Brain Age Estimation at Tract Group Level and its Association with Daily Life Measures, Cardiac Risk Factors and Genetic Variants

Ahmed Salih (ahmedmahdeeabdo.salih@univr.it)
University of Verona
Ilaria Boscolo Galazzo
University of Verona
Zahra Raisi-Estabragh
William Harvey Research Institute, Queen Mary University of London
Elisa Rauseo
William Harvey Research Institute, Queen Mary University of London
Polyxeni Gkontra
University of Barcelona
Steffen E. Petersen
William Harvey Research Institute, Queen Mary University of London
Karim Lekadir
University of Barcelona
André Altmann
University College London
Petia Radeva
University of Barcelona
Gloria Menegaz
University of Verona

Research Article

Keywords: Brain age, MRI , Tract Group Level , Daily Life Measures, Cardiac Risk Factors ,Genetic Variants

Posted Date: May 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-518759/v1

Brain age estimation at tract group level and its association with daily life measures, cardiac risk factors and genetic variants

Ahmed Salih[1,*], Ilaria Boscolo Galazzo[1], Zahra Raisi-Estabragh[2,3], Elisa Rauseo[2,3], Polyxeni Gkontra[4], Steffen E. Petersen[2,3], Karim Lekadir[4], André Altmann[5], Petia Radeva[4], and Gloria Menegaz[1]

[1]University of Verona, Department of Computer Science, Verona, 37134, Italy
[2]William Harvey Research Institute, NIHR Barts Biomedical Research Centre, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, UK
[3]Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, West Smithfield, London, EC1A 7BE, UK
[4]University of Barcelona, Dept. de Matemàtiques i Informàtica, Barcelona, Spain
[5]Centre for Medical Image Computing (CMIC), Department of Medical Physics and Biomedical Engineering, University College London, London, UK
[*]ahmedmahdeeabdo.salih@univr.it

ABSTRACT

Brain age can be estimated using different MRI modalities including diffusion MRI. Recent studies demonstrated that white matter (WM) tracts that share the same function might experience similar alterations. Therefore, in this work, we sought to investigate such issue focusing on five WM bundles holding that feature that is Association, Brainstem, Commissural, Limbic and Projection fibers, respectively. For each tract group, we estimated brain age for 15335 healthy participants from United Kingdom Biobank relying on diffusion MRI image derived endophenotypes, Bayesian ridge regression modeling and 10 fold-cross validation. Furthermore, we estimated brain age for an ensemble model that gathers all the considered white matter bundles. Association analysis was subsequently performed between the estimated brain age delta as resulting from the six models, that is for each tract group as well as for the ensemble model, and 38 daily life style measures, 14 cardiac risk factors and cardiovascular magnetic resonance imaging features and genetic variants. Our study revealed that the limbic tracts experience less brain aging compared to other tract groups, Brainstem tracts ages relatively faster while the other tract groups present similar brain aging patterns. The model achieved 5.86 mean absolute error (MAE) for brainstem, 5.08 for limbic tracts and around 5.2 for the other three tract groups while, the ensemble model 4.55. Moreover, the results suggest that limbic tracts are more affected by daily lifestyle factors than the other tract groups. Lastly, two SNPs were significantly (p-value < 5E-8) associated with brain age delta in projection fibers. Those SNPs are mapped to *HIST1H1A* and *SLC17A3* genes.

1 Introduction

Aging is a complex process with substantial impact across multiple organ systems, yet to be fully characterised. In the specific case of the brain, previous studies have found evidence of considerable structural alterations of white and grey matter (WM/GM) structures as well as of morphological and functional connectivity changes across different areas[1]. These modifications are associated with distinct aspects of cognitive functions, emotions, and neurodegenerative disorders[2]. Several neuroimaging modalities can be adopted to estimate the so-called brain age which allows monitoring the longitudinal progression of brain during lifecourse. This is defined as the apparent biological age of the brain, when comparing individuals' data against a population dataset spanning a range of ages[3,4]. The difference between predicted brain age and actual (chronological) age, generally referred to as "predicted age delta" (brain-PAD), is often computed to verify whether a subject's brain appears younger or older than their chronological age. Indeed, since humans do not experience brain aging at the same rate and pronounced differences possibly related to genetic and environmental factors are present, brain-PAD can be exploited as a novel biomarker to assess brain aging progression in both healthy and diseased populations. Greater brain age (positive brain-PAD) has been associated with increased risk of neurodegenerative diseases, whilst younger brain age (negative or small brain-PAD) correlates with healthy environmental exposures and lifestyle habits[5]. Among these factors, daily lifestyle, physical activity, electronic device use, and sleeping habits have all shown significant effects on brain progress during the lifecourse[4,6], with smoking and greater alcohol intake frequency closely linked to increased brain-PAD for instance. Similarly, genetic factors also have

