PubMed]

36. Hannum, G.; Guinney, J.; Zhao, L.; Zhang, L.; Hughes, G.; Sadda, S.; Klotzle, B.; Bibikova, M.; Fan, J.B.; Gao, Y.; et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **2013**, *49*, 359–367. [CrossRef]

37. Sayed, N.; Huang, Y.; Nguyen, K.; Krejciova-Rajaniemi, Z.; Grawe, A.P.; Gao, T.; Tibshirani, R.; Hastie, T.; Alpert, A.; Cui, L.; et al. An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. *Nat. Aging* **2021**, *1*, 598–615. [CrossRef] [PubMed]

38. Shapley, L.S. A value for n-person games. *Contrib. Theory Games* **1953**, *2*, 307–317.

39. Gómez-Ramírez, J.; Ávila-Villanueva, M.; Fernández-Blázquez, M.Á. Selecting the most important self-assessed features for predicting conversion to Mild Cognitive Impairment with Random Forest and Permutation-based methods. *Sci. Rep.* **2020**, *10*, 1–15. [CrossRef] [PubMed]

40. Sanz-Blasco, R.; Ruiz-Sánchez de León, J.M.; Ávila-Villanueva, M.; Valentí-Soler, M.; Gómez-Ramírez, J.; Fernández-Blázquez, M.A. Transition from mild cognitive impairment to normal cognition: Determining the predictors of reversion with multi-state Markov models. *Alzheimer Dement.* **2021**. [CrossRef]

41. McCarthy, C.S.; Ramprashad, A.; Thompson, C.; Botti, J.A.; Coman, I.L.; Kates, W.R. A comparison of FreeSurfer-generated data with and without manual intervention. *Front. Neurosci.* **2015**, *9*, 379. [CrossRef]

42. Fischl, B. FreeSurfer. *NeuroImage* **2012**, *62*, 774–781. [CrossRef] [PubMed]

43. Gómez-Ramírez, J.; González-Rosa, J.J. Intra-and interhemispheric symmetry of subcortical brain structures: a volumetric analysis in the aging human brain. *Brain Struct. Funct.* **2022**, *227*, 451–462. [CrossRef] [PubMed]

44. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

45. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]

46. Domingues, R.; Filippone, M.; Michiardi, P.; Zouaoui, J. A comparative evaluation of outlier detection algorithms: Experiments and analyses. *Pattern Recognit.* **2018**, *74*, 406–421. [CrossRef]

47. Alaverdyan, Z. Unsupervised Representation Learning for Anomaly Detection on Neuroimaging. Application to Epilepsy Lesion Detection on Brain MRI. Ph.D. Thesis, Université de Lyon, Lyon, France, 2019.

48. Shen, L.; Saykin, A.J.; Kim, S.; Firpi, H.A.; West, J.D.; Risacher, S.L.; McDonald, B.C.; McHugh, T.L.; Wishart, H.A.; Flashman, L.A. Comparison of manual and automated determination of hippocampal volumes in MCI and early AD. *Brain Imaging Behav.* **2010**, *4*, 86–95. [CrossRef]

49. Malone, I.B.; Leung, K.K.; Clegg, S.; Barnes, J.; Whitwell, J.L.; Ashburner, J.; Fox, N.C.; Ridgway, G.R. Accurate automatic estimation of total intracranial volume: A nuisance variable with less nuisance. *NeuroImage* **2015**, *104*, 366–372. [CrossRef]

50. Wolf, H.; Julin, P.; Gertz, H.J.; Winblad, B.; Wahlund, L.O. Intracranial volume in mild cognitive impairment, Alzheimer's disease and vascular dementia: Evidence for brain reserve? *Int. J. Geriatr. Psychiatry* **2004**, *19*, 995–1007. [CrossRef]

51. Caspi, Y.; Brouwer, R.M.; Schnack, H.G.; van de Nieuwenhuijzen, M.E.; Cahn, W.; Kahn, R.S.; Niessen, W.J.; van der Lugt, A.; Pol, H.H. Changes in the intracranial volume from early adulthood to the sixth decade of life: A longitudinal study. *NeuroImage* **2020**, *220*, 116842. [CrossRef]

52. Heinen, R.; Bouvy, W.H.; Mendrik, A.M.; Viergever, M.A.; Biessels, G.J.; De Bresser, J. Robustness of automated methods for brain volume measurements across different MRI field strengths. *PLoS ONE* **2016**, *11*, e0165719.

