Multimodal brain-age prediction and cardiovascular risk: The Whitehall II MRI sub-study

Ann-Marie G. de Lange[*a,b,c], Melis Anatürk[a,d], Tobias Kaufmann[c], James H. Cole[e,f], Ludovica Griffanti[d], Enikő Zsoldos[a,d], Daria Jensen[a,d], Sana Suri[a,d], Nicola Filippini[a,d], Archana Singh-Manoux[g,h], Mika Kivimäki[h], Lars T. Westlye[b,c], Klaus P. Ebmeier[a]

[a]*Department of Psychiatry, University of Oxford, Oxford, UK*
[b]*Department of Psychology, University of Oslo, Oslo, Norway*
[c]*NORMENT, Institute of Clinical Medicine, University of Oslo, & Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway*
[d]*Wellcome Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK*
[e]*Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK*
[f]*Dementia Research Centre, Institute of Neurology, University College London, London, UK*
[g]*Epidemiology of Ageing and Neurodegenerative diseases, Université de Paris, INSERM U1153, Paris France.*
[h]*Department of Epidemiology and Public Health, University College London, London, UK*

Abstract

Brain age is becoming a widely applied imaging-based biomarker of neural aging and potential proxy for brain integrity and health. We estimated multimodal and modality-specific brain age in the Whitehall II MRI cohort using machine learning and imaging-derived measures of gray matter morphology, diffusion-based white matter microstructure, and resting state functional connectivity. Ten-fold cross validation yielded multimodal and modality-specific brain age estimates for each participant, and additional predictions based on a separate training sample was included for comparison. The results showed equivalent age prediction accuracy between the multimodal model and the gray and white matter models (R^2 of 0.34, 0.31, and 0.31, respectively), while the functional connectivity model showed a lower prediction accuracy (R^2 of 0.01). Cardiovascular risk factors, including high blood pressure, alcohol intake, and stroke risk score, were each associated with more apparent brain aging, with consistent associations across modalities.

Keywords: Multimodal MRI, brain-age prediction, machine learning, biomedical factors

*correspondence: ann-marie.delange@psych.ox.ac.uk

1. Introduction

In older age, the human brain undergoes structural changes including reductions in brain volume, cortical thinning, and decline in white-matter microstructure [1], and large-scale resting-state networks become less segregated [2, 3]. Aging-related changes in brain structure and functional connectivity are associated with decreased cognitive performance in domains including memory and processing speed [2, 4, 5], and comprise an increased risk for neurodegenerative disorders such as dementia [6]. Although the senescent deterioration of the brain is well-known, older populations are characterised by substantial variation in neurobiological aging trajectories [1], and recent neuroimaging studies have focused on developing potential markers for brain aging [7, 8]. Brain-age prediction based on machine-learning algorithms estimates an individual's 'brain age' using structural and functional brain characteristics derived from magnetic resonance imaging (MRI) [9, 10, 11, 8]. Subtracting chronological age from estimated brain age provides an estimate of brain aging, the *brain-age delta*. For instance, if a 70 year old individual exhibits a brain-age delta of $+5$ years, their typical aging pattern resembles the brain structure of a 75 year old, i.e. their estimated brain age is older than what is expected for their chronological age [11]. Individual variation in delta estimations are associated with a range of cognitive and biological measures [10, 11, 12, 13, 14, 15, 16, 17], including cardiovascular health [14], and differences in brain age have been established between patient groups and healthy controls: individuals with conditions such as Alzheimer's disease, multiple sclerosis, epilepsy, and psychiatric disorders show on average older brain age relative to their chronological age [9, 18, 19, 20, 21, 22]. Longitudinal studies have documented highly reliable brain age prediction in stroke patients [23], and accelerated brain aging in patients with schizophrenia and multiple sclerosis [21, 24, 25]. Combined with studies on the association between brain-age delta and biomedical factors in healthy population cohorts [13, 15], the documented reliability and clinical sensitivity supports the utility of brain-age estimation as a candidate biomarker for neurological senescence and disease [8].

Modality-specific brain age models (based on e.g. gray and white matter separately) provide information about tissue-specific aging processes [26, 27]. For instance, imaging-derived

2

measures of gray matter are known to detect cortical atrophy in older age-groups [28], while changes in diffusion MRI measures reflect age-related decline in white matter microstructure, as well as white matter lesions, which are more prevalent in aging, relative to young, populations [29]. Functional MRI (fMRI) measures are indicative of brain network connectivity, which may change with advancing age [2, 3]. Cardiovascular risk factors may influence these neural aging processes differently [26, 27, 30], and in a recent Whitehall II (WHII) MRI study using voxelwise analyses, allostatic load, metabolic syndrome, and multifactorial stroke risk predicted gray matter density measured decades later, while only cumulative stroke risk measured by the Framingham stroke risk score [31] predicted white matter integrity in terms of fractional anisotropy and mean diffusivity [32].

In this study of the WHII MRI cohort (N = 671), we investigated whether machine learning using neuroimaging data could produce reliable biomarkers of brain aging, and whether cardiovascular risk factors including blood pressure and alcohol intake, and cumulative risk as indicated by the Framingham stroke risk score were associated with modality-specific brain-age markers. We estimated brain age using ten-fold cross validation in separate models based on I) gray matter (GM) measures, II) white matter (WM) measures derived from diffusion tensor imaging (DTI) and white matter hyperintensites (WMH), III) functional connectivity measures derived from resting state fMRI (rs-fMRI), and IV) a multimodal model that included all of the brain measures. A detailed description of the methodology is provided below.

2. Materials and Methods

2.1. Sample:

The WHII study was established in London in 1985, and included an initial cohort of 10,308 civil servants. Between 2012 and 2013, 6035 individuals participated in the Phase 11 assessment, from which a random sample of 800 participants was enrolled in an MRI sub-study including brain scans and biomedical assessments [33] (www.psych.ox.ac.uk/research/neurobiology-of-ageing/research-projects-1/whitehall-oxford). The current sample was drawn from the WHII MRI sub-study, and included 715 participants with multimodal MRI data. Forty-four participants were excluded based on neurological disease and incidental

MRI findings, yielding a final sample of 671 participants. Sample demographics are provided in Table 1.

Table 1: Sample demographics. Age range (mean age $\pm$ standard deviation = 54.70 ± 7.29), percentage male (M) and female (F) participants, percentage with white (W) and non-white (NW) ethnic background, and percentage with educational qualifications U = university degree, PG = post-graduate / masters / PhD, Pr = Professional qualifications, A = A levels or equivalent, O = O levels or equivalent, N = No qualifications.