a crucial role in brain aging. In a recent study, Jonsson and colleagues[3] demonstrated the presence of two single nucleotide polymorphisms (SNPs) significantly associated with brain-PAD by relying on a genome wide association study (GWAS) on healthy subjects from the United Kingdom Biobank (UKB) database. These variants were correlated with reduced white matter surface area and reduced sulcal width[3]. Other studies identified several SNPs associated with brain-PAD, with the most significant ones located in *MAPT*[6] and *TMEM106B* genes[7]. These two genes have been shown to be closely associated with frontotemporal dementia[8], and *MAPT* has also been considered as a model of interaction in Parkinson's disease between functional disease outcomes and genetic[9]. Recently, 156 loci resulted as significantly associated with brain-PAD values estimated using several brain modes assembling more than 3900 imaging features, as reported in[10]. These peaks involved different SNPs that were correlated with several factors, such as blood pressure, migration of neurons and Alzheimer's disease (AD) endophenotypes[10].

Furthermore, there is a growing evidence suggesting complex cross-system interactions between brain and cardiovascular systems[11,12]. Cardiovascular functions are controlled by large regions in cortical and subcortical structures through the process of sympathetic and parasympathetic outflow. The brain and heart are therefore closely inter-connected and it is highly probable that changes in one organ system might lead to alterations in the other[13]. Indeed, cardiovascular risk factors (CRFs) have been already associated with poorer cognitive function. Precisely, higher body mass index (BMI) has been linked to poorer performance across multiple cognitive indications including working memory, attention, delayed recall, and category fluency[14]. In addition, other risk factors such as diabetes and hypertension have been associated with unhealthy brain aging, abnormal neuroanatomical alterations, and increasing risk of developing AD[15]. Ann-Marie et, al.[16] investigated brain aging in the Whitehall II (WHII) MRI cohort using measures derived from grey white matter microstructure, grey matter morphology and resting state functional connectivity. Then, they performed association of brain aging and CRFs. Their results indicate that CRFs, such as stroke risk score and alcohol intake are associated with older brain age. Moreover, blood pressure was more associated with white matter than grey matter. All these elements deserve further investigations to better understand whether they might influence the brain aging processes differently.

In this context, neuroimaging data derived from Magnetic Resonance Imaging (MRI) sequences have demonstrated to provide accurate estimates of the apparent age of individuals' brains, generally relying on age regression models[17]. Most brain-age models only use T1-weighted structural MRI, reflecting brain volumes. However, the possibility to use complementary modalities mapping different aspects of brain structure and function, has opened the way to the estimation of modality-specific brain ages. In particular, diffusion MRI (dMRI), resting-state/task functional MRI (fMRI) and susceptibility weighted imaging are currently exploited in different studies to extract novel image-derived phenotypes (IDPs) to be used in specific brain-age models, thanks to the new opportunities offered by large-scale multimodal databases such as the UKB[18]. While these separate ages should be closely related in healthy subjects, even though differences related to regional age-related changes might exist, for patient groups distinct patterns of aberrant brain aging could emerge[5]. Statistical methods for modeling brain age using neuroimaging data are generally highly accurate, with MAE of predictions in the range of 4-5 years for most of the studies relying on different regression approaches such as simple linear regression, support vector regression (SVR) and least absolute shrinkage and selection operator (LASSO) models age[4–6,17]. Complex models grounding on deep learning approaches have been also applied mainly to raw T1-weighted scans, achieving lower MAE values (< 3 years)[3,19].

In addition, most of these previous studies have demonstrated better results when including multimodal neuroimaging data rather than a single modality in the models[16,20,21]. In particular, findings from these multimodality studies suggest that dMRI measures have higher accuracy in predicting brain age compared to those derived from fMRI, SWI or even anatomical images in some cases[5]. Indeed, Niu and colleagues[1] found that the model built on dMRI and GM volume measures could reach the highest accuracy, and that replacing diffusion with functional data worsened the performance. In addition, Cole[5] demonstrated that structural and diffusion data led to the best performances when single modalities were used to predict brain age, and that, when considering all the data together and performing a leave-one-out prediction, the biggest decrease in performance occurred when excluding dMRI data. These diffusion-based features are generally extracted starting from the microstructural maps estimated using different models, such as the diffusion tensor imaging (DTI) and the neurite orientation dispersion and density (NODDI), and then averaging the corresponding values over several WM tracts. Fractional anisotropy (FA) along with indices of diffusivity (mean/axial/radial [MD/AD/RD]) can be estimated from the DTI model, informing on the degree of anisotropy/diffusivity of a diffusion process[22]. Conversely, more complex indices are derived from NODDI, a compartmental model where brain microstructure is described in terms of a set of predefined parameters that is neurite orientation dispersion (OD), representing the overall coherence of modeled axons, isotropic volume fraction (ISOVF), showing the unhindered water volume fraction, and intracellular volume fraction (ICVF) that represents neuronal density[23,24]. Previous works have demonstrated the importance of DTI and NODDI IDPs for estimating brain age in both healthy and diseased populations[25,26].