53. Gómez-Ramírez, J.; Fernandez-Blazquez, M.A.; González-Rosa, J. The aging human brain: A causal analysis of the effect of sex and age on brain volume. *bioRxiv* **2020**. . [CrossRef]

54. Garthwaite, P.H. An interpretation of partial least squares. *J. Am. Stat. Assoc.* **1994**, *89*, 122–127. [CrossRef]

55. Chen, T.; He, T.; Benesty, M.; Khotilovich, V.; Tang, Y.; Cho, H. *Xgboost: Extreme Gradient Boosting*; R Package Version 0.4-2; 2015, Volume 1, pp. 1–4.

56. Ritzberger, K. *Foundations of Non-Cooperative Game Theory*; OUP Catalogue; Oxford University Press: Oxford, UK, 2002.

57. Tripathi, S.; Hemachandra, N.; Trivedi, P. Interpretable feature subset selection: A Shapley value based approach. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5463–5472.

58. Leonard, C.M.; Towler, S.; Welcome, S.; Halderman, L.K.; Otto, R.; Eckert, M.A.; Chiarello, C. Size matters: Cerebral volume influences sex differences in neuroanatomy. *Cereb. Cortex* **2008**, *18*, 2920–2931. [CrossRef] [PubMed]

59. Lotze, M.; Domin, M.; Gerlach, F.H.; Gaser, C.; Lueders, E.; Schmidt, C.O.; Neumann, N. Novel findings from 2838 adult brains on sex differences in gray matter brain volume. *Sci. Rep.* **2019**, *9*, 1–7. [CrossRef]

60. Mole, J.P.; Fasano, F.; Evans, J.; Sims, R.; Kidd, E.; Aggleton, J.P.; Metzler-Baddeley, C. APOE-ε4-related differences in left thalamic microstructure in cognitively healthy adults. *Sci. Rep.* **2020**, *10*, 1–25. [CrossRef]

61. Fleisher, A.; Grundman, M.; Jack, Clifford R., J.; Petersen, R.C.; Taylor, C.; Kim, H.T.; Schiller, D.H.B.; Bagwell, V.; Sencakova, D.; Weiner, M.F.; et al. Sex, Apolipoprotein E e4 Status, and Hippocampal Volume in Mild Cognitive Impairment. *Arch. Neurol.* **2005**, *62*, 953–957. doi: [CrossRef]

62. Orihashi, R.; Mizoguchi, Y.; Imamura, Y.; Yamada, S.; Ueno, T.; Monji, A. Oxytocin and elderly MRI-based hippocampus and amygdala volume: A 7-year follow-up study. *Brain Commun.* **2020**, *2*, fcaa081. [CrossRef]

63. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.

64. Franke, K.; Ziegler, G.; Klöppel, S.; Gaser, C.; Initiative, A.D.N. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *NeuroImage* **2010**, *50*, 883–892. [CrossRef]

65. Wang, J.; Li, W.; Miao, W.; Dai, D.; Hua, J.; He, H. Age estimation using cortical surface pattern combining thickness with curvatures. *Med Biol. Eng. Comput.* **2014**, *52*, 331–341. [CrossRef]

66. Mwangi, B.; Hasan, K.M.; Soares, J.C. Prediction of individual subject's age across the human lifespan using diffusion tensor imaging: A machine learning approach. *NeuroImage* **2013**, *75*, 58–67. [CrossRef] [PubMed]

67. Dosenbach, N.U.; Nardos, B.; Cohen, A.L.; Fair, D.A.; Power, J.D.; Church, J.A.; Nelson, S.M.; Wig, G.S.; Vogel, A.C.; Lessov-Schlaggar, C.N.; et al. Prediction of individual brain maturity using fMRI. *Science* **2010**, *329*, 1358–1361. [CrossRef]

68. Brown, T.T.; Kuperman, J.M.; Chung, Y.; Erhart, M.; McCabe, C.; Hagler, D.J., Jr.; Venkatraman, V.K.; Akshoomoff, N.; Amaral, D.G.; Bloss, C.S.; et al. Neuroanatomical assessment of biological maturity. *Curr. Biol.* **2012**, *22*, 1693–1698. [CrossRef] [PubMed]

69. Gomez-Ramirez, J.; Wu, J. Network-based biomarkers in Alzheimer's disease: Review and future directions. *Front. Aging Neurosci.* **2014**, *6*, 12. [CrossRef] [PubMed]

70. Eckerström, C.; Klasson, N.; Olsson, E.; Selnes, P.; Rolstad, S.; Wallin, A. Similar pattern of atrophy in early-and late-onset Alzheimer's disease. *Alzheimer Dementia Diagn. Assess. Dis. Monit.* **2018**, *10*, 253–259. [CrossRef]

71. Kaye, J.A.; Swihart, T.; Howieson, D.; Dame, A.; Moore, M.; Karnos, T.; Camicioli, R.; Ball, M.; Oken, B.; Sexton, G. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. *Neurology* **1997**, *48*, 1297–1304. [CrossRef] [PubMed]

72. Bettio, L.E.; Rajendran, L.; Gil-Mohapel, J. The effects of aging in the hippocampus and cognitive decline. *Neurosci. Biobehav. Rev.* **2017**, *79*, 66–86. [CrossRef]

73. Schultz, M.B.; Kane, A.E.; Mitchell, S.J.; MacArthur, M.R.; Warner, E.; Vogel, D.S.; Mitchell, J.R.; Howlett, S.E.; Bonkowski, M.S.; Sinclair, D.A. Age and life expectancy clocks based on machine learning analysis of mouse frailty. *Nat. Commun.* **2020**, *11*, 1–12. [CrossRef]

74. Niu, X.; Zhang, F.; Kounios, J.; Liang, H. Improved prediction of brain age using multimodal neuroimaging data. *Hum. Brain Mapp.* **2020**, *41*, 1626–1643. [CrossRef]

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

76. Pardoe, H.R.; Kuzniecky, R. NAPR: A cloud-based framework for neuroanatomical age prediction. *Neuroinformatics* **2018**, *16*, 43–49. [CrossRef] [PubMed]

77. Schnack, H.G.; Van Haren, N.E.; Nieuwenhuis, M.; Hulshoff Pol, H.E.; Cahn, W.; Kahn, R.S. Accelerated brain aging in schizophrenia: A longitudinal pattern recognition study. *Am. J. Psychiatry* **2016**, *173*, 607–616. [CrossRef] [PubMed]

78. Cole, J.H.; Poudel, R.P.; Tsagkrasoulis, D.; Caan, M.W.; Steves, C.; Spector, T.D.; Montana, G. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* **2017**, *163*, 115–124. [CrossRef] [PubMed]

79. Habeck, C.; Gazes, Y.; Razlighi, Q.; Stern, Y. Cortical thickness and its associations with age, total cognition and education across the adult lifespan. *PLoS ONE* **2020**, *15*, e0230298. [CrossRef]

80. Cabeza, R.; Nyberg, L.; Park, D.C. *Cognitive Neuroscience of Aging: Linking Cognitive and Cerebral Aging*; Oxford University Press: Oxford, UK, 2016.

81. Morcom, A.M.; Henson, R.N. Increased prefrontal activity with aging reflects nonspecific neural responses rather than compensation. *J. Neurosci.* **2018**, *38*, 7303–7313. [CrossRef]

82. Smith, E.E.; Biessels, G.J.; De Guio, F.; De Leeuw, F.E.; Duchesne, S.; Düring, M.; Frayne, R.; Ikram, M.A.; Jouvent, E.; MacIntosh, B.J.; et al. Harmonizing brain magnetic resonance imaging methods for vascular contributions to neurodegeneration. *Alzheimer Dementia Diagn. Assess. Dis. Monit.* **2019**, *11*, 191–204. [CrossRef]

83. Valverde, J.M.; Imani, V.; Abdollahzadeh, A.; De Feo, R.; Prakash, M.; Ciszek, R.; Tohka, J. Transfer learning in magnetic resonance brain imaging: A systematic review. *J. Imaging* **2021**, *7*, 66. [CrossRef]

84. Chen, C.L.; Hsu, Y.C.; Yang, L.Y.; Tung, Y.H.; Luo, W.B.; Liu, C.M.; Hwang, T.J.; Hwu, H.G.; Tseng, W.Y.I. Generalization of diffusion magnetic resonance imaging–based brain age prediction model through transfer learning. *NeuroImage* **2020**, *217*, 116831. [CrossRef]

85. Shafto, M.A.; Tyler, L.K.; Dixon, M.; Taylor, J.R.; Rowe, J.B.; Cusack, R.; Calder, A.J.; Marslen-Wilson, W.D.; Duncan, J.; Dalgleish, T.; 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.* **2014**, *14*, 1–25. [CrossRef]

86. Beheshti, I.; Ganaie, M.; Paliwal, V.; Rastogi, A.; Razzak, I.; Tanveer, M. Predicting brain age using machine learning algorithms: A comprehensive evaluation. *IEEE J. Biomed. Health Inform.* **2021**, *26*, 1432–1440. [CrossRef]

87. Cole, J.H.; Leech, R.; Sharp, D.J.; Initiative, A.D.N. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Ann. Neurol.* **2015**, *77*, 571–581. [CrossRef] [PubMed]

88. Wang, J.; Knol, M.J.; Tiulpin, A.; Dubost, F.; de Bruijne, M.; Vernooij, M.W.; Adams, H.H.; Ikram, M.A.; Niessen, W.J.; Roshchupkin, G.V. Gray matter age prediction as a biomarker for risk of dementia. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21213–21218. [CrossRef] [PubMed]