N	Age range	Sex %	Ethnicity %	Educational qualification %
671	60.43 - 84.58	M79 | F21	W95 | NW5	U27 | PG22 | Pr12 | A18 | O14 | C5 | N3

2.2. MRI data acquisition and processing:

MRI data were acquired using a 3 Tesla Siemens Magnetom Verio (n. of participants = 473) with a 32-channel receive head coil (between April 2012 – Dec 2014) and, following scanner updates, a 3 Tesla Siemens Magnetom Prisma (n. of participants = 198) with a 64-channel head-neck coil (June 2015 – Dec 2016). MRI images for all participants were processed using the analysis pipeline described in [33], including automated surface-based morphometry and subcortical segmentation as implemented in FreeSurfer 6.0 [34], and residualized with respect to scanner, relative head motion during the acquisition of rs-fMRI images [35, 36], intracranial volume (ICV [37]), sex, and ethnic background using linear models.

2.2.1. Gray matter:

In line with recent large-scale implementations [16, 19], we utilized a fine-grained cortical parcellation scheme [38] to extract cortical thickness, area, and volume for 180 regions of interest per hemisphere, in addition to the classic set of subcortical and cortical summary statistics from FreeSurfer [34]. This yielded a total set of 1118 structural brain imaging features (360/360/360/38 for cortical thickness/area/volume, as well as cerebellar/subcortical and cortical summary statistics, respectively). To exclude potential outliers, we tested for values that were $\pm$ 4 standard deviations (SD) away from the average on the global MRI measures mean cortical or subcortical gray matter volume. No outliers were identified based on the general T1-derived measures.

2.2.2. White matter microstructure:

Global and tract-specific estimates of fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD) and mode of anisotropy (MO) were calculated for each individual. In accordance with established methods [39], tract-specific estimates of each DTI metric were derived using 48 standard-space masks available from the ICBM-DTI-81 White-Matter Labels Atlas [40, 41], producing a total of 245 DTI features. Global WMH volumes were automatically extracted from FLAIR images with Brain Intensity AbNormality Classification Algorithm (BIANCA) [42]. To avoid scanner-specific biases in these estimates, BIANCA was initially trained with WMH masks manually delineated in a sub-sample of individuals scanned on the Prisma (n = 24) and Verio (n = 24) scanners and an independent sample from the UK Biobank study (n = 12) [43]. Subjects with values ± 4 SD away from mean FA or mean MD were excluded, yielding a total of 668 included subjects.

2.2.3. Functional connectivity:

Spatial maps of large-scale resting state networks were derived by applying MELODIC group Independent Component Analysis (group-ICA, n of components = 25) [44] to the rs-fMRI images of 678 individuals of the Whitehall II MRI sub-study. All non-artefactual group-ICA components (n = 19) and subject-specific rs-fMRI timeseries (extracted with dual regression [45, 46]) were used in FSLNets (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLNets) in order to generate a 19x19 matrix representing subject-specific correlations between each pair of networks. Partial correlations (derived using L2 Regularization, setting rho = 0.01 in Ridge Regression option in FSLNETS) were examined in the present study, as these estimates allow for a more direct estimate of the connectivity between each pair of nodes [47]. These partial correlations were then z-transformed using Fisher's transformation. Subsequently, the upper triangle of the matrix was converted into a row vector, producing a total of 171 rs-fMRI features for each participant. Subjects with values ± 4 SD away from the average on any of the fMRI measures were excluded, yielding a total of 651 included subjects.

2.3. Brain age prediction:

2.3.1. K-fold cross-validation

The XGBoost regressor model https://xgboost.readthedocs.io/en/latest/python) was used to run the brain-age prediction, including an algorithm that has been used in recent large-scale brain age studies [16, 19]. Parameters were set to maximum depth = 3, number of estimators = 100, and learning rate = 0.1 (defaults). Age was estimated in ten-fold cross validations yielding a multimodal brain-age estimate for each individual, as well as modality-specific brain-age estimates based on gray matter, white matter, and functional connectivity measures. To investigate the prediction accuracy, correlation analyses were run for predicted versus chronological age, and R^2, root mean square error (RMSE), and mean absolute error (MAE) were calculated for each model. After removing outliers for each respective modality, the multimodal model included 649 subjects. For this model, average RMSE was calculated from a cross validation with ten splits and ten repetitions, and compared to a null distribution calculated from 1000 permutations. The result is shown in Figure 1.
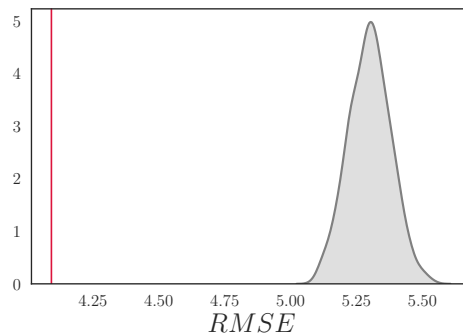

Figure 1: The mean ± SD root mean square error (RMSE) for the multimodal brain age model was 4.09 ± 0.32 based on the cross validation (red vertical line). The null distribution calculated from 1000 permutations is shown in grey, with a mean ± SD of 5.30 ± 0.07. The number of permuted results from the null distribution that exceeded the mean from the cross validation was 0 ($p < 0.001$).

2.3.2. Data quality analyses

To investigate effects of including versus excluding low-quality imaging data in the models, we trained two separate models for each modality (with and without low-quality data) on 60% of the sample, and tested them on the remaining 40% of the data. The categorization of data was based on manual quality checks and images were classified as "poor quality" if

they contained evidence of excessive motion and/or scanner-related artefacts. For diffusion weighted images, images with > 5 volumes missing from their scans (which were excluded due to containing 10 or more outlier slices) were also labelled as "poor quality". The percentage of subjects with low-quality data in the training sets was 9% for the gray matter model, 12% for the white matter model, 13% for the RS functional connectivity model, and 11% for the multimodal model. To test for differences in model performance before and after excluding low-quality data, we used a Z test for correlated samples [48]:

$$Z = (r_{\text{All}} - r_{\text{QC}})/\sqrt{\sigma_{\text{All}}^2 + \sigma_{\text{QC}}^2 - 2\rho\sigma_{\text{All}}\sigma_{\text{QC}}} \qquad (1)$$

where "All" represents the full sample, "QC" represents the sample with low-quality data removed, the r terms represent the Pearson's correlation coefficients of predicted versus chronological age, the σ terms represent their errors, and ρ represents the correlation between the two sets of model predictions.

2.3.3. External training sets

As a cross check, we estimated gray-matter based brain age in the current sample using an external model trained on a 27,200 subjects from UK Biobank, with a mean age $\pm$ SD of 55.40 ± 7.46, and 48% male / 52% female subjects (*External model 1*). The model contained the same 1118 structural MRI variables as those included in the WHII k-folding model (see section 2.2.1). The MRI data were residualized with respect to scanning site, data quality and motion using Euler numbers [49] extracted from FreeSurfer, ICV, sex, and ethnic background using linear models. Subjects with known brain disorders were excluded based on ICD10 diagnose (chapter V and VI, field F and G, excluding G5, http://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=41270). 43 subjects with values $\pm$ 4 SD away from the average on the measures mean cortical or subcortical gray matter volume were excluded, yielding a total of 27,157 subjects. The correlation between the brain-age estimates based on this training set and the estimates based on the k-folding procedure for gray matter was $r = 0.49, p < 0.001$, 95% confidence interval (CI) $= [0.43, 0.54]$. When applied to the WHII dataset, the RSME was 10.28, and the correlation between predicted and chronological age was $r = 0.49, p < 0.001$, 95% CI $= [0.43, 0.54]$.

To test the performance of a sex-specific external model with a larger age range, we applied an additional gray matter model trained on a separate sample of 35,474 individuals aged 3–89 (mean age $\pm$ SD = 47.30 $\pm$ 17.10) [19] (*External model 2*). The model was derived from https://github.com/tobias-kaufmann/brainage. For this analysis, the WHII data were corrected for scanning site, estimated in-scanner head motion, ICV, and ethnic background. The external model was trained on men and women separately [19], and applied to each sex in the WHII sample. The correlation between the brain-age estimates based on this training set and the estimates based on the k-folding procedure for gray matter was $r = 0.64, p < 0.001$, $95\%\text{CI} = [0.59, 0.68]$. When applied to the WHII dataset, the RSME was 9.58, and the correlation between predicted and chronological age was $r = 0.47, p < 0.001, 95\%$ CI = [0.41, 0.53]. As the two external models showed similar performance, we included only the sex-specific *external model 2* with larger age range in the subsequent analyses, herein referred to as "the external model".

2.4. Brain age delta and biomedical measures:

To investigate associations with biomedical variables, the brain-age delta (predicted age – chronological age) was used as a measure of apparent brain aging. Clinical measures included systolic and diastolic blood pressure, alcohol intake measured by units per week (see [50] for details), and the Framingham stroke risk score [31], which includes cardio-metabolic measures, smoking habits, diabetes, sex, and age (see [32] for full description). Blood pressure and alcohol intake were corrected for sex, ethnic background, and educational level using linear models. Stroke risk score was corrected for ethnic background and educational level, as sex was already accounted for in the score. 624 subjects had data on all biomedical variables as well as demographic variables. Subjects with values $\pm$ 4 SD away from the average on any of the variables were excluded from the analyses, yielding 616 subjects in total. The mean $\pm$ SD for each measure is shown in Table 2.

Table 2: Mean $\pm$ standard deviation on each of the biomedical measures.

N	Systolic BP	Diastolic BP	Alcohol intake	Framingham Stroke Risk score
616	140.52 ± 16.48	77.17 ± 10.47	14.55 ± 13.47	11.04 ± 5.46

To obtain a direct comparison of β values, both the brain-age deltas and the clinical variables were standardized (subtracting the mean and dividing by the SD) before a series of multiple regressions were run. In order to adjust for a frequently observed bias in brain-age prediction leading to an overestimation of brain age in younger subjects and an underestimation of brain age in older subjects [14, 15, 30, 51, 52], chronological age was included as a covariate. Correction for multiple comparisons was performed using false-discovery rate correction [53]. The statistical analyses were conducted using Python 3.7.0.

3. Results

3.1. Brain age prediction:

Figure 2 shows the correlations between predicted age and chronological age for each of the models. The multimodal, gray matter, and white matter models showed consistent performance with explained variance (R^2) $\pm$ standard error (SE) of 0.33 ± 0.02, 0.31 ± 0.02 and 0.31 ± 0.03, respectively, while the functional connectivity model showed lower prediction accuracy with an $R^2 \pm$ SE of 0.02 ± 0.04. The explained variance of the predictions based on the external gray matter model was 0.22 ± 0.03. The accuracy of each model's prediction measured by R^2, adjusted R^2, RMSE, and MAE are shown in Table 3. As a follow-up analysis, we re-ran the multimodal model without the rs-fMRI data to test for prediction improvements. The results showed comparable performance with and without the rs-fMRI data (RMSE = 4.14 vs 4.17, $z = -1.14, p = 0.26$). Figure 3 shows the proportion of variance in chronological age explained by each of the models. Figure 4 shows the correlations between the multimodal and modality-specific brain-age deltas. The delta values were first corrected for age-bias [14, 15, 30, 51, 52] using linear models, and the residuals were used in the correlation analyses. Excluding low-quality data improved the multimodal model ($z = 2.07, p = 0.04$), but none of the other models, as shown in Table 4 and 5.
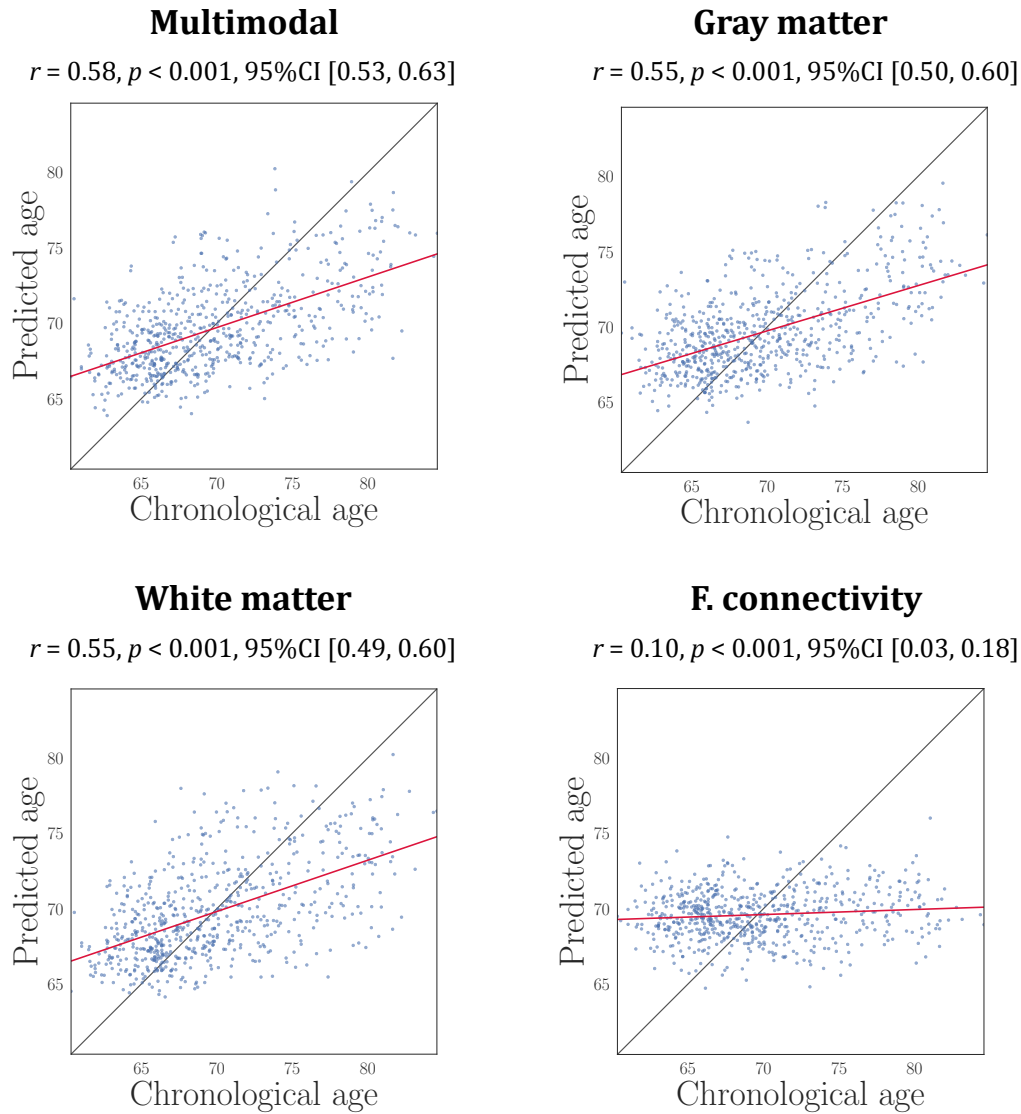
Figure 2: The association between predicted and chronological age for each of the models is shown by the red line. The black line illustrates a perfect relationship between x and y.

Table 3: Number of MRI variables, R^2, adjusted R^2, root mean square error (RMSE), and mean absolute error (MAE) for each of the brain-age models. RMSE and MAE are reported in years.

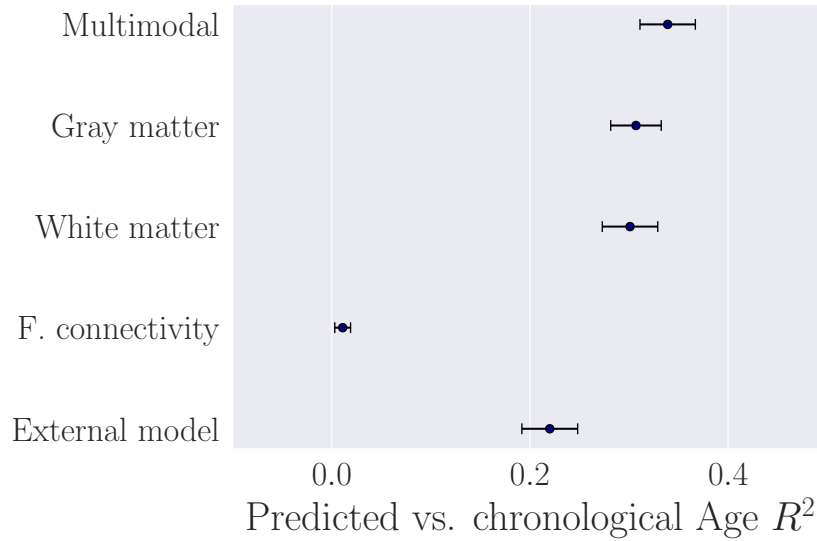| Model	MRI variables	R^2	Adj R^2	RMSE	MAE
Multimodal	1535	0.339	0.337	4.138	3.335
Gray matter	1118	0.307	0.305	4.246	3.420
White matter	246	0.301	0.299	4.268	3.383
F. connectivity	171	0.011	0.008	5.195	4.212
External model	1118	0.221	0.218	9.584	8.279

Figure 3: The proportion of variance in chronological age explained ($R^2 \pm$ standard error) by the multimodal model, the modality-specific models, and the external model based on a separate training sample.
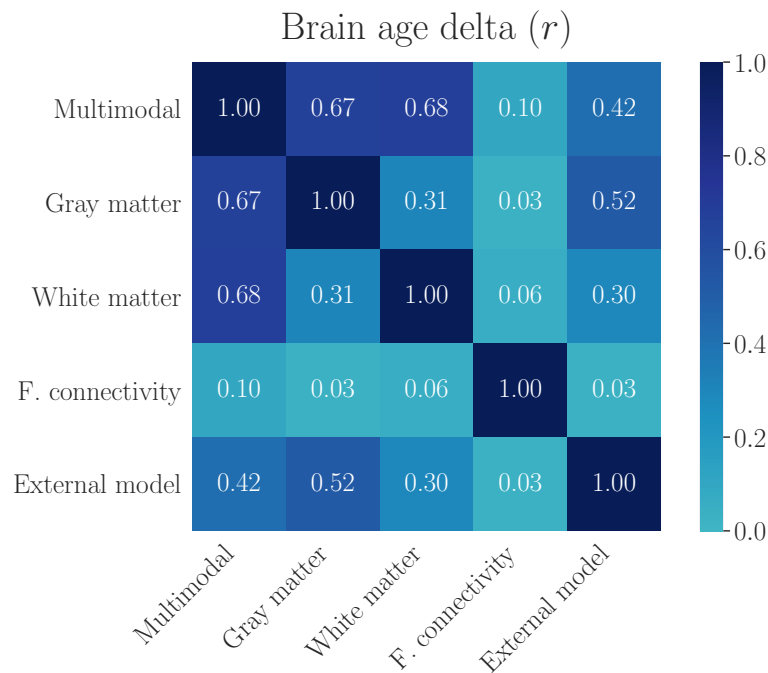

Figure 4: The correlations (Pearson's r) between brain-age deltas of the multimodal model, the modality-specific models, and the external model based on a separate training sample. The delta values are corrected for chronological age.

11

Table 4: The correlation (Pearson's r) between predicted and chronological age based on training sets with and without low-quality data applied to the same test set. 95% confidence intervals are indicated in square brackets. Z represents the difference in r values expressed in standard deviations, accounting for the correlated samples (Eq.1.)

Model	All data included	QC data excluded	Z	p
Multimodal	$r = 0.54, [0.44, 0.62]$	$r = 0.57, [0.48, 0.64]$	2.07	0.04
Gray matter	$r = 0.56, [0.47, 0.64]$	$r = 0.59, [0.50, 0.66]$	1.70	0.09
White matter	$r = 0.51, [0.42, 0.59]$	$r = 0.50, [0.40, 0.58]$	-0.72	0.47
F. connectivity	$r = 0.08, [-0.04, 0.20]$	$r = 0.07, [-0.05, 0.19]$	-1.57	0.12

Table 5: Root mean square error (RMSE) and mean square error (MAE) for the predictions based on training sets with and without low-quality data (LQD) applied to the same test set. RMSE and MAE are reported in years.

Model	RMSE	$\text{RMSE}_{LQD-excluded}$	MAE	$\text{MAE}_{LQD-excluded}$
Multimodal	4.38	4.23	3.44	3.34
Gray matter	4.25	4.18	3.45	3.43
White matter	4.48	4.51	3.60	3.64
F. connectivity	5.43	5.41	4.33	4.28

3.2. Biomedical predictors:

The associations between blood pressure (BP), alcohol intake, and stroke risk and brain-age deltas are shown in Figure 5 and Table 6. The associations were consistent across models. To test for non-linear relationships, polyfits including both a linear and a quadratic term (γ) were run for the multimodal brain age and each variable. No non-linear associations were found (**Systolic BP**: $\gamma = 0.002 \pm 0.028, p = 0.948$; **Diastolic BP**: $\gamma = -0.003 \pm 0.028, p = 0.928$;; **Alcohol intake**: $\gamma = 0.023 \pm 0.028, p = 0.413$; **Framingham stroke risk score**: $\gamma = 0.004 \pm 0.028, p = 0.887$).

Table 6: Relationships between biomedical variables and brain-age delta for each modality, including age as a co-variate. p-values are reported before and after FDR correction. Corrected p-values below 0.05 are marked with an asterisk.

| Model	β	SE	t	p	p_{corr}
BP Systolic					
Multimodal	0.04	0.02	1.79	0.073	0.112
Gray matter	0.00	0.02	0.22	0.826	0.826
White matter	0.05	0.02	2.00	0.046	0.077
F. connectivity	0.03	0.01	2.78	0.006	0.022*
External model	0.03	0.03	0.83	0.408	0.429
BP Diastolic					
Multimodal	0.05	0.02	2.35	0.019	0.039*
Gray matter	0.02	0.02	1.00	0.317	0.354
White matter	0.04	0.02	1.73	0.084	0.113
F. connectivity	0.03	0.01	2.39	0.017	0.039*
External model	0.03	0.03	1.00	0.318	0.354
Alcohol intake					
Multimodal	0.07	0.02	2.94	0.003	0.017*
Gray matter	0.08	0.02	3.73	< 0.001	0.004*
White matter	0.04	0.02	1.75	0.081	0.113
F. connectivity	0.03	0.01	2.02	0.044	0.077
External model	0.12	0.03	3.59	< 0.001	0.077
Stroke risk					
Multimodal	0.09	0.03	3.47	0.001	0.004*
Gray matter	0.06	0.03	2.34	0.020	0.039*
White matter	0.07	0.03	2.52	0.012	0.039*
F. connectivity	0.03	0.01	2.38	0.018	0.039*
External model	0.06	0.04	1.49	0.136	0.169
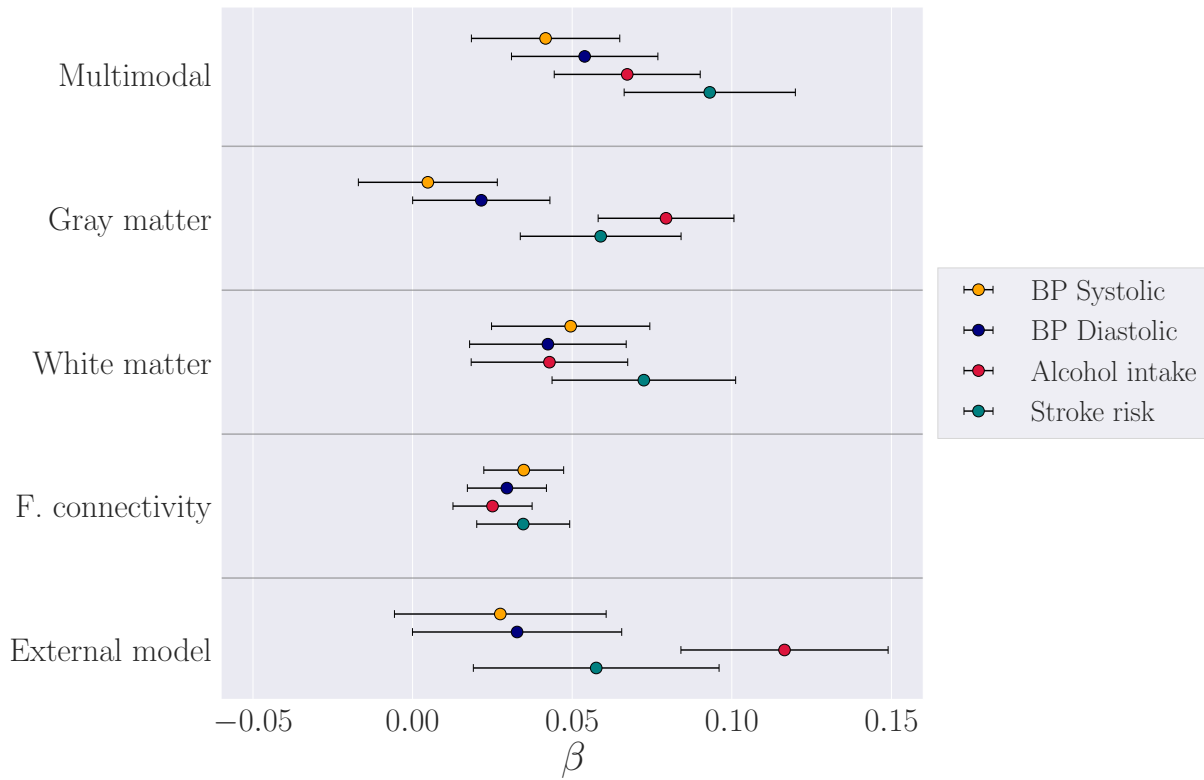
Figure 5: The associations ($\beta\pm$ standard error) between standardized measures of brain-age delta and blood pressure, alcohol intake, and Framingham stroke risk score for each of the brain-age models. The analyses included age as a covariate.

4. Discussion

Our findings demonstrate that machine learning models provide meaningful imaging-based biomarkers for brain aging in healthy population cohorts [13, 15]. We found that a) the multimodal, gray matter, and white matter based models showed comparable performance, while the functional connectivity model showed lower prediction accuracy, b) blood pressure, alcohol intake, and stroke risk scores were each associated with brain-age deltas, and c) these associations were consistent across modalities.

Although previous studies have suggested better prediction with multiple imaging modalities [13, 14, 30, 54], the current study showed equivalent prediction accuracy between the multimodal model and the gray and white matter models. The exclusion of low-quality data improved the performance of the multimodal model, suggesting that established procedures for data quality control may have implications for model performance [39, 55]. Although there was also a tendency for the gray matter model to improve with the exclusion of low-quality data, discarding such data did not affect the performance of the white matter and functional connectivity models. The external gray matter model that was trained on an independent sample showed less accurate prediction of brain age in our data compared to the k-folding based gray matter model. While both datasets were corrected for factors including scanner site, motion, and ICV, such discrepancies indicate that confounding factors including recruitment procedures, scanner equipment and data-processing pipelines may influence prediction accuracy across datasets [56].

The lowest prediction accuracy was observed for rs-fMRI, indicating that the included measures of functional connectivity were less closely related to chronological age compared to structural measures. Although it is possible that voxel-wise functional connectivity measures could improve the model performance [57], the lower age-sensitivity of rs-fMRI measurements may be explained by these metrics reflecting a state, rather than trait based, assessment [58, 59, 60]. While resting-state networks, including the default mode network, are characterized as highly replicable both within-participants and across studies [45, 61, 62], investigations employing dynamic rs-fMRI conversely suggest that the connectivity between these networks may vary even within a single session of scanning [63]. Alternatively, resting-state networks and their connectivity to other networks may be preserved though plasticity

15

despite age-related structural changes [64]. Another potential explanation is that age-related changes to network connectivity occur in a modular, rather than gradual manner [65, 66]. While gray and white matter based brain-age models provide relatively accurate predictions across studies [13, 15, 17, 19, 26], a recent application of brain-age estimation in the UK Biobank cohort similarly highlights fMRI-based brain-age prediction as a weaker correlate of chronological age ($r = 0.434$), relative to gray matter ($r = 0.685$) and DTI-based ($r = 0.668$) predictions [14].

Despite poorer performance accuracy of the functional connectivity model, the fMRI-based brain-age deltas showed associations with the biomedical variables that were similar to the other modalities. In line with recent findings from UK Biobank [14, 15], positive associations were found between brain-age deltas and diastolic blood pressure, alcohol intake, and stroke risk, concurring with previous WHII studies [32, 50], and demonstrating that the brain age-delta measure reflects individual variation in neural aging processes [30]. The associations with biomedical variables were consistent across models (see Figure 5), indicating that while modality-specific brain age models may be informative in patient groups where tissue types are differently affected by disease [26, 27, 54, 67, 68], such models may be more closely related in healthy cohorts [14]. It is possible that regional modelling of modality-specific brain aging patterns may be more suitable to detect specific associations with biomedical and clinical measures [19], which could get lost in machine learning models that summarise aging across the whole brain to produce a single global prediction [27].

In conclusion, machine-learning based brain age prediction can reduce the dimensionality of neuroimaging data to provide meaningful biomarkers of individual brain aging. While the presented imaging-derived markers can help to assess general effects of clinical and biomedical risk factors on the brain, models of distinct and regional neural aging patterns may result in more refined biomarkers that can capture additional biological detail [19, 26, 27].

Acknowledgements

References

[1] A. M. Fjell, L. McEvoy, D. Holland, A. M. Dale, K. B. Walhovd, A. D. N. Initiative, et al., What is normal in normal aging? effects of aging, amyloid and alzheimer's disease on the cerebral cortex and the hippocampus, Progress in neurobiology 117 (2014) 20–40.

[2] R. Sala-Llonch, D. Bartrés-Faz, C. Junqué, Reorganization of brain networks in aging: a review of functional connectivity studies, Frontiers in psychology 6 (2015) 663.

[3] A. M. Mowinckel, T. Espeseth, L. T. Westlye, Network-specific effects of age and in-scanner subject motion: a resting-state fmri study of 238 healthy adults, Neuroimage 63 (2012) 1364–1373.

[4] C. Grady, The cognitive neuroscience of ageing, Nature Reviews Neuroscience 13 (2012) 491.

[5] R. Cabeza, M. Albert, S. Belleville, F. I. Craik, A. Duarte, C. L. Grady, U. Lindenberger, L. Nyberg, D. C. Park, P. A. Reuter-Lorenz, et al., Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing, Nature Reviews Neuroscience 19 (2018) 701–710.

[6] T. Vos, A. D. Flaxman, M. Naghavi, R. Lozano, C. Michaud, M. Ezzati, K. Shibuya, J. A. Salomon, S. Abdalla, V. Aboyans, et al., Years lived with disability (ylds) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the global burden of disease study 2010, The lancet 380 (2012) 2163–2196.

[7] T. T. Brown, J. M. Kuperman, Y. Chung, M. Erhart, C. McCabe, D. J. Hagler Jr, V. K. Venkatraman, N. Akshoomoff, D. G. Amaral, C. S. Bloss, et al., Neuroanatomical assessment of biological maturity, Current biology 22 (2012) 1693–1698.

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

[9] K. Franke, G. Ziegler, S. Klöppel, C. Gaser, A. D. N. Initiative, et al., Estimating the age of healthy subjects from t1-weighted mri scans using kernel methods: exploring the influence of various parameters, Neuroimage 50 (2010) 883–892.

[10] J. H. Cole, K. Franke, Predicting age using neuroimaging: innovative brain ageing biomarkers, Trends in neurosciences 40 (2017) 681–690.

[11] K. Franke, C. Gaser, Ten years of brainage as a neuroimaging biomarker of brain aging: What insights have we gained?, Frontiers in Neurology 10 (2019) 789.

[12] J. H. Cole, S. J. Ritchie, M. E. Bastin, M. V. Hernández, S. M. Maniega, N. Royle, J. Corley, A. Pattie, S. E. Harris, Q. Zhang, et al., Brain age predicts mortality, Molecular psychiatry 23 (2018) 1385.

[13] J. H. Cole, R. E. Marioni, S. E. Harris, I. J. Deary, Brain age and other bodily 'ages': implications for neuropsychiatry, Molecular psychiatry 24 (2019) 266–281.

[14] J. H. Cole, Multi-modality neuroimaging brain-age in uk biobank: relationship to biomedical, lifestyle and cognitive factors, bioRxiv (2019). URL: https://www.biorxiv.org/content/early/2019/10/21/812982. doi:10.1101/812982. arXiv:https://www.biorxiv.org/content/early/2019/10/21/812982.full.pdf.

[15] S. M. Smith, D. Vidaurre, F. Alfaro-Almagro, T. E. Nichols, K. L. Miller, Estimation of brain age delta from brain imaging, NeuroImage (2019).

[16] A.-M. G. de Lange, T. Kaufmann, D. van der Meer, L. A. Maglanoc, D. Alnæs, T. Moberget, G. Douaud, O. A. Andreassen, L. T. Westlye, Population-based neuroimaging reveals traces of childbirth in the maternal brain, Proceedings of the National Academy of Sciences (2019).

[17] A.-M. G. de Lange, C. Barth, T. Kaufmann, I. Maximov, D. van der Meer, I. Agartz, L. T. Westlye, Cumulative estrogen exposure, apoe genotype, and women's brain aging-a population-based neuroimaging study, bioRxiv (2019) 826123.

[18] K. Franke, C. Gaser, Longitudinal changes in individual brainage in healthy aging, mild cognitive impairment, and alzheimer's disease, GeroPsych (2012).

[19] T. Kaufmann, D. van der Meer, N. T. Doan, E. Schwarz, M. J. Lund, I. Agartz, D. Alnæs, D. M. Barch, R. Baur-Streubel, A. Bertolino, et al., Common brain disorders are associated with heritable patterns of apparent aging of the brain, Nature Neuroscience, In press (2019).

[20] H. R. Pardoe, J. H. Cole, K. Blackmon, T. Thesen, R. Kuzniecky, H. E. P. Investigators, et al., Structural brain changes in medically refractory focal epilepsy resemble premature brain aging, Epilepsy research 133 (2017) 28–32.

[21] N. Koutsouleris, C. Davatzikos, S. Borgwardt, C. Gaser, R. Bottlender, T. Frodl, P. Falkai, A. Riecher-Rössler, H.-J. Möller, M. Reiser, et al., Accelerated brain aging in schizophrenia and beyond: a neuroanatomical marker of psychiatric disorders, Schizophrenia bulletin 40 (2013) 1140–1153.

[22] H. G. Schnack, N. E. Van Haren, M. Nieuwenhuis, H. E. Hulshoff Pol, W. Cahn, R. S. Kahn, Accelerated brain aging in schizophrenia: a longitudinal pattern recognition study, American Journal of Psychiatry 173 (2016) 607–616.

[23] G. Richard, K. Kolskår, K. M. Ulrichsen, T. Kaufmann, D. Alnæs, A.-M. Sanders, E. S. Dørum, J. M. Sánchez, A. Petersen, H. Ihle-Hansen, et al., Brain age prediction in stroke patients: Highly reliable but limited sensitivity to cognitive performance and response to cognitive training, NeuroImage: Clinical (2019) 102159.

[24] E. A. Høgestøl, T. Kaufmann, G. O. Nygaard, M. K. Beyer, P. Sowa, J. E. Nordvik, K. Kolskår, G. Richard, O. A. Andreassen, H. F. Harbo, et al., Cross-sectional and longitudinal mri brain scans reveal accelerated brain aging in multiple sclerosis, Frontiers in neurology 10 (2019) 450.

[25] J. Cole, J. Raffel, T. Friede, A. Eshaghi, W. Brownlee, D. Chard, N. De Stefano, C. Enzinger, L. Pirpamer, M. Filippi, et al., Accelerated brain ageing and disability in multiple sclerosis, bioRxiv (2019) 584888.

[26] G. Richard, K. Kolskår, A.-M. Sanders, T. Kaufmann, A. Petersen, N. T. Doan, J. M. Sanchez, D. Alnaes, K. M. Ulrichsen, E. S. Dørum, et al., Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry, PeerJ 6 (2018) e5908.

[27] S. M. Smith, L. T. Elliott, F. Alfaro-Almagro, P. McCarthy, T. E. Nichols, G. Douaud, K. L. Miller, Brain aging comprises multiple modes of structural and functional change with distinct genetic and biophysical associations, BioRxiv (2019) 802686.

[28] A. M. Fjell, K. B. Walhovd, Structural brain changes in aging: courses, causes and cognitive consequences, Reviews in the Neurosciences 21 (2010) 187–222.

[29] S. Debette, H. Markus, The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis, Bmj 341 (2010) c3666.

[30] X. Niu, F. Zhang, J. Kounios, H. Liang, Improved prediction of brain age using multimodal neuroimaging data, Human Brain Mapping (2019).

[31] R. B. D'Agostino, P. A. Wolf, A. J. Belanger, W. B. Kannel, Stroke risk profile: adjustment for antihypertensive medication. the framingham study., Stroke 25 (1994) 40–43.

[32] E. Zsoldos, N. Filippini, A. Mahmood, C. E. Mackay, A. Singh-Manoux, M. Kivimäki, M. Jenkinson, K. P. Ebmeier, Allostatic load as a predictor of grey matter volume and white matter integrity in old age: The whitehall ii mri study, Scientific reports 8 (2018) 6411.

[33] N. Filippini, E. Zsoldos, R. Haapakoski, C. E. Sexton, A. Mahmood, C. L. Allan, A. Topiwala, V. Valkanova, E. J. Brunner, M. J. Shipley, et al., Study protocol: the whitehall ii imaging sub-study, BMC psychiatry 14 (2014) 159.

[34] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, et al., Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain, Neuron 33 (2002) 341–355.

[35] K. L. Miller, F. Alfaro-Almagro, N. K. Bangerter, D. L. Thomas, E. Yacoub, J. Xu, A. J. Bartsch, S. Jbabdi, S. N. Sotiropoulos, J. L. Andersson, et al., Multimodal population brain imaging in the uk biobank prospective epidemiological study, Nature neuroscience 19 (2016) 1523.

[36] S. M. Smith, T. E. Nichols, D. Vidaurre, A. M. Winkler, T. E. Behrens, M. F. Glasser, K. Ugurbil, D. M. Barch, D. C. Van Essen, K. L. Miller, A positive-negative mode of population covariation links brain connectivity, demographics and behavior, Nature neuroscience 18 (2015) 1565.

[37] O. Voevodskaya, A. Simmons, R. Nordenskjöld, J. Kullberg, H. Ahlström, L. Lind, L.-O. Wahlund, E.-M. Larsson, E. Westman, A. D. N. Initiative, et al., The effects of intracranial volume adjustment approaches on multiple regional mri volumes in healthy aging and alzheimer's disease, Frontiers in aging neuroscience 6 (2014) 264.

[38] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, et al., A multi-modal parcellation of human cerebral cortex, Nature 536 (2016) 171.

[39] F. Alfaro-Almagro, M. Jenkinson, N. K. Bangerter, J. L. Andersson, L. Griffanti, G. Douaud, S. N. Sotiropoulos, S. Jbabdi, M. Hernandez-Fernandez, E. Vallee, et al., Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank, Neuroimage 166 (2018) 400–424.

[40] S. Mori, S. Wakana, P. C. Van Zijl, L. Nagae-Poetscher, MRI atlas of human white matter, Elsevier, 2005.

[41] S. Wakana, A. Caprihan, M. M. Panzenboeck, J. H. Fallon, M. Perry, R. L. Gollub, K. Hua, J. Zhang, H. Jiang, P. Dubey, et al., Reproducibility of quantitative tractography methods applied to cerebral white matter, Neuroimage 36 (2007) 630–644.

[42] L. Griffanti, G. Zamboni, A. Khan, L. Li, G. Bonifacio, V. Sundaresan, U. G. Schulz, W. Kuker, M. Battaglini, P. M. Rothwell, et al., Bianca (brain intensity abnormality classification algorithm): a new tool for automated segmentation of white matter hyperintensities, Neuroimage 141 (2016) 191–205.

[43] L. Griffanti, I. Bertani, V. Bordin, I. Mattioli, G. Zamboni, S. Suri, E. Zsoldos, K. P. Ebmeier, M. M. Laganà, G. Baselli, M. Jenkinson, C. E. Mackay, E. Duff, Harmonising white matter hyperintensities measures across studies: impact of bianca training options, Poster presentation at the Organisation for Human Brain Mapping Conference (Rome 2019).

[44] C. F. Beckmann, S. M. Smith, Probabilistic independent component analysis for functional magnetic resonance imaging, IEEE transactions on medical imaging 23 (2004) 137–152.

[45] C. F. Beckmann, C. E. Mackay, N. Filippini, S. M. Smith, Group comparison of resting-state fmri data using multi-subject ica and dual regression (????).

[46] N. Filippini, B. J. MacIntosh, M. G. Hough, G. M. Goodwin, G. B. Frisoni, S. M. Smith, P. M. Matthews, C. F. Beckmann, C. E. Mackay, Distinct patterns of brain activity in young carriers of the apoe-ε4 allele, Proceedings of the National Academy of Sciences 106 (2009) 7209–7214.

[47] S. Suri, A. Topiwala, N. Filippini, E. Zsoldos, A. Mahmood, C. E. Sexton, A. Singh-Manoux, M. Kivimäki, C. E. Mackay, S. Smith, et al., Distinct resting-state functional connections associated with episodic and visuospatial memory in older adults, NeuroImage 159 (2017) 122–130.

[48] D. W. Zimmerman, Correcting two-sample" z" and" t" tests for correlation: An alternative to one-sample tests on difference scores., Psicologica: International Journal of Methodology and Experimental Psychology 33 (2012) 391–418.

[49] A. F. Rosen, D. R. Roalf, K. Ruparel, J. Blake, K. Seelaus, L. P. Villa, R. Ciric, P. A. Cook, C. Davatzikos, M. A. Elliott, et al., Quantitative assessment of structural image quality, Neuroimage 169 (2018) 407–418.

[50] A. Topiwala, C. L. Allan, V. Valkanova, E. Zsoldos, N. Filippini, C. Sexton, A. Mahmood, P. Fooks, A. Singh-Manoux, C. E. Mackay, et al., Moderate alcohol consumption as risk factor for adverse brain outcomes and cognitive decline: longitudinal cohort study, bmj 357 (2017) j2353.

[51] H. Liang, F. Zhang, X. Niu, Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders, Human brain mapping (2019).

[52] T. T. Le, R. T. Kuplicki, B. A. McKinney, H.-w. Yeh, W. K. Thompson, M. P. Paulus, T. . Investigators, et al., A nonlinear simulation framework supports adjusting for age when analyzing brainage, Frontiers in aging neuroscience 10 (2018).

[53] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, Journal of the Royal statistical society: series B (Methodological) 57 (1995) 289–300.

[54] F. Liem, G. Varoquaux, J. Kynast, F. Beyer, S. K. Masouleh, J. M. Huntenburg, L. Lampe, M. Rahim, A. Abraham, R. C. Craddock, et al., Predicting brain-age from multimodal imaging data captures cognitive impairment, Neuroimage 148 (2017) 179–188.

[55] M. S. Graham, I. Drobnjak, H. Zhang, A supervised learning approach for diffusion mri quality control with minimal training data, NeuroImage 178 (2018) 668–676.

[56] C. R. Madan, Advances in studying brain morphology: the benefits of open-access data, Frontiers in human neuroscience 11 (2017) 405.

[57] H. Li, T. D. Satterthwaite, Y. Fan, Brain age prediction based on resting-state functional connectivity patterns using convolutional neural networks, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, 2018, pp. 101–104.

[58] A. B. Waites, A. Stanislavsky, D. F. Abbott, G. D. Jackson, Effect of prior cognitive state on resting state networks measured with functional connectivity, Human brain mapping 24 (2005) 59–68.

[59] J. Blautzik, C. Vetter, I. Peres, E. Gutyrchik, D. Keeser, A. Berman, V. Kirsch, S. Mueller, E. Pöppel, M. Reiser, et al., Classifying fmri-derived resting-state connectivity patterns according to their daily rhythmicity, Neuroimage 71 (2013) 298–306.

[60] B. J. Harrison, J. Pujol, H. Ortiz, A. Fornito, C. Pantelis, M. Yücel, Modulation of brain resting-state networks by sad mood induction, PLoS One 3 (2008) e1794.

[61] J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, C. F. Beckmann, Consistent resting-state networks across healthy subjects, Proceedings of the National Academy of Sciences 103 (2006) 13848–13853. URL: https://www.pnas.org/content/103/37/13848. doi:10.1073/pnas.0601417103. arXiv:https://www.pnas.org/content/103/37/13848.full.pdf.

[62] M. D. Fox, M. Greicius, Clinical applications of resting state functional connectivity, Frontiers in systems neuroscience 4 (2010) 19.

[63] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, V. D. Calhoun, Tracking whole-brain connectivity dynamics in the resting state, Cerebral cortex 24 (2014) 663–676.

[64] G. Benson, A. Hildebrandt, C. Lange, C. Schwarz, T. Köbe, W. Sommer, A. Flöel, M. Wirth, Functional connectivity in cognitive control networks mitigates the impact of white matter lesions in the elderly, Alzheimer's research & therapy 10 (2018) 109.

[65] J. Song, R. M. Birn, M. Boly, T. B. Meier, V. A. Nair, M. E. Meyerand, V. Prabhakaran, Age-related reorganizational changes in modularity and functional connectivity of human brain networks, Brain connectivity 4 (2014) 662–676.

[66] R. F. Betzel, L. Byrge, Y. He, J. Goñi, X.-N. Zuo, O. Sporns, Changes in structural and functional connectivity among resting-state networks across the human lifespan, Neuroimage 102 (2014) 345–357.

[67] A. Cherubini, M. E. Caligiuri, P. Péran, U. Sabatini, C. Cosentino, F. Amato, Importance of multimodal mri in characterizing brain tissue and its potential application for individual age prediction, IEEE journal of biomedical and health informatics 20 (2016) 1232–1239.

[68] A. R. Groves, S. M. Smith, A. M. Fjell, C. K. Tamnes, K. B. Walhovd, G. Douaud, M. W. Woolrich, L. T. Westlye, Benefits of multi-modal fusion analysis on a large-scale dataset: life-span patterns of inter-subject variability in cortical morphometry and white matter microstructure, Neuroimage 63 (2012) 365–380.