Several studies have demonstrated that groups of WM tracts that share the same function experience similar alterations during the life course and in specific brain disorders. In particular, Yang et al.[27] investigated the association of brain aging with WM integrity and functional connectivity in a group of healthy subjects. Their findings suggest that some WM tracts are

preferentially affected during lifecourse. For example, projection, association and commissural fibers were substantially affected by aging resulting in a significant reduction of their WM integrity, while Brainstem tracts were relatively preserved. In another study, Bender et al. compared FA, AD, and RD indices estimated over Association, Commissural and Projection fibers again in a healthy population[28], demonstrating a greater microstructural decline over time in the first fiber group compared to the Commissural and Projection ones, and a differential aging of cerebral WM. Another analysis compared DTI-based measures in Limbic, Commissural, Association and Projection fiber groups in patients with AD, mild cognitive impairment and with normal cognition (CN)[29]. Their results revealed distinctive structural changes among the fibers, and different values between AD and CN among the four fiber groups, such as FA, MD and RD group differences in Limbic and Commissural fibers, or diffusivity changes in Association fibers. Therefore, existing work suggests differential aging-related changes depending on the specific WM fiber groups, which might result in diverse patterns of disease and cognitive impairment. However, the determinants of these different alteration patterns have not been adequately investigated so far.

In this study, we aimed at estimating and comparing diffusion-specific brain ages in a large cohort free from clinically diagnosed neurological disease from the UKB database, relying on dMRI measures of different fiber groups in order to assess the impact of aging on WM at the tract-group level. Indeed, investigating brain aging for tracts with shared functionality may permit a more accurate and disease-specific risk assessment compared to brain aging for the whole brain. In addition, for each fiber group, we evaluated the relationship between brain predicted ages and several factors spanning across different scales, relating in particular to daily lifestyle, health, cardiac measures and genetics to verify whether a differential association might be present in specific WM tracts. This will also allow to identify those factors that can negatively impact brain aging, providing further insights on its complex mechanisms.

2 Materials and Methods

2.1 Datasets

Participants. Data from n = 16394 participants with complete brain and cardiac MRI assessment were initially downloaded from the UKB database. Of these, 1059 subjects who reported neurological disorders that could directly affect cognitive function were excluded in order to include only people who met criteria for being neurologically intact at the time of scanning. These were identified using the self-reported medical conditions at baseline extracted from detailed questionnaires that the UKB participants had to answer, the relevant ICD-10 code, hospital episode statistics, and algorithmically-defined outcomes. This led to a final group of 15335 subjects (mean age 54.79 $\pm$ 7.45, 7277 males, 8058 females). The complete list of conditions and ICD-10 codes used as inclusion/exclusion criteria are available in supplementary table 2.

All the methods were conducted in accordance with the relevant guidelines and regulations and all participants provided informed consent. UKB received ethical approval from the NHS National Research Ethics Service on 17th June 2011 (Ref 11/NW/0382) and extended on 10th May 2016 (Ref 16/NW/0274). More details can be found on the UKB resource page https://biobank.ndph.ox.ac.uk/showcase/catalogs.cgi. The present analyses were conducted under data application number 2964.

Brain and Cardiac MRI features. The UKB brain imaging protocol was implemented on a 3T Siemens scanner (Skyra, VD13A SP4, Siemens Healthcare, Erlangen, Germany) and included six different sequences, covering structural, diffusion and functional imaging for a total of 35 minutes scan time. In particular, a multi-shell protocol has been used for dMRI data, with two b-values (b = 1000, 2000 s/mm^2), a 2-mm isotropic resolution and a multiband acceleration factor of 3. 50 diffusion-encoding directions were acquired per shell, covering a total of 100 distinct directions over the two b-values. Full details on the neuroimaging data can be found at https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf. Cardiac MRI was performed on a 1.5T Siemens scanner (MAGNETOM Aera, Syngo Platform VD13A, Siemens Healthcare) according to a pre-defined protocol[30,31]. Left and right ventricular (LV, RV) function was assessed using standard long and short axis acquisitions.

Genotype data. UKB genotyped genetic data for 488377 participants were obtained using two genotyping arrays. A small subsets of the participants (49950) involved in UKB Lung Exome Variant Evaluation (UK BiLEVE) study were genotyped using the Applied Biosystems UK BiLEVE Axiom Array by Affymetrix. Conversely, the majority of the participants (438427) was genotyped using the closely related Applied Biosystems UKB Axiom Array. More details about genotyping and genotype calling steps can be found in[32].

2.2 Feature extraction

Brain microstructure feature extraction. In the current study, we relied on the IDPs derived centrally by the researchers involved in the UKB project and made available via the data showcase (https://biobank.ctsu.ox.ac.uk/crystal/index.cgi). Of these, we focused on the 675 dMRI IDPs extracted for each participant using the following pipeline. First, both the diffusion tensor and the NODDI models were fitted to the pre-processed data leading to nine voxelwise microstructural maps, namely FA, MD, axial diffusivity (L1), radial diffusivities (L2, L3) and mode of anisotropy (MO) from DTI, and ICVF,

ISOVF, and OD from NODDI. Two sets of measures were used as microstructural features, both obtained from the UKB repository and extracted following two different approaches[18,33]. The first used tract-based spatial statistics (TBSS). Each individual dMRI map was aligned to a standard-space WM tract skeleton and a series of ROIs was then defined as the overlap of this skeleton with 48 standard-space tract masks from the JHU ICBM-DTI-81 atlas[34]. For each skeletonised microstructural index, the mean value was calculated in each region, leading to a total of 432 IDPs (that is 48 ROIs times 9 IDPs). The second relied on probabilistic tractography. A total of 27 major tracts were identified using standard-space start/stop ROI masks. The mean value of each DTI/NODDI parameter was calculated across each tract and weighted by the tractography output as in Alfaro at, al.[33] in order to emphasize values in regions most likely to belong to the tract of interest, resulting in a total of 243 IDPs (27 tracts times 9 IDPs). Each ROI and tract were subsequently assigned to one out of five fiber groups (FG) following the fiber tract-based atlas[35]. In particular, the following FG were considered: i) Association; (cortex-cortex connections); ii) Brainstem; iii) Commissural (left-right hemispheric connection); iv) Limbic; and v) Projection (cortex-brainstem, cortex–spinal cord and cortex-thalamus connections) fibers. Each FG consisted of a different number of tracts, that is: 22 for Association, 13 for Brainstem, 13 for Commissural, 9 for Limbic, and 18 for Projection. An illustration of these five fiber families is reported in figure 1, where the different tracts are depicted in different colors. *Association* fibers interconnect different cortical

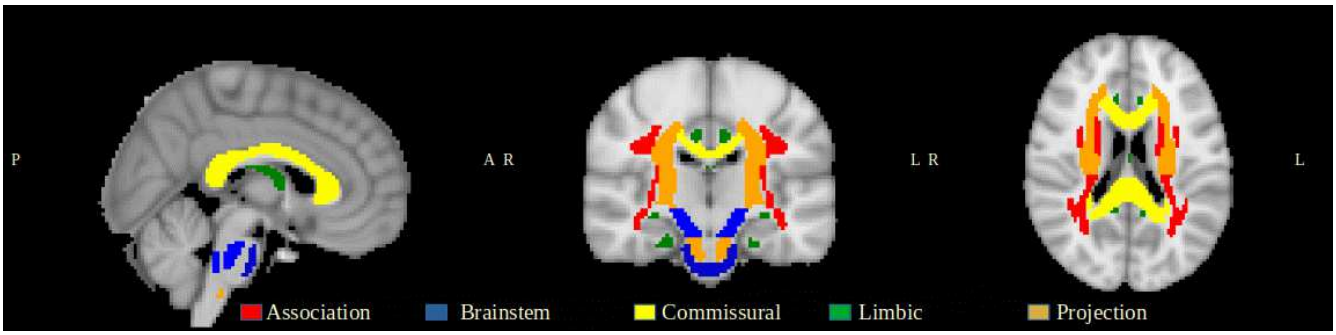


Figure 1. White matter tract groups

areas in the same hemisphere[36]. These might be short association fibers that connect adjacent gyri, or long association fibers linking more distant parts of the cerebral cortex. Important examples of this category are the superior/inferior longitudinal fasciculus, inferior fronto-occipital fasciculus, and uncinate fasciculus[37]. *Brainstem* fibers involve the tracts that connect cerebrum to the spinal cord and cerebellum[38]. These includes the corticospinal tract, the posterior column-medial lemniscus pathway and the spinothalamic tract. *Commissural* fibers interconnect corresponding cortical regions of the two hemispheres and are mainly represented by the corpus callosum and anterior commissure[36]. *Limbic* fibers involve structure in both sides of thalamus[39]. Fornix is one of the main vital tract of this system[40], alongside the Cingulum bundle that connects parietal, frontal and temporal lobe[41]. Finally, *Projection* fibers connