Neurobiology of Aging 92 (2020) 34–42

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors

James H. Cole*

Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK
Dementia Research Centre, Institute of Neurology, University College London, London, UK

ARTICLE INFO

Article history:
Received 21 October 2019
Received in revised form 5 March 2020
Accepted 24 March 2020
Available online 8 April 2020

Keywords:
Brain aging
Neuroimaging
Multimodality MRI
UK Biobank
Biomedical measures

ABSTRACT

The brain-age paradigm is proving increasingly useful for exploring aging-related disease and can predict important future health outcomes. Most brain-age research uses structural neuroimaging to index brain volume. However, aging affects multiple aspects of brain structure and function, which can be examined using multimodality neuroimaging. Using UK Biobank, brain-age was modeled in n = 2205 healthy people with T1-weighted MRI, T2-FLAIR, T2*, diffusion-MRI, task fMRI, and resting-state fMRI. In a held-out healthy validation set (n = 520), chronological age was accurately predicted (r = 0.78, mean absolute error = 3.55 years) using LASSO regression, higher than using any modality separately. Thirty-four neuroimaging phenotypes were deemed informative by the regression (after bootstrapping); predominantly gray-matter volume and white-matter microstructure measures. When applied to new individuals from UK Biobank (n = 14,701), significant associations with multimodality brain-predicted age difference (brain-PAD) were found for stroke history, diabetes diagnosis, smoking, alcohol intake and some, but not all, cognitive measures (corrected $p < 0.05$). Multimodality neuroimaging can improve brain-age prediction, and derived brain-PAD values are sensitive to biomedical and lifestyle factors that negatively impact brain and cognitive health.

1. Introduction

Aging has a pronounced effect on the human brain, resulting in cognitive decline and an increased risk of neurodegenerative diseases and dementia. Though brain aging is ubiquitous, differences between individuals can be substantial. Some people experience cognitive decline, neurodegeneration and age-related brain diseases in midlife, while others retain the majority of their cognitive function well into their 10th decade (Deary et al., 2009; Mendez, 2017; Wyss-Coray, 2016).

Poorer brain-health during aging has a pronounced negative impact for individuals, their families and for society (Wasay et al., 2016). Hence, ways of identifying people at risk of poorer brain health during aging have become an important research goal. One such approach, the so-called 'brain-age paradigm' (Cole and Franke, 2017; Franke and Gaser, 2019), aims to define the brain's 'biological age' (Jackson et al., 2003; Ludwig and Smoke, 1980). The idea here is that genetic and environmental factors can influence the rate at which age-associated biological changes accumulate, and that one's biological age might be a better predictor of disease-risk, functional capacity and residual lifespan, than chronological age. By defining a statistical model of healthy brain aging, using neuroimaging data to predict chronological age, one can then evaluate neuroimaging data from new individuals, and consider whether their 'brain-age' appears younger or older than their chronological age.

Having an older-appearing brain has previously associated with markers of physiological aging (e.g., grip strength, lung function, walking speed) and cognitive aging (Cole et al., 2018), suggesting that brain and body aging are related (Cole et al., 2019c). An older brain age has also been associated with poor future outcomes, including progression from mild cognitive impairment to dementia (Franke and Gaser, 2012; Gaser et al., 2013) and mortality (Cole et al., 2018). Brain-age can also be moderated by a range of different neurological and psychiatric diseases (reviewed in Cole et al., 2019b).

Statistical methods for modeling brain age using neuroimaging are generally highly accurate, with most approaches able to account for >90% of the variance in chronological age. The mean absolute error (MAE) of predictions are generally around 4–5 years (when predicting between 18–90 years), though the winners of a recent

* Corresponding author at: Dept. of Computer Science, University College London, 66-72 Gower Street, London WC1E 6EA, UK. Tel.: +44 (0)20 7679 7214; fax: +44 (0) 20 7387 1397.
E-mail address: james.cole@ucl.ac.uk.

https://doi.org/10.1016/j.neurobiolaging.2020.03.014

competition achieved MAE <3 years using an ensemble deep-learning method (https://www.photon-ai.com/pac2019). Nevertheless, further improving model accuracy is important step toward application in clinical settings, where predictions will be made at the individual level. Potential clinical uses include screening for poorer brain-health in cognitively normal middle-aged adults, in stratifying clinical trial recruitment or as a surrogate outcome measure of neuroprotective treatments.

Most brain-age models use only T1-weighted structural MRI, reflecting brain volumes. Given that aging impacts many aspects of brain structure and function, and that these can be measured with other neuroimaging modalities, one could potentially improve accuracy by incorporating complementary data on brain connectivity, white-matter hyperintensities, iron deposition and brain activity during tasks and at rest. A multimodality approach has the benefit of providing a richer and more comprehensive explanation of the mechanisms underlying individual differences in brain aging. For example, changes in white-matter microstructure may precede alterations in brain volume, or there may be regional differences in the susceptibility to age-related brain changes. Including more measurements should capture more age-related variance and thus improve model accuracy. Multimodality imaging could also be used to generate an array of modality-specific brain "ages". For example, an individual could have a structural brain-age, a diffusion brain-age and a functional connectivity brain-age. In healthy people, these separate "ages" should be closely related, according to the concept of biological age, but for patient groups, distinct patterns of aberrant brain aging could emerge.

Several previous brain-age studies have used 2 or 3 modalities (Brown et al., 2012; Cherubini et al., 2016; Groves et al., 2012; Liem et al., 2017; Niu et al., 2019; Richard et al., 2018). In one of the first brain age studies, Brown and colleagues (2012) used T1-weighted measures of regional volumes and cortical thickness, T2-weighted normalised signal intensity and diffusion-MRI measures to predict age during neurodevelopment (ages 3–20 years). The resulting multimodality model was highly accurate in this age range (mean error = 1.1 years), though they did not explicitly compare this model with the performance of its unimodal constituents. Liem and colleagues (2017) studied n = 2354 people, finding improved accuracy when combining T1-weighted structural MRI with resting-state fMRI (MAE = 4.29 years). Structural MRI resulted in higher accuracy than fMRI when used independently. Meanwhile, Groves and colleagues (2012) using linked independent component analysis to merge T1-weighted and diffusion-MRI in n = 484 participants, finding that the resulting largest component could predict age with high accuracy (correlation between chronological age and brain age $r = 0.95$). Similarly, Richard and colleagues (2018) used T1-weighted and diffusion-MRI in their study of n = 877 participants. Again, the combination of modalities achieved top performance ($r = 0.86$, MAE = 6.14 years), while T1-derived values resulted in higher age-prediction accuracy than white-matter microstructure measures from diffusion-MRI. Cherubini and colleagues (2016) used 3 modalities; T1-weighted, T2* relaxometry and diffusion-MRI data. This small study (n = 140) achieved excellent performance ($r = 0.96$) with these 3 modalities, although accuracy was still generally high with single modalities. Most recently, Niu and colleagues (2019), combined T1-weighted structural MRI, diffusion-MRI and resting-state fMRI to predict age in n = 839 participants aged 8–21 years, evaluating the performance of 36 different feature set and statistical model combinations. This comprehensive study found the lowest MAE with a deep neural network model combining all 3 modalities (MAE = 1.38 years), though other statistical models and combinations of 2 out of the 3 modalities reached similar performance levels (MAE range 1.38–1.55 years).

Findings from these multimodality brain age studies suggest that T1-weighted MRI is the best single predictor, but that adding other modalities results in higher accuracy than in any single modality, despite the noticeable levels of collinearity between modalities. These results are promising but still limited to 2 or 3 modalities. This is largely due to data availability; the UK Biobank imaging study (Miller et al., 2016) presents a new opportunity to overcome this limitation. With a planned total of n = 100,000 individuals undergoing standardized neuroimaging, using 4 identical, dedicated MRI scanners, a wealth of multimodality MRI data are becoming available. The modalities include T1-weigthed MRI, T2-FLAIR, susceptibility-weighted imaging (SWI), diffusion-MRI, task fMRI, and resting-state fMRI. In addition to neuroimaging, participants provide detailed information on current health, lifestyle, and medical history, as well as participating in cognitive testing and providing blood for biological assessments.

Here, I tested whether brain-age prediction in healthy people can be improved by combining data from these 6 modalities, as well as evaluating their independent performance. I then applied the resulting multimodality brain-age model to a held-out sample of UK Biobank participants to test the relationship between brain-prediction age and factors relating to health, lifestyle, and cognitive function. I hypothesized that i) highest accuracy (i.e., lowest MAE, highest variance explained in age) would be achieved by combining data from all modalities; ii) that T1-weighted MRI would provide the highest independent accuracy; iii) that all modalities could be used to significantly predict chronological age independently; iv) that smoking, alcohol intake, major physiological health conditions, and poorer cognition would be associated with having an older-appearing brain.

2. Materials and methods

2.1. Participants

Data from n = 22,392 participants from UK Biobank were analyzed. Of these, n = 20,237 completed the brain MRI assessment; which included n = 2776 participants who were missing one or more neuroimaging modality. This left n = 17,461 participants (aged 45–80 years, n = 9274 females, n = 8187 males) with complete neuroimaging data. Two subsets were then defined; one for training the statistical model (the training set), one for testing associations with other UK Biobank measures (the test set). Training data included only people who met criteria for being healthy at the time of scanning. Exclusion criteria were an ICD-10 diagnosis, a self-reported long-standing illness disability or infirmity (UK Biobank data field #2188), self-reported diabetes (field #2443), stroke history (field #4056), not having good or excellent self-reported health (field #2178). This gave n = 2725 (mean age = 61.47 ± 7.2 years, 1343 females and 1382 males). The test set comprised the remaining n = 14,701 participants (mean age = 62.64 ± 7.5, 7914 females, 6787 males). All UK Biobank data were downloaded and reformatted using the R package ukbtools (Hanscombe et al., 2019).

All participants provided informed consent. UK Biobank has ethical approval from the North West Multi-Centre Research Ethics Committee (MREC). Further details on the UK Biobank Ethics and Governance framework are available here: https://www.ukbiobank.ac.uk/the-ethics-and-governance-council/. The present analyses were conducted under data application number 40933, *Optimising neuroimaging biomarkers of brain aging to identify genetic and environmental risk factors for poor brain health*.

2.2. Data acquisition

Full details on the UK Biobank neuroimaging data are provided here: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. In brief, T1-weighted MRI used an MPRAGE sequence with

1-mm isotropic resolution. The T2 protocol used a fluid-attenuated inversion recovery (FLAIR) contrast with the 3D SPACE optimized readout, with a 1.05 × 1 × 1mm resolution. The SWI protocol used a dual-echo 3D gradient echo acquisition at 0.8 × 0.8 × 3mm resolution (TEs = 9.4, 20 ms). T2* values (i.e., signal decay times) were estimated from the magnitude images at the 2 echo times. Diffusion-MRI data were acquired with 2 b-values (b = 1,000, 2000 s/mm^2) at 2-mm isotropic resolution, with a multiband acceleration factor of 3. For both diffusion-weighted shells, 50 diffusion-encoding directions were acquired (covering 100 distinct directions over the 2 b-values). Task and resting-state fMRI use the following acquisition parameters, with 2.4-mm spatial resolution and TR = 0.735s, with a multiband acceleration factor of 8. The total neuroimaging acquisition protocol lasted 32 minutes per participant.

2.3. Data processing

Data used in the current analysis were the imaging-derived phenotypes developed centrally by researchers involved in UK Biobank (Miller et al., 2016) and distributed via the data showcase (http://biobank.ctsu.ox.ac.uk/crystal/index.cgi). These data were the available summary metrics for T1-weighted MRI, diffusion-MRI, T2-FLAIR, SWI (i.e., T2*) and task fMRI. Data from resting state fMRI were so-called bulk data, in this case the 25-dimension partial correlation matrices (data field #25752). These matrices had been converted into vectors of length 210, representing the pairwise correlations between the BOLD timeseries from 21 separate resting state networks (details on page 16 here: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf). Four of the twenty-five networks were classified as noise by UK Biobank researchers and removed prior to being made available. The figure of 210 values is arrived at by 21 networks multiplied by 20 (i.e., excluding the identity correlation), then divided by 2 (i.e., as the matrix is diagonally symmetrical). The use of partial correlations aims to provide a better estimate of direct "connection" strengths between networks than using full correlations.

The final set of neuroimaging phenotypes numbered as follows: T1-weighted = 165, T2-FLAIR = 1 (total volume of white-matter hyperintensities), T2* = 14, diffusion-MRI = 675, task fMRI = 14, resting-state fMRI = 210. This gave 1079 phenotypes in total (see Table A1 in Appendix).

2.4. Statistical analysis

All statistical analyses were carried using R version 3.5.2 (R Core Team, 2015). The R Markdown notebook containing all statistical code can be found here: https://james-cole.github.io/UKBiobank-Brain-Age/.

To define a multimodality healthy aging model, the training data were randomly split into separate training (80%, n = 2205) and validation sets (20%, n = 520), to ensure that model accuracy could be evaluated in an unbiased manner. All neuroimaging phenotypes were normalized (i.e., scaled by standard deviation and mean centered) to account for the different measurement scales used by the different modalities.

To predict age from neuroimaging data, a least absolute shrinkage and selection operator (LASSO) regression was run, with age as the outcome variable and neuroimaging phenotypes as the predictors. LASSO regression imposes an L1-norm penalty, in which the goal is to minimize the absolute value of the beta coefficients in the model. Coefficients that shrink below a threshold (lambda) are automatically set to zero. This leads to a sparse solution, which reduces the variance in the model (i.e., regularization); in effect working to select informative features and remove uninformative ones. To derive the optimal value lambda, 10-fold cross-validation was first run using a range of lambdas, and the highest value within one standard error of the minimum was used in subsequent analysis.

Since LASSO can be susceptible to biased solutions caused by multicollinearity, bootstrapping (i.e., resampling with replacement) was used to generate a distribution of coefficients for each predictor variable. Here, 1000 bootstraps were run, and the mean coefficient and 95% confidence interval (calculated using the basic bootstrap method) were computed. Neuroimaging phenotypes with bootstrapped 95% confidence intervals that did not overlap zero were considered to be informative for the prediction of age.

The LASSO regression procedure was then run per modality (T1-weighted, T2-FLAIR, diffusion-MRI, SWI, task fMRI, resting-state fMRI), where phenotypes from that modality alone were included. This time bootstrapping was not conducted, as the goal was to compare model-level performance across modalities, rather than identify important features within each model. Next, the LASSO was repeated using a leave-one-modality-out approach, the inverse of the single modality analysis, to test whether including some modalities decreased model performance.

To test for associations between brain-age and health-related variables in UK Biobank, the LASSO model was applied to predict age in people from the testing set. As recent research has highlighted a proportional bias in brain-age calculation, whereby the difference between chronological age and brain-predicted age is negatively correlated with chronological age (Le et al., 2018; Liang et al., 2019; Smith et al., 2019), an age-bias correction procedure was used. This entailed calculating the regression line between age (predictor) and brain-predicted age (outcome) in the training set, then using the slope (i.e., coefficient) and intercept of that line to adjust brain-predicted age values in the testing set (by subtracting the intercept and then dividing by the slope). After applying the age-bias correction the brain-predicted age difference (brain-PAD) was calculated; chronological age subtracted from brain-predicted age. This gives a resulting value in unit years, with positive values representing an older-appearing brain and negative values a younger-appearing brain.

Associations between brain-PAD values and demographic, biomedical, cognitive, and lifestyle measures were then tested, using linear regression models. In the models, brain-PAD was the outcome measure, the variable of interest was a predictor, alongside age, age^2, sex, height, volumetric scaling from T1-weighted MRI to standard (data field #25000), and mean task fMRI head motion (averaged across space and time points; data field #25742) as covariates. These covariates were chosen due to face validity (i.e., theoretically likely to relate to brain structure) and statistical correlation with brain-PAD.

Biomedical measures (from visit #2, the imaging visit) tested were systolic blood pressure, diastolic blood pressure, weight, body mass index (BMI), hip circumference, diabetes diagnosis; stroke diagnosis, and facial aging. Lifestyle measures were smoking status, alcohol intake frequency, duration of moderate activity or vigorous activity per day. Cognitive performance measures were fluid intelligence score, Trail making task: duration to complete numeric path trail 1, duration to complete alphanumeric path trail 2, Matrix pattern completion: number of puzzles correctly solved, duration spent answering each puzzle and Tower rearranging: number of puzzles correct. Full details on the coding for these variables are reported in the UK Biobank data showcase (http://biobank.ctsu.ox.ac.uk/crystal/search.cgi). Of particular note, the coding for alcohol intake frequency has lower values reflecting higher intake (1 = daily or almost daily; 2 = 3 or 4 times a week; 3 = once or twice a week; 4 = one to 3 times a month; 5 = special occasions only; 6 = never; −3 = prefer not to answer). Multiple testing corrections for

these 18 separate measures were conducted using false-discovery rate correction (Benjamini and Hochberg, 1995).

3. Results

3.1. Demographics of healthy participants in training and test set

After subdividing UK Biobank into a healthy training set and separate test set, the 2 groups were broadly comparable (Table 1). Age was slightly higher in the test set, which also had a higher proportion of females. The groups were equivocal in terms of blood pressure, BMI, weight, and hip circumference. By design, rates of stroke history and a diagnosis of diabetes were higher in the test set (1.4% and 5.7%, respectively), as these were exclusion criteria for the healthy training set.

3.2. Multimodality neuroimaging can predict chronological age in healthy people

When applying the brain-model model to the validation data set, the correlation between chronological age and brain-predicted age was r = 0.786, R^2 = 0.618, MAE = 3.515, with an age-bias of r = −0.65 (Fig. 1).

From the initial input of 1079 neuroimaging phenotypes, 221 were set to nonzero in the LASSO regression model. After bootstrapping, 34 neuroimaging phenotypes had 95% confidence intervals that did not overlap zero (Table 2). This is despite 886 out of the 1079 neuroimaging phenotypes being significantly correlated with age at $p < 0.05$; even using Bonferroni correction 704 neuroimaging were significantly correlated with age (adjusted $p = 4.6*10^{-5}$).

These 34 "informative" neuroimaging phenotypes were predominantly from T1-weighted MRI or diffusion-MRI, although T2*, task fMRI and resting-fMRI variables were included. Given that the variables were scaled across modalities, the relative beta coefficients from the model can be considered. The variable with the largest absolute coefficient was the volume of gray matter (normalized for head size) beta = −1.466, 95% confidence intervals (−1.85, −1.24), suggesting that lower gray matter volume is associated with higher brain-age estimates. Lower weighted mean intracellular volume fraction (ICVF) in the forceps minor was also associated with higher brain-age estimates, suggesting that neurite density in the anterior corpus callosum and related regions is lower in older-appearing brains. Increasing volume of the fourth ventricle, equating to an enlargement of this CSF space, was also associated with higher brain-age estimates. Other notable associations include reductions in volume and changes in diffusion metrics in the cerebellum.

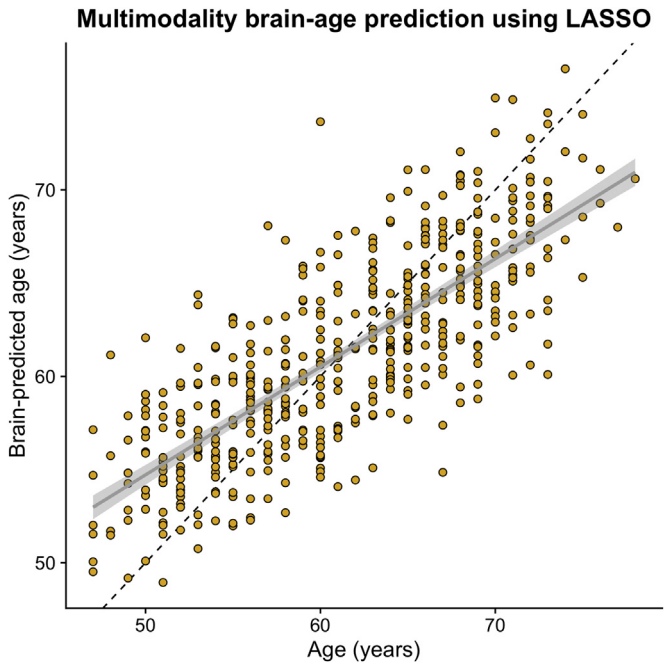

Fig. 1. Brain-predicted age from multimodality LASSO regression model. Scatterplot depicting chronological age (x-axis) by brain-predicted age (y-axis) in UK Biobank validation set (n = 520). Black line is the line of identity. Gray line is the regression line of age on brain-predicted age with shaded errors representing the 95% confidence intervals. No age-bias correction had been applied at this stage.

3.3. Single modality brain-age prediction

Variables from each neuroimaging modality were used separately to train a model to predict age in the validation data set (Table 3). The best performing modalities (i.e., highest correlation with age, lowest MAE) were T1-weighted MRI and diffusion-MRI. The other 4 modalities were only able to explain a limited amount of variance in age, particularly the task fMRI. The pattern of correlations across different "ages" predicted by each modality is shown in Fig. 2. The highest correlation between 2 "ages" is between T1-weighted and diffusion-MRI ($r = 0.73$).

3.4. Leave-one-modality-out brain-age prediction

To further explore the influence of different modalities on brain-age prediction, I conducted a leave-one-modality-out analysis, essentially the inverse of the single modality approach above. The motivation behind this analysis was to test whether the inclusion of some modalities actually decreased model performance. Here, each modality was omitted in turn, and the LASSO regression trained on

Table 1
UK Biobank neuroimaging participant characteristics

Characteristic	Healthy training set	Test set
N	2725	14,701
Age, mean ± SD	61.47 ± 7.21	62.64 ± 7.45
Female, % (n)	49.3% (1343)	53.8% (7914)
Body mass index, median (IQR)	25.94 [23.58–28.68]	25.94 [23.56–28.88]
Weight, kg, median (IQR)	75 [65.6–85.4]	74.6 [65.2–85]
Hip circumference, cm, median (IQR)	100 [95.5–105]	100 [96–106]
Diastolic blood pressure, median (IQR)	79 [72–87]	78 [71–85]
Systolic blood pressure, median (IQR)	139 [126–151.75]	137 [125–150]
ICD-10 diagnosis, % (n)	0% (0)	73% (12,753)
Diabetes, % (n)	0% (0)	5.72% (836)
Stroke, % (n)	0% (0)	1.37% (201)

Table 2
Informative neuroimaging phenotypes consistently for predicting age in a multimodality LASSO regression.

Neuroimaging phenotype	UK Biobank data field #	Modality	Coefficient [95% confidence intervals]
Volume of gray matter normalized for head size	25005	T1-weighted	−1.466 [−1.851, −1.242]
Weighted mean ICVF in the tract forceps minor	25661	Diffusion-MRI	−1.029 [−1.789, −0.699]
Volume of the brain stem fourth ventricle	25025	T1-weighted	0.692 [0.484, 1.139]
Volume of gray matter in the ventral striatum left	25890	T1-weighted	−0.666 [−1.022, −0.393]
Weighted mean L1 in the tract anterior thalamic radiation right	25572	Diffusion-MRI	0.513 [0.218, 1.026]
Mean ISOVF in the fornix on FA skeleton	25445	Diffusion-MRI	0.501 [0.219, 1.003]
Mean FA in the middle cerebellar peduncle on FA skeleton	25056	Diffusion-MRI	−0.477 [−0.91, −0.275]
Volume of gray matter in the putamen right	25883	T1-weighted	0.460 [0.262, 0.911]
Volume of the thalamus right	25012	T1-weighted	−0.454 [−0.909, −0.116]
Mean FA in the cerebral peduncle on FA skeleton left	25071	Diffusion-MRI	−0.453 [−0.906, −0.226]
Mean FA in the superior cerebellar peduncle on FA skeleton left	25069	Diffusion-MRI	0.429 [0.053, 0.859]
Mean L1 in the middle cerebellar peduncle on FA skeleton	25200	Diffusion-MRI	−0.428 [−0.815, −0.162]
Mean L3 in the posterior thalamic radiation on FA skeleton right	25324	Diffusion-MRI	0.418 [0.129, 0.835]
Mean L2 in the fornix cres/stria terminalis on FA skeleton left	25287	Diffusion-MRI	0.393 [0.062, 0.787]
Mean ICVF in the body of the corpus callosum on FA skeleton	25347	Diffusion-MRI	0.381 [0.031, 0.762]
Mean L1 in the anterior limb of the internal capsule on FA skeleton left	25217	Diffusion-MRI	0.349 [0.095, 0.698]
Mean MO in the fornix cres/stria terminalis on FA skeleton left	25191	Diffusion-MRI	−0.336 [−0.61, −0.061]
Volume of gray matter in the VI cerebellum right	25899	T1-weighted	−0.330 [−0.661, −0.059]
Volume of gray matter in the frontal operculum cortex right	25863	T1-weighted	−0.321 [−0.582, −0.146]
Mean OD in the superior longitudinal fasciculus on FA skeleton left	25433	Diffusion-MRI	0.316 [0.169, 0.633]
Volume of gray matter in the vermis crus II cerebellum	25904	T1-weighted	0.306 [0.115, 0.532]
Weighted mean L3 in the tract uncinate fasciculus left	25648	Diffusion-MRI	0.302 [0.002, 0.603]
Volume of the putamen left	25015	T1-weighted	−0.284 [−0.567, −0.011]
Mean MD in the cingulum hippocampus on FA skeleton right	25140	Diffusion-MRI	−0.272 [−0.544, −0.109]
Mean L2 in the splenium of the corpus callosum on FA skeleton	25252	Diffusion-MRI	−0.271 [−0.542, −0.014]
Weighted mean MO in the tract anterior thalamic radiation left	25544	Diffusion-MRI	0.263 [0.065, 0.526]
90th percentile of BOLD effect in a group-defined mask for shapes activation	25761	Task fMRI	0.229 [0.05, 0.425]
Weighted mean MO in the tract forceps minor	25553	Diffusion-MRI	−0.226 [−0.452, −0.031]
Median T2* in the putamen left	25030	T2*	−0.221 [−0.442, −0.016]
Volume of gray matter in the thalamus right	25879	T1-weighted	0.217 [0.005, 0.435]
Median BOLD effect in a group-defined mask for faces-shapes contrast	25048	Task fMRI	−0.212 [−0.423, −0.06]
Weighted mean L1 in the tract parahippocampal part of the cingulum left	25575	Diffusion-MRI	−0.174 [−0.348, −0.001]
Resting-state partial correlation 25-dimension IC variable 157	25752	Resting-state fMRI	0.165 [0.024, 0.329]
Mean L2 in the cingulum cingulate gyrus on FA skeleton right	25282	Diffusion-MRI	−0.119 [−0.237, −0.009]

Key: FA, fractional anisotropy; ICVF, intracellular volume fraction; ISOVF, isotropic volume fraction; MO, mode of anisotropy; OD, orientation dispersion; v#, variable from resting-state fMRI partial correlation matrix, using 25-dimension independent component analysis.

the remaining 5 modalities and then tested on the validation data set (Table 4). Compared with using all 6 modalities, removing the task fMRI or the resting-state fMRI had no appreciable impact on performance. The biggest decrease in performance occurred when excluding diffusion-MRI data, although excluding T1-weighted phenotypes also decreased model accuracy. Notably, model performance was never improved when excluding a single modality.

3.5. Brain predicted-age, health, lifestyle, and cognitive performance

When applied to the n = 14,701 test data set participants from UK Biobank, performance was similar to that in the validation set: r = 0.803, R^2 = 0.644, MAE = 3.555 years (Fig. 3A). A pronounced age bias was apparent (r = −0.640), which was adjusted for using the slope (0.59) and intercept (24.7) of the age on brain-predicted age relationship. After adjustment, the age bias was minimal (Fig. 3B).

Using the bias-adjusted brain-PAD values, the relationships with the selected outcome measures were assessed (Table 5). Each analysis used a linear regression model, adjusting for age, age^2, sex, height, volumetric T1-weighted MRI scaling, and head motion. Resulting *p*-values were false-discovery rate corrected (18 tests). After multiple comparison correction, increasing brain-PAD was associated with: higher blood pressure, a diagnosis of diabetes, a history of stroke, past or present smoking, greater frequency of alcohol intake, lower fluid intelligence, longer duration to complete the alphanumeric path trail 2, fewer correct matrix pattern puzzles complete, and fewer correct Tower rearranging puzzles. There was no association with anthropometric measures, facial aging, physical activity levels, and some of the cognitive tasks.

4. Discussion

Using UK Biobank, chronological age can be accurately predicted in healthy people by combining data from 6 different neuroimaging modalities (T1-weighted MRI, T2-FLAIR, T2*, diffusion-MRI, task fMRI, resting-state fMRI). T1-weighted MRI and diffusion-MRI phenotypes were generally the most informative for age

Table 3
Brain-age prediction performance from single modalities

Modality	Number of entered variables	Correlation between age and brain-age (r)	Variance in age explained (R^2)	Mean absolute error (years)	Age bias (correlation between age and age-difference)
T1-weighted	165	0.684	0.468	4.140	−0.721
T2-FLAIR	1	0.308	0.095	5.653	−0.942
T2*	14	0.329	0.108	5.780	−0.987
Diffusion-MRI	675	0.730	0.533	3.897	−0.638
Task fMRI	14	0.161	0.026	5.929	−0.986
Resting-state fMRI	210	0.444	0.194	5.261	−0.921

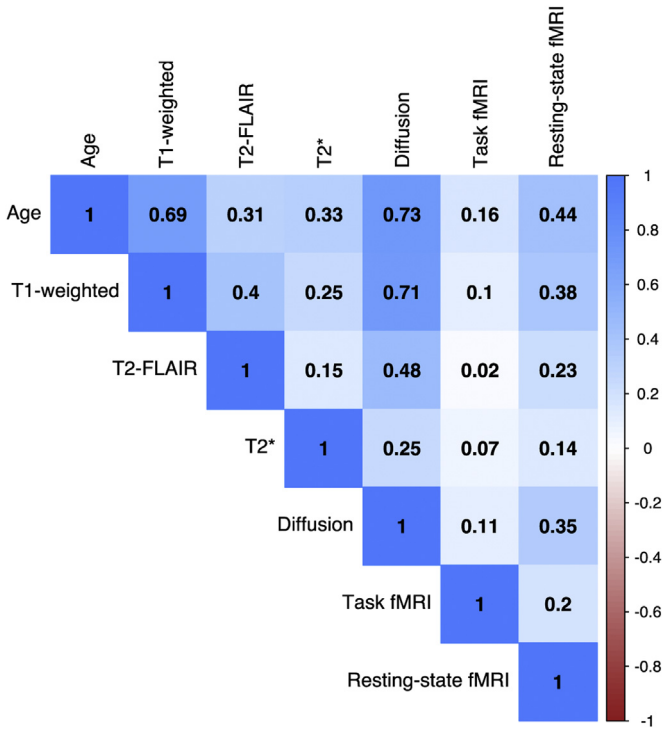

Fig. 2. Correlation matrix of age and brain-age predicted by 6 different modalities. Bivariate correlations between chronological age and brain-age values derived from each of the 6 neuroimaging modalities, in the validation set (n = 520). Values are Pearson's r for each pairwise correlation. Darker blue colors indicate higher positive correlations. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

prediction in this combined model. Bootstrapping highlighted 34 variables as informative, and the multimodality model outperformed prediction models from any modality independently, as hypothesized. This indicates that much of the age-related variation can be captured by these 34 phenotypes alone. Of the independent modality predictors, only T1-weighted MRI and diffusion-MRI achieved reasonable prediction accuracy. Performance was slightly higher for diffusion-MRI, contrary to the expectation that T1-weighted MRI would provide the most accurate separate model. When applying the multimodality model to held-out test data, aberrant brain aging (quantified by brain-PAD) was associated with higher blood pressure, a history of stroke, diabetes, smoking, frequent alcohol intake, and poor cognitive performance.

The majority of extant brain-age studies use T1-weighted MRI alone (Cole et al., 2019b), although previous multimodality studies have used 2 or 3 modalities (Brown et al., 2012; Cherubini et al., 2016; Groves et al., 2012; Liem et al., 2017; Niu et al., 2019; Richard et al., 2018). Thanks to UK Biobank, I was able to combine and compare 6 different modalities. As anticipated, T1-weighted MRI proved important for brain-age prediction here, with normalized gray matter volume being the most informative neuroimaging phenotype. Diffusion-MRI phenotypes were also informative in the LASSO model, and T2* in the putamen, BOLD response during the faces and shapes tasks, and independent components derived from resting-state fMRI were also included. Volumes of white-matter hyperintensities (from T2-FLAIR) were not retained by the model. Several reasons may explain the importance placed on T1-weighted and diffusion-MRI phenotypes, from biological, technical, and statistical perspectives. Firstly, the phenotypes reflect biological processes shown by neuropathology studies to change with age, specifically brain volume loss and myelin changes (Peters, 2002; Savva et al., 2009); hence, the ground-truth relationship is strong. Secondly, these modalities have relatively high signal-to-noise ratio, particularly T1-weighted MRI (Lu et al., 2005; Perthen et al., 2008; Polzehl and Tabelow, 2016), meaning that the age-related signal is less likely to be obfuscated by factors linked to data acquisition. Thirdly, the T1-weighted and diffusion-MRI phenotypes were far more numerous in the data set than the other modalities (e.g., n = 675 diffusion-MRI phenotypes compared with n = 1 for T2-FLAIR). This means that not only is there much greater anatomical specificity (as many of the phenotypes were local ROIs) but also by simple probability there is more chance of these phenotypes detecting the age-related signal in the data.

The use of bootstraps (n = 1000) to derive confidence intervals provides assurance of the robustness of the selected features, despite the relatively small contributory effects of each of the 34 informative phenotypes. Interestingly, the predictive performance of the multimodality model ($r = 0.79$) was lower than some previous reports using multiple modalities (e.g., Cherubini et al., $r = 0.96$), despite the much larger training sample size (n = 2205 versus n = 140). The current use of derived summary measures (i.e., imaging-derived phenotypes) may negatively impact brain-age prediction performance since these summary measures collapse complex high-dimensional information into single values, potentially removing considerable amounts of age-related variance. In fact, brain-age prediction performance using voxelwise T1-weighted MRI data alone tends to be highly accurately (e.g., r = 0.96; Cole et al., 2017a). Future work will use the raw neuroimaging data from each UK Biobank modality to fuse multimodality data in a more high-dimensional manner.

One important finding of the current study is that although many of the 1079 neuroimaging phenotypes significantly correlated with age, with this large sample (n = 2725) correlations with $r > 0.05$ will be significant at $p < 0.05$. Hence, many of these statistically significant effects are so small as to be negligible, and only 34 phenotypes were selected as being informative for age-prediction. This means that although multiple patterns of brain structure and function are statistically associated with age, only a few measures are potentially suitably for deriving individualized brain-age predictions with sufficient accuracy for clinical utility. For brain-age, as with other research, it is important to be mindful of the differences

Table 4
Brain-age prediction performance, leaving out single modalities

Excluded modality	Number of entered variables	Correlation between age and brain-age (r)	Variance in age explained (R^2)	Mean absolute error (y)	Age bias (correlation between age and age-difference)
All included	1079	0.786	0.618	3.515	−0.650
T1-weighted	914	0.751	0.565	3.752	−0.654
T2-FLAIR	1078	0.778	0.605	3.572	−0.642
T2*	1065	0.773	0.598	3.598	−0.641
Diffusion-MRI	404	0.715	0.511	3.975	−0.692
Task fMRI	1065	0.781	0.610	3.569	−0.634
Resting-state fMRI	869	0.779	0.607	3.522	−0.615

Table 5
Biomedical, lifestyle, and cognitive measures in relation to brain-age

Measure	UK Biobank data field #	Estimate	Standard error	T-value	*p*	FDR corrected-P	ηp^2
Biomedical							
Diastolic blood pressure	4079	0.049	0.005	9.520	<0.001	<0.001	0.0074
Systolic blood pressure	4080	0.028	0.003	9.433	<0.001	<0.001	0.0073
Body mass index	21001	0.008	0.013	0.646	0.518	1.000	0
Weight	21002	0.006	0.005	1.400	0.162	1.000	0.0001
Hip circumference	49	−0.080	0.006	−1.381	0.167	1.000	0.0001
Diabetes[a]	2443	2.115	0.207	10.208	<0.001	<0.001	0.0071
Stroke[a]	4056	2.695	0.407	6.614	<0.001	<0.001	0.0030
Facial aging[a]	1757	−0.550	0.462	−1.201	0.230	1.000	0.0001
Lifestyle							
Smoking status[a]	20116	0.879	0.102	8.636	<0.001	<0.001	0.0070
Alcohol intake frequency[a]	1558	−0.997	0.147	−6.776	<0.001	<0.001	0.0090
Duration of moderate activity	894	0.001	0.001	0.279	0.780	1.000	0
Duration of vigorous activity	914	−0.001	0.001	−0.461	0.645	1.000	0
Cognitive performance							
Fluid intelligence	20016	−0.147	0.024	−5.998	<0.001	<0.001	0.0027
Trail making task: duration to complete numeric path trail 1	6348	0.003	0.001	2.712	0.007	0.121	0.0012
Trail making task: duration to complete alphanumeric path trail 2	6350	0.002	0.001	5.667	<0.001	<0.001	0.0054
Matrix pattern completion: number of puzzles correctly solved	6373	−0.218	0.037	−5.882	<0.001	<0.001	0.0059
Matrix pattern completion: duration spent answering each puzzle	6333	0.010	0.007	1.452	0.147	1.000	0.0004
Tower rearranging: number of puzzles correct	6382	−0.117	0.021	−5.468	<0.001	<0.001	0.0050

All other variables are continuous.
Key: FDR, false discovery rate. ηp^2, partial eta-squared effect size.
[a] Indicates categorical variable.

between statistical significance, effect size, and predictive power in new data.

Another noteworthy issue is the validity of the processing methods for deriving the neuroimaging phenotypes. For example, in the current study, subcortical gray matter volume phenotypes were generated in 2 ways; using FSL FIRST (Patenaude et al., 2011) and FSL FAST (Zhang et al., 2001), the latter with a subcortical regional mask applied to the tissue-segmented image. Despite theoretically measuring the same underlying construct, the correlation between FIRST thalamus volume and FAST thalamus volume, for example, is only $r = 0.44$ and $r = 0.41$ for the left and right, respectively. This lack of internal consistency highlights the need for researchers who use "big data" resources such as UK Biobank to carefully consider the nuances of the data processing and to select phenotypes explicitly, as opposed to blindly analyzing all available data. This is an important caveat when interpreting the current and other works using these data. Proper validation will come from not only using truly independent data sets but also different processing methods to converge toward the same constructs (i.e., brain region segmentation), thus overcoming the assumptions and errors of each separate method.

By using healthy people only for training, a new individual's deviation from this healthy brain aging model can be indexed. This key point makes a distinct between models trained on broad inclusive population samples and models trained specifically on healthy people. With the latter, the goal is always to improve model prediction, while with the former, improved model prediction may mean that anyone's age can be accurately predicted, but the concept of deviation loses value. Previous research has not always screened for healthy people during model training, or at least has not

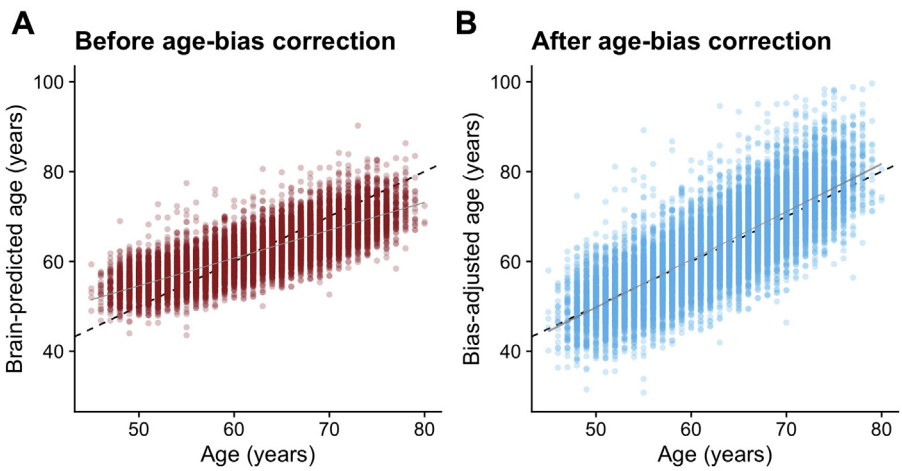

Fig. 3. Brain-predicted age by chronological age in the test set, with and without age-bias adjustment. Scatterplots depicting chronological age (x-axis) by brain-predicted age (y-axis) in UK Biobank test set (n = 14,701). Black line is the line of identity. Gray line is the regression line of age on brain-predicted age with shaded areas representing the 95% confidence intervals. A) Brain-predicted age values generated from application of previously trained LASSO regression model. The age bias is evident from the slope of the regression line. B) Brain-predicted age values have been adjusted by the slope and intercept of the age-bias line in the training set.

reported that this was the case (Smith et al., 2019; Sun et al., 2019). This means that potentially the brain-age difference metrics (i.e., brain-age gap, brain-age delta) do not reflect actual deviations from healthy brain aging and thus may be less sensitive to subsequent relationships with other measures or more prone to false positives. The current study has a large sample with detailed biomedical data, allowing thorough screening of the healthy training data set. Although it is true that even some of these healthy participants may have undetected or prodromal pathologies, this is the case for all case-control research and can only be properly assessed using long-term follow-up. This will also be possible in future, thanks to the study design of UK Biobank.

An older-appearing brain was associated with higher diastolic and systolic blood pressure, a history of stroke, a diagnosis of diabetes, smoking, alcohol intake, and some facets of cognitive performance. BMI, weight, hip circumference, facial aging, and other aspects of cognitive performance were not associated with brain-PAD. These findings partially concur with the previous reports. Franke and colleagues (2014) reported brainAGE (equivalent to brain-PAD) associations with diastolic blood pressure and BMI; here only the former was replicated. Obesity has previously been linked with added brain-aging (Kolenic et al., 2018; Ronan et al., 2016), and BMI has been shown to influence the brain structure (Cole et al., 2013), although no BMI and brain-PAD association was found here. Exercise duration was not associated with brain-PAD, contrary to the findings of Steffener and colleagues (2016). Nevertheless, many different approaches to measuring physical activity are possible, with the current self-report measures not necessarily being the most valid.

Regarding medical history, Franke and colleagues (2013) found that diabetes was associated with an increased brainAGE = 4.6 years, whereas here diabetes increased brain-PAD by 2 years (after covariate adjustment). Stroke was also associated with an increased brain-PAD = 2.6 years; interestingly neither diabetes nor stroke were associated with brain-PAD in the Lothian Birth Cohort 1936 (Cole et al., 2018). Discrepancies between the current and previous works may be due to sampling biases, such as cohort effects or recruitment strategies, although statistical power is certainly in favor of the current study, thanks to UK Biobank's large sample size. Another reason may be the analytic strategy used, particularly the use of covariates. Here, multiple covariates were used to increase sensitivity to brain-PAD relationships, most notably head motion (averaged across the task fMRI session), which accounted for 4.7% of the variance in brain-PAD. Although several biomedical, lifestyle, and cognitive factors related to brain-PAD, effect sizes were generally small. For instance, the effects of diabetes, stroke, current smoking (+2.1 years brain-PAD), and daily alcohol intake (+0.97 years) are substantially smaller than previously reported effects of Alzheimer's (+10 years) or multiple sclerosis (+10 years) on apparent brain-age (Cole et al., 2020; Franke et al., 2010). Although it makes intuitive sense that such factors are associated with poorer aging-related brain health, the magnitude of the effects should not be overstated. Future work will investigate gene-environment interactions to better understand individual differences in the impact of medical history and lifestyle on brain health during aging.

Fluid intelligence has previously been linked to brain-PAD (Cole et al., 2018), as has performance on the trail-making task (Cole et al., 2015, 2017b), and other studies report moderate significant relationships between brain-aging and cognitive performance (Liem et al., 2017; Richard et al., 2018). Here I was able to replicate that relationship, although more detailed analysis of how brain-aging relates to cognitive aging will require more comprehensive cognitive testing. The presence of data on facial aging presented an interesting opportunity to test the relationship between brain-PAD and self-report of whether participants felt people said they looked "younger than you are," "older than you are," or "about the same age." There was no relationship detected, although interesting related work has reported a relationship between subjective "age" and brain-age (Kwak et al., 2018). In the current study, only 1% of respondents reported that people think they look older, potentially questioning the validity of this subjective measure.

The current study benefits from the extremely large sample size and in-depth biomedical data of UK Biobank. This enabled detailed screening of healthy individuals to include in the training data set, to minimize the impact of disease-related effects on the brain that may potentially confound models of healthy brain aging. Despite 84% of the neuroimaging sample from UK Biobank used herein having at least one ICD-10 diagnosis, a self-reported longstanding illness, poor self-rated health status, or medical history of stroke or diabetes, the remaining 16% still left greater than 2500 participants for training and validation. Another benefit here is the use of LASSO regression with bootstrapping, combining the strength of LASSO for identifying important features for prediction (by shrinking uninformative variables to zero) with the robustness of bootstrapping, which overcomes the limitation of using LASSO with highly correlated predictor variables. A weakness of the study is the use of summary-level neuroimaging phenotypes. As noted, model accuracy is substantially below that commonly achieved with voxelwise data. Another potential limitation is the current limited use of biomedical, lifestyle, and cognitive variables. Potentially many more associations with biological, psychological, and behavioral parameters could have been assessed; however, adding further tests adds to the multiple-testing burden and decreases sensitivity to real effects. Hence, the decision was taken to test only variables with previous research evidence and strong face validity. Another weakness is that the current study is cross-sectional; hence, the long-term consequences of having an older-appearing brain cannot be tested and no causal inference made. The planned long-term medical follow-up of UK Biobank participants makes this an obvious avenue for future research.

5. Conclusions

In summary, brain-age can be predicted by combining 6 different MRI modalities, with the strongest predictors being T1-weighted and diffusion-MRI phenotypes. Deviations from healthy brain aging are related to medical history, smoking, and alcohol intake. Poorer cognitive performance was also related to having an older-appearing brain, suggesting that biomedical and lifestyle factors can negatively impact brain aging and cognitive aging. The brain-age paradigm presents a potential screening tool to detect the negative impact of medical and lifestyle factors on brain health in asymptomatic people, and the planned longitudinal nature of UK Biobank offers the opportunity to validate this potential.

Disclosure statement

James Cole is a scientific advisor to and shareholder in Brain Key, a medical image analysis software company.

Acknowledgements

The author would like to thank the UK Biobank participants for giving up their time for this unique project.

Funding sources: James Cole is funded by a UK Research & Innovation (UKRI) Innovation Fellowship (MR/R024790/1; MR/R024790/2).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neurobiolaging.2020.03.014.

References

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. J. Roy Stat. Soc. Series B Stat. Methodol. 57, 289–300.

Brown, T., Kuperman Joshua, M., Chung, Y., Erhart, M., McCabe, C., Hagler Jr., D.J., Venkatraman Vijay, K., Akshoomoff, N., Amaral David, G., Bloss Cinnamon, S., Casey, B.J., Chang, L., Ernst Thomas, M., Frazier Jean, A., Gruen Jeffrey, R., Kaufmann Walter, E., Kenet, T., Kennedy David, N., Murray Sarah, S., Sowell Elizabeth, R., Jernigan Terry, L., Dale Anders, M., 2012. Neuroanatomical assessment of biological maturity. Curr. Biol. 22, 1693–1698.

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

Cole, J.H., Boyle, C.P., Simmons, A., Cohen-Woods, S., Rivera, M., McGuffin, P., Thompson, P.M., Fu, C.H.Y., 2013. Body mass index, but not FTO genotype or major depressive disorder, influences brain structure. Neuroscience 252, 109–117.

Cole, J.H., Franke, K., 2017. Predicting age using neuroimaging: innovative brain ageing biomarkers. Trends Neurosci. 40, 681–690.

Cole, J.H., Franke, K., Cherbuin, N., 2019b. Quantification of the biological age of the brain using neuroimaging. In: Moskalev, A. (Ed.), Biomarkers of Human Aging. Springer International Publishing, Cham, pp. 293–328.

Cole, J.H., Leech, R., Sharp, D.J., 2015. For the Alzheimer's Disease Neuroimaging Initiative. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. Ann. Neurol. 77, 571–581.

Cole, J.H., Marioni, R.E., Harris, S.E., Deary, I.J., 2019c. Brain age and other bodily 'ages': implications for neuropsychiatry. Mol. Psychiatry 24, 266–281.

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

Cole, J., Raffel, J., Friede, T., Eshaghi, A., Brownlee, W., Chard, D., De Stefano, N., Enzinger, C., Pirpamer, L., Filippi, M., Gasperini, C., Rocca, M., Rovira, A., Ruggieri, S., Sastre-Garriga, J., Stromillo, M., Uitdehaag, B., Vrenken, H., Barkhof, F., Nicholas, R., Ciccarelli, O., MAGNIMS Study group, 2020. Longitudinal assessment of multiple sclerosis with the brain-age paradigm. Annals of Neurology. https://doi.org/10.1002/ana.25746. https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.25746, 2020. Epub ahead of print.

Cole, J.H., Ritchie, S.J., Bastin, M.E., Valdes Hernandez, M.C., Munoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S.E., Zhang, Q., Wray, N.R., Redmond, P., Marioni, R.E., Starr, J.M., Cox, S.R., Wardlaw, J.M., Sharp, D.J., Deary, I.J., 2018. Brain age predicts mortality. Mol. Psychiatry 23, 1385–1392.

Cole, J.H., Underwood, J., Caan, M.W., De Francesco, D., van Zoest, R.A., Leech, R., Wit, F.W., Portegies, P., Geurtsen, G.J., Schmand, B.A., Schim van der Loeff, M.F., Franceschi, C., Sabin, C.A., Majoie, C.B., Winston, A., Reiss, P., Sharp, D.J., 2017b. Increased brain-predicted aging in treated HIV disease. Neurology 88, 1349–1357.

Deary, I.J., Corley, J., Gow, A.J., Harris, S.E., Houlihan, L.M., Marioni, R.E., Penke, L., Rafnsson, S.B., Starr, J.M., 2009. Age-associated cognitive decline. Br. Med. Bull. 92, 135–152.

Franke, K., Gaser, C., 2012. Longitudinal changes in individual BrainAGE in healthy aging, mild cognitive impairment, and Alzheimer's Disease. J. GeroPsych. Geriatr. Psychiatry 25, 235–245.

Franke, K., Gaser, C., 2019. Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained? Front. Neurol. 10, 789.

Franke, K., Gaser, C., Manor, B., Novak, V., 2013. Advanced BrainAGE in older adults with type 2 diabetes mellitus. Front. Aging Neurosci. 5, 90.

Franke, K., Ristow, M., Gaser, C., 2014. Gender-specific impact of personal health parameters on individual brain aging in cognitively unimpaired elderly subjects. Front. Aging Neurosci. 6, 94.

Franke, K., Ziegler, G., Klöppel, S., Gaser, C., 2010. For the Alzheimer's Disease Neuroimaging Initiative. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. Neuroimage 50, 883–892.

Gaser, C., Franke, K., Klöppel, S., Koutsouleris, N., Sauer, H., 2013. For the Alzheimer's Disease Neuroimaging Initiative. BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer's disease. PLoS One 8, e67346.

Groves, A.R., Smith, S.M., Fjell, A.M., Tamnes, C.K., Walhovd, K.B., Douaud, G., Woolrich, M.W., Westlye, L.T., 2012. 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, 365–380.

Hanscombe, K.B., Coleman, J., Traylor, M., Lewis, C., 2019. Ukbtools: an R package to manage and query UK Biobank data. PLoS One 14.

Jackson, S.H.D., Weale, M.R., Weale, R.A., 2003. Biological age - what is it and can it be measured? Arch. Gerontol. Geriatr. 36, 103–115.

Kolenic, M., Franke, K., Hlinka, J., Matejka, M., Capkova, J., Pausova, Z., Uher, R., Alda, M., Spaniel, F., Hajek, T., 2018. Obesity, dyslipidemia and brain age in first-episode psychosis. J. Psychiatr. Res. 99, 151–158.

Kwak, S., Kim, H., Chey, J., Youm, Y., 2018. Feeling how old I Am: subjective age is associated with estimated brain age. Front. Aging Neurosci. 10, 168.

Le, T.T., Kuplicki, R.T., McKinney, B.A., Yeh, H.-W., Thompson, W.K., Paulus, M.P., , TI, Aupperle, R.L., Bodurka, J., Cha, Y.-H., Feinstein, J.S., Khalsa, S.S., Savitz, J., Simmons, W.K., Victor, T.A., 2018. A nonlinear simulation framework supports adjusting for age when analyzing BrainAGE. Front. Aging Neurosci. 10, 317.

Liang, H., Zhang, F., Niu, X., 2019. Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders. Hum. Brain Mapp. 40, 3143–3152.

Liem, F., Varoquaux, G., Kynast, J., Beyer, F., Kharabian Masouleh, S., Huntenburg, J.M., Lampe, L., Rahim, M., Abraham, A., Craddock, R.C., Riedel-Heller, S., Luck, T., Loeffler, M., Schroeter, M.L., Witte, A.V., Villringer, A., Margulies, D.S., 2017. Predicting brain-age from multimodal imaging data captures cognitive impairment. Neuroimage 148, 179–188.

Lu, H., Nagae-Poetscher, L.M., Golay, X., Lin, D., Pomper, M., van Zijl, P.C.M., 2005. Routine clinical brain MRI sequences for use at 3.0 Tesla. J. Magn. Reson. Imaging 22, 13–22.

Ludwig, F.C., Smoke, M.E., 1980. The measurement of biological age. Exp. Aging Res. 6, 497–522.

Mendez, M.F., 2017. Early-onset Alzheimer disease. Neurol. Clin. 35, 263–281.

Miller, K.L., Alfaro-Almagro, F., Bangerter, N.K., Thomas, D.L., Yacoub, E., Xu, J., Bartsch, A.J., Jbabdi, S., Sotiropoulos, S.N., Andersson, J.L.R., Griffanti, L., Douaud, G., Okell, T.W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., Matthews, P.M., Smith, S.M., 2016. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat. Neurosci. 19, 1523–1536.

Niu, X., Zhang, F., Kounios, J., Liang, H., 2019. Improved prediction of brain age using multimodal neuroimaging data. Hum. Brain Mapp. 41, 1626–1643.

Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 56, 907–922.

Perthen, J.E., Bydder, M., Restom, K., Liu, T.T., 2008. SNR and functional sensitivity of BOLD and perfusion-based fMRI using arterial spin labeling with spiral SENSE at 3 T. Magn. Reson. Imaging 26, 513–522.

Peters, A., 2002. The effects of normal aging on myelin and nerve fibers: a review. J. Neurocytol 31, 581–593.

Polzehl, J., Tabelow, K., 2016. Low SNR in diffusion MRI models. J. Am. Stat. Assoc. 111, 1480–1490.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Richard, G., Kolskar, K., Sanders, A.M., Kaufmann, T., Petersen, A., Doan, N.T., Monereo Sanchez, J., Alnaes, D., Ulrichsen, K.M., Dorum, E.S., Andreassen, O.A., Nordvik, J.E., Westlye, L.T., 2018. Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. PeerJ 6, e5908.

Ronan, L., Alexander-Bloch, A.F., Wagstyl, K., Farooqi, S., Brayne, C., Tyler, L.K., Fletcher, P.C., 2016. Obesity associated with increased brain age from midlife. Neurobiol. Aging 47, 63–70.

Savva, G.M., Wharton, S.B., Ince, P.G., Forster, G., Matthews, F.E., Brayne, C., 2009. Age, neuropathology, and dementia. N. Engl. J. Med. 360, 2302–2309.

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

Steffener, J., Habeck, C., O'Shea, D., Razlighi, Q., Bherer, L., Stern, Y., 2016. Differences between chronological and brain age are related to education and self-reported physical activity. Neurobiol. Aging 40, 138–144.

Sun, H., Paixao, L., Oliva, J.T., Goparaju, B., Carvalho, D.Z., van Leeuwen, K.G., Akeju, O., Thomas, R.J., Cash, S.S., Bianchi, M.T., Westover, M.B., 2019. Brain age from the electroencephalogram of sleep. Neurobiol. Aging 74, 112–120.

Wasay, M., Grisold, W., Carroll, W., Shakir, R., 2016. World Brain Day 2016: celebrating brain health in an ageing population. Lancet Neurol. 15, 1008.

Wyss-Coray, T., 2016. Ageing, neurodegeneration and brain rejuvenation. Nature 539, 180–186.

Zhang, Y.Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans. Med. Imaging 20, 45–57.