Learning patterns of the ageing brain in MRI using deep convolutional networks

Emma Bluemke[1,2], Stephen Smith[3], Zobair Arya[3], Diego Vidaurre[3,4], Mark Jenkinson[3], Ana I. L. Namburete[2]

[1]EPSRC & MRC Centre for Doctoral Training in Biomedical Imaging;[2] Institute of Biomedical Engineering, Department of Engineering Science; [3]Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences; [4]Department of Psychiatry, University of Oxford

Abstract

Motivation:

- Both normal ageing and neurodegenerative diseases cause morphological changes of the brain. Age-related changes to the brain are subtle, nonlinear, and spatially and temporally heterogeneous both within a subject and across a population.
- This presents a challenge for discerning normal ageing from early stages of disease, as many diseases can appear as advanced ageing early on.
- **The ability to differentiate early signs of disease from normal ageing holds great clinical value, as many current therapies are most effective with early intervention. To distinguish between normal and abnormal cases, it is necessary to develop a model of the healthy ageing brain.**
- Machine learning models are particularly suited to address problems with these properties, and can produce a model that is sensitive to the large variety in healthy brain appearance.

Approach:

In this project, we harness the power of convolutional neural networks and the rich UK Biobank to produce a 'brain age' prediction model. We developed a convolutional neural network, designed to predict age, and used 14,490 T1-weighted MRI images as a 3D input. We use the prediction 'error':

Delta = Predicted Age – Actual Age

as a measurement for 'brain age'. For the purpose of exploration, we examined the relationship between these deltas and:

- **Subject measurements in the Biobank (i.e. Clinical, Lifestyle)**
- **Imaging Derived Phenotypes (IDPs) from all other imaging modalities in the UK Biobank**

Results:

- The 'deltas' from the CNN correlated weakly to 96 and 60 clinical measurements from the Biobank in the female and male groups, respectively.
- The 'deltas' showed weak correlation to 698 and 694 IDPs from dMRI, T1-weighted, and T2 FLAIR images in the female and male groups, respectively. These relationships can give us insight into which modalities could be most useful for future age prediction models.
- Due to the longitudinal aspect of the UK Biobank study, in the future we will have the ability to look at whether the deltas from models such as this were predictive of any health outcomes.

Data

Data processing:

- 14,490 T1-weighted structural MRI from the UK Biobank database (age distribution shown in Figure 2) were used in this experiment.
- All non-brain tissues were removed from the images (FSL BET) and were both linearly aligned and nonlinearly warped to the standard MNI152 brain template (FSL FLIRT, FNIRT). 20 slices used shown in **Figure 2.**
- The subjects were split into two sex groups to train two separate models.

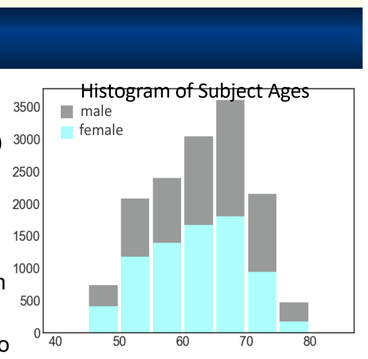

Figure 1. The distribution of ages across all 14,490 subjects. Minimum age of 44 y, maximum 80 y, mean 63.7y. Female mean: 62.0 y, male mean: 63.4 y.

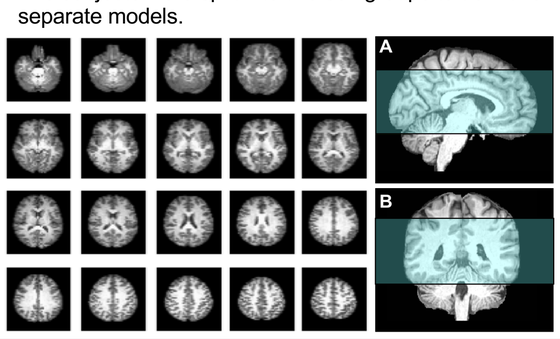

Figure 2. An example of structures contained in the axial view (128x128 pixels) of the 20 slices used. The region covered by the extracted slices is illustrated in the (A) sagittal and (B) coronal planes.

Results: Age Prediction

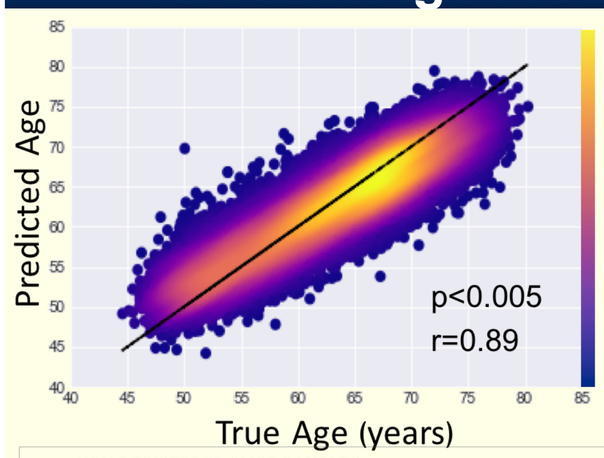

Figure 6. A density plot of 14,490 predicted ages vs the true subject age (r=0.89). The 3D CNN achieved a MAE of 2.97 and 3.26 on the female and male groups, respectively (r=0.89,0.88),

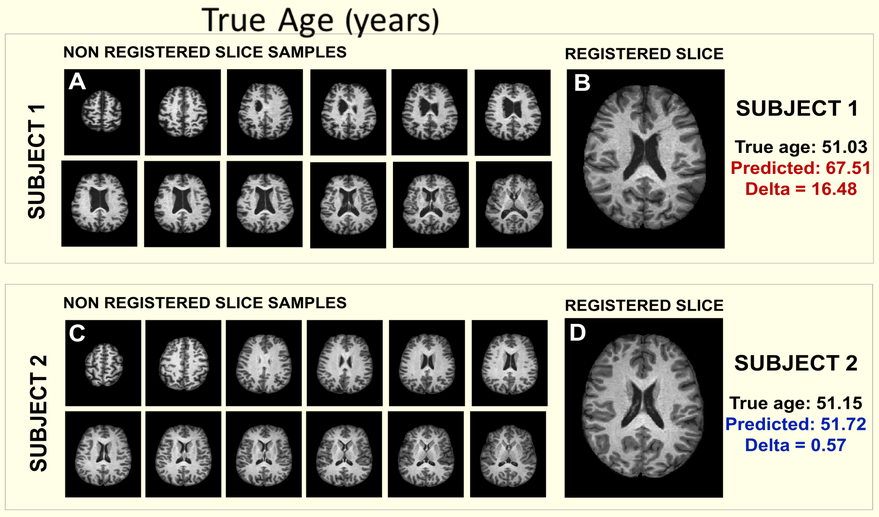

Figure 7. The MRI images from two subjects, both 51 years old. B and D: the non-linearly registered T1 slices of Subject 1 and Subject 2 are displayed. A and C: 12 slices from the original non-registered brain of each subject.

Although the MAE and MSE scores in this experiment are comparable to the current literature, it is important to note that achieving the lowest MAE on this UK Biobank dataset, would not yield the most clinically relevant model, since all 14,490 subjects are not necessarily healthy. This concept is most easily illustrated by Subjects 1 and 2 in Figure 7.

Method

CNN Architecture and Training:

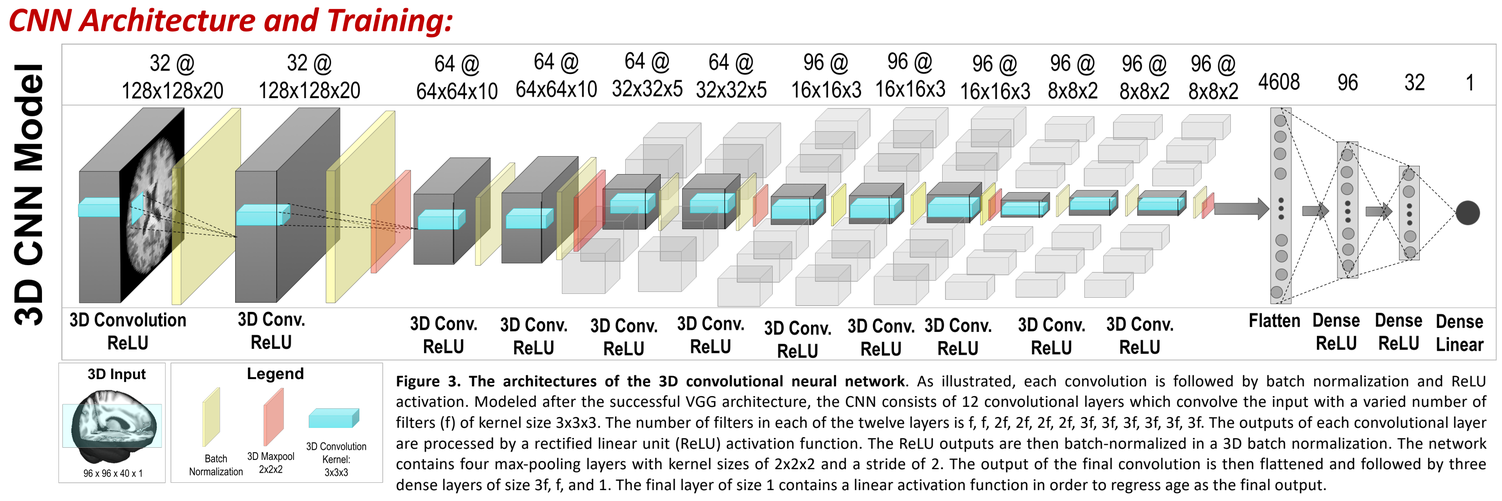

Figure 3. The architectures of the 3D convolutional neural network. As illustrated, each convolution is followed by batch normalization and ReLU activation. Modeled after the successful VGG architecture, the CNN consists of 12 convolutional layers which convolve the input with a varied number of filters (f) of kernel size 3x3x3. The number of filters in each of the twelve layers is f, f, 2f, 2f, 2f, 2f, 3f, 3f, 3f, 3f, 3f, 3f. The outputs of each convolutional layer are processed by a rectified linear unit (ReLU) activation function. The ReLU outputs are then batch-normalized in a 3D batch normalization. The network contains four max-pooling layers with kernel sizes of 2x2x2 and a stride of 2. The output of the final convolution is then flattened and followed by three dense layers of size 3f, f, and 1. The final layer of size 1 contains a linear activation function in order to regress age as the final output.

- The training of both the 2D and 3D network was implemented using the Keras (v2.1.5) framework with a Tensorflow (v1.8.0) backend and performed on an NVIDIA Tesla P100 GPU.
- Optimization was achieved using the Keras RMSprop algorithm and mean squared error loss function over a maximum of 1000 epochs. To minimize overfitting, the mean squared error (MSE) on the validation set was monitored during training and the network was set to stop training if the validation MSE did not improve for 30 epochs.
- The initial learning rate was 10^{-2} and was set to decrease by a factor of 2 if the validation MSE did not decrease over 15 epochs, with a minimum learning rate of 10^{-6}. In every epoch, both the 2D and 3D models were trained with a batch size of 16.
- In the evaluation on each sex group, 10% of the subjects were reserved for cross-validation testing, while the remaining 90% was split five-fold for training and validation during training.

Correlation with Biobank Variables & IDPs:

- Sixteen UK Biobank variable categories were chosen to be of interest, which include:
  - Six lifestyle factors (***Early Lifestyle Factors, Lifestyle General, Exercise and Work, Alcohol, Tobacco***),
  - Eight physiological or medical measurements (***Physical General, Bone Density and Size, Cardiac, Blood Assays, Eye Test, Physical Activity Measures, Cognitive Phenotypes***),
  - and all variables contained in the **Medical History** and **Mental Health Self-Report** categories.
- Univariate statistics were carried out using Pearson correlation analyses between the deltas and 5456 variables from the UK Biobank database, and between the deltas and 941 IDPs collected by Miller *et al*[1].
- The variables examined had all confounding variables (such as sex, age, head size, head motion, scanner table position, and imaging centre) removed and then gaussianized (inormal Gaussianization). The resulting p-values were sorted by category and plotted together with the Bonferroni correction threshold, shown in **Results: UK Biobank**.

Results: UK Biobank Relationships

Correlations with Biobank Variables:

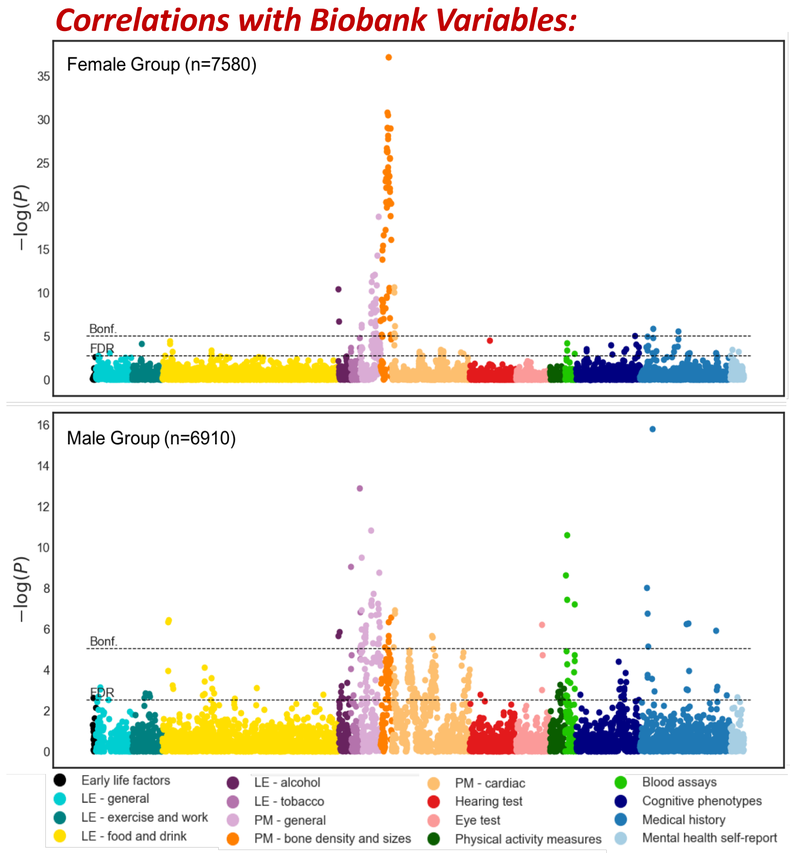

Figure 4. Manhattan plot relating each of the non-imaging Biobank variables with the delta measurements from the 3D CNN on female and male subjects. Each dot represents the statistical significance of the correlation between one Biobank variable and the deltas. The 'deltas' from the CNN correlated weakly (passing Bonferroni threshold shown with dotted line - corresponding to $|r|$=0.26) to 96 and 60 clinical measurements from the Biobank in the female and male groups, respectively.

Correlations with IDPs:

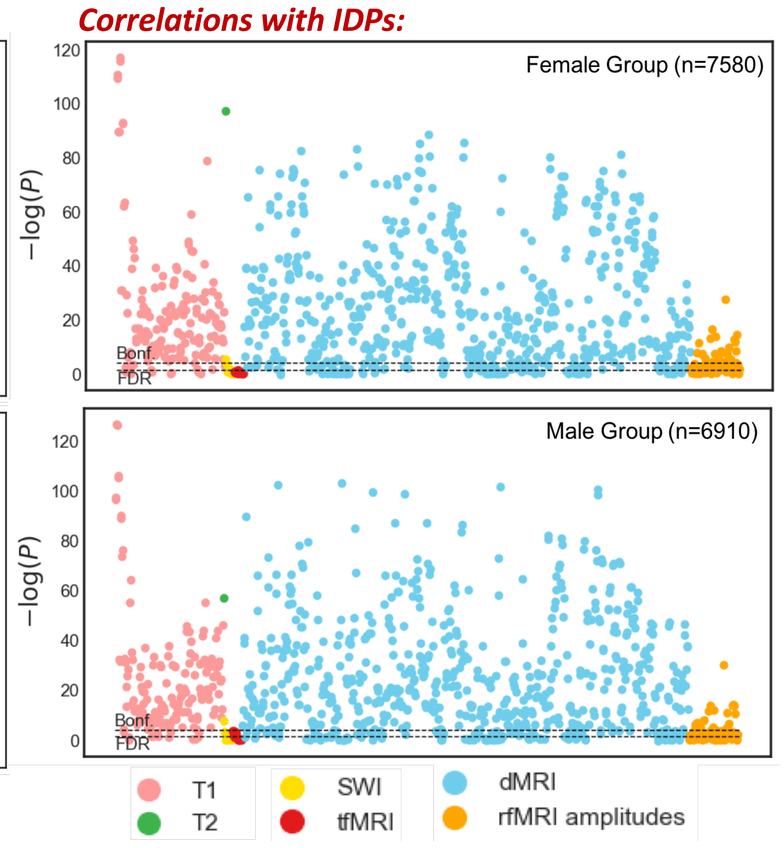

Figure 5. Manhattan plot relating each of the IDPs with the delta measurements from the 3D CNN on female and male subjects. Each dot represents the statistical significance of the correlation between one IDP type and the deltas. Total number of IDPs tested n=941. Interestingly, although our network was only trained using T1 images, the deltas showed weak correlation to 698 IDPs (Bonferroni threshold (corresponding to $|r|$=0.15) from dMRI, T1, and T2 FLAIR images.

Correlations with IDPs, a closer look:

dMRI Tract-Based Spatial Statistics IDPs: Correlation with Delta
Each TBSS IDP passing Bonferroni Threshold and Respective Pearson r value

Tract Name (TBSS)		Female FA	ICVF	ISOVF	L1	L2	L3	MD	MO	OD	Total	Male FA	ICVF	ISOVF	L1	L2	L3	MD	MO	OD	Total
Posterior limb of internal capsule	L	-0.1	-0.2		0.09		0.16	0.15	0.09	-0.1	7		-0.1	0.06	0.17		0.13	0.16	0.15	-0.1	7
	R		-0.2		0.11		0.15	0.15	0.09	-0.1	6		-0.1	0.08	0.17		0.13	0.17	0.15	-0.1	7
Anterior limb of internal capsule	L	-0.1	-0.2	0.07	0.13	0.08	0.2	0.17	0.14		8	-0.1	-0.2	0.06	0.16		0.2	0.17	0.19		7
	R	-0.1	-0.2		0.13	0.06	0.2	0.16	0.15		7	-0.1	-0.2	0.06	0.17		0.21	0.17	0.21		7
Retrolenticular part of internal capsule	L	-0.1	-0.2		0.12	0.1	0.14	0.15			6		-0.1		0.13	0.08	0.11	0.14			5
	R	-0.1	-0.2	0.05	0.15	0.09	0.13	0.15		-0.1	8		-0.1		0.13	0.07	0.12	0.13			5
Anterior corona radiata	L	-0.2	-0.2	0.05	0.1	0.21	0.22	0.2	-0.1	0.08	9	-0.2	-0.2	0.09	0.12	0.22	0.23	0.21	-0.1	0.08	9
	R	-0.2	-0.2		0.12	0.19	0.21	0.2			6	-0.2	-0.2		0.11	0.2	0.22	0.2	-0.1	0.05	8
Posterior corona radiata	L	-0.1	-0.2	0.11	0.22	0.14	0.22	0.22	0.11	-0.1	9	-0.1	-0.2	0.12	0.22	0.12	0.2	0.2	0.13	-0.1	9
	R	-0.1	-0.2	0.12	0.23	0.12	0.22	0.22	0.13	-0.1	9	-0.1	-0.2	0.14	0.25	0.08	0.2	0.21	0.17	-0.2	9
Superior corona radiata	L	-0.2	-0.2	0.09	0.16	0.13	0.24	0.22	0.1		8	-0.1	-0.2	0.13	0.19	0.11	0.22	0.22	0.12	-0.1	9
	R	-0.1	-0.2	0.11	0.16	0.14	0.23	0.22	0.09		8	-0.1	-0.2	0.13	0.18	0.11	0.21	0.21	0.11	-0.1	9
Superior Longitudinal Fasciculus	L	-0.1	-0.2	0.07	0.16	0.15	0.17	0.2			7	-0.1	-0.2	0.08	0.17	0.12	0.16	0.18			7
	R	-0.1	-0.2	0.09	0.18	0.15	0.18	0.2			7	-0.1	-0.2	0.11	0.17	0.12	0.15	0.18			7
Superior fronto-occipital fasciculus	L	-0.2	-0.2	0.05	0.16	0.18	0.21	0.22			7	-0.2	-0.3	0.09	0.17	0.21	0.22	0.22	-0.1		8
	R	-0.1	-0.2		0.13	0.17	0.2	0.2			6	-0.2	-0.3	0.07	0.15	0.21	0.22	0.23			7
Uncinate fasciculus	L	-0.1	-0.2		0.06	0.11	0.09	0.12	-0.1		7	-0.1	-0.2		0.07	0.11	0.09	0.12			6
	R	-0.1	-0.1		0.07	0.11	0.09	0.13	-0.1		7		-0.2	0.06	0.11	0.12	0.1	0.16			6
External capsule	L	-0.1	-0.2		0.12	0.18	0.13	0.16	-0.1		7	-0.1	-0.2		0.13	0.18	0.14	0.17	-0.1		7
	R	-0.1	-0.2		0.12	0.19	0.14	0.17	-0.1		7	-0.1	-0.2	0.05	0.14	0.2	0.15	0.19	-0.1		8
Tapetum	L	-0.1	-0.2	0.08	0.14	0.08	0.14	0.13	0.14	-0.1	9	-0.1	-0.2	0.11	0.18	0.1	0.17	0.17	0.17	-0.1	9
	R	-0.1	-0.2	0.12	0.16	0.13	0.17	0.17	0.1	-0.1	9	-0.1	-0.2	0.15	0.19	0.15	0.19	0.19	0.09	-0.1	9
Cerebral peduncle L	L	-0.2		0.1		0.15	0.15	0.09	-0.1	0.08	7	-0.1		0.09		0.1	0.14	0.08			5
	R	-0.1		0.11		0.14	0.15	0.1	-0.1	0.05	7	-0.1		0.11		0.11	0.14	0.11			5
Inferior cerebellar peduncle	L	-0.1	-0.1	0.1		0.15	0.14	0.12	-0.1	0.09	8	-0.1	-0.1	0.08		0.12	0.14	0.11	-0.1	0.08	8
	R	-0.1		0.11		0.16	0.13	0.11	-0.1	0.12	7	-0.1	-0.1	0.1		0.14	0.16	0.13	-0.1	0.1	8
Middle cerebellar peduncle		-0.1	-0.1			0.08	0.08	0.06	-0.1	0.07	7	-0.1	-0.1		0.08	0.07	0.1	0.1			6
Superior cerebellar peduncle	L			0.12	0.13	0.09	0.1	0.14			5	-0.1		0.15	0.13	0.12	0.12	0.16	-0.1		7
	R			0.12	0.13	0.09	0.09	0.13			5	-0.1		0.15	0.13	0.12	0.12	0.15			6
Cingulum cingulate gyrus	L	-0.1	-0.2			0.16	0.14	0.13	-0.1	0.07	7	-0.2	-0.2			0.16	0.14	0.11	-0.1	0.1	7
	R	-0.1	-0.2			0.17	0.14	0.16	-0.1	0.06	7	-0.2	-0.2	-0.1		0.17	0.14	0.12	-0.1	0.13	8
Cingulum hippocampus	L		-0.1	0.08		0.08	0.09	0.09			5		-0.1	0.09		0.1	0.1	0.1			5
	R		-0.1		0.08	0.07	0.1	0.11			5		-0.1	0.06	0.06	0.07	0.1	0.1			6
Posterior thalamic radiation	L	-0.2	-0.2	0.11	0.12	0.19	0.19	0.19			7	-0.2	-0.1	0.12	0.07	0.18	0.17	0.16	-0.1	0.07	9
	R	-0.2	-0.2	0.1	0.11	0.19	0.19	0.18	-0.1		8	-0.2	-0.1	0.09		0.19	0.16	0.14	-0.1	0.09	8
Sagittal stratum	L	-0.1	-0.2		0.12	0.12	0.15	0.16			6	-0.1	-0.1	0.05	0.09	0.12	0.15	0.15			7
	R	-0.1	-0.2	0.06	0.12	0.13	0.16	0.16			7	-0.1	-0.1	0.07	0.1	0.12	0.14	0.14			7
Corticospinal tract	L	-0.1	-0.1								2	-0.1				0.05	0.06				3
	R	-0.1	-0.1			0.05	0.06		-0.1	0.07	6										
Fornix		-0.2	-0.1	0.22	0.22	0.21	0.22	0.22	-0.1	0.12	9	-0.3	-0.1	0.27	0.27	0.27	0.27	0.27	-0.1	0.21	9
Fornix cres+Stria terminalis L	L	-0.2	-0.1	0.16	0.08	0.21	0.23	0.21	-0.1	0.1	9	-0.2	-0.1	0.17	0.07	0.22	0.23	0.19	-0.1	0.13	9
	R	-0.2	-0.1	0.18	0.14	0.22	0.24	0.23	-0.1	0.06	9	-0.2	-0.1	0.22	0.14	0.25	0.25	0.24	-0.1	0.12	9
Body of corpus callosum		-0.2	-0.1	0.2	0.15	0.17	0.19	0.21	-0.2	0.06	9	-0.2		0.23	0.15	0.19	0.18	0.21	-0.2	0.12	8
Splenium of corpus callosum			-0.1	0.09	0.16		0.15	0.17	0.11	-0.1	7		-0.1	0.09	0.19		0.14	0.16	0.14	-0.2	7
Genu of corpus callosum		-0.2	-0.1	0.15	0.15	0.16	0.19	0.2			7	-0.2	-0.1	0.21	0.18	0.18	0.22	0.22	-0.1		8
Medial lemniscus	L	-0.1	-0.1			0.08	0.08	0.05		0.06	6	-0.1	-0.1			0.1	0.09	0.06	-0.1	0.1	7
	R	-0.1	-0.1			0.09	0.08	0.05	-0.1	0.07	7	-0.1	-0.1			0.1	0.09	0.06	-0.1	0.1	7
Pontine crossing tract		-0.1	-0.1								2	-0.1	-0.1				0.06			0.08	4
Grand Total		42	43	28	35	43	46	45	30	23	335	39	41	36	35	41	47	45	30	24	338

Figure 8. Chart of each IDP passing the Bonferroni threshold in Figure 6 and the corresponding correlation coefficients. The white matter tract associated with each listed IDP is displayed, based on the JHU white-matter tractography atlas49 on top of the FMRIB58-FA standard-space FA template. Acronyms: fractional anisotropy (FA), mean diffusivity (MD) intra-cellular volume fraction (ICVF); isotropic volume fraction (ISOVF), tract complexity/fanning (OD); L1, L2, and L3 are the three eigenvectors of the diffusion tensor.

- Overall, one very consistent finding was that intracellular volume fraction showed some of the strongest correlations with delta, and the association was always negative. ICVF is related to neuronal density, which means that a higher predicted age than true age correlated to decreased neuronal density.
- IDPs relating to T1 grey matter and white matter volumes correlated negatively with delta, while CSF and white matter hyperintensity (WMH) volume correlated positively. This is expected, since CSF volume in the brain will either increase with age, disease or loss of other brain tissue, and there is a well known association between WMH and ageing.

Correlations with Biobank Variables, a closer look:

In the female group, the deltas correlated weakly with:

- 2 Lifestyle Factors: Alcohol
- 25 Physical (General) measurements
- 60 Bone Density and Size measurements
- 5 Physical (Cardiac) measurements
- 3 Medical History variables

In the male group, the deltas correlated weakly with:

- 2 Lifestyle Factors: Food & Drink
- 2 Lifestyle Factors: Alcohol
- 3 Lifestyle Factors: Tobacco
- 31 Physical (General) measurements
- 5 Bone Density and Size measurements
- 1 Eye Test measurement
- 4 Blood assay measurements
- 7 Medical History variables

- These findings support the idea that this model is producing a biologically relevant measure when it predicts age from an image. It is possible that by involving certain important IDPs as additional inputs in future models, the biomedical relevance of the delta measurement could be greatly improved.

Acknowledgements

1. Miller, K., F, A. & Nature, B.-N. Multimodal population brain imaging in the UK biobank prospective epidemiological study. *Nature* (2016). DOI 10.1038/nn.4393.

Contact: emma.bluemke@dtc.ox.ac.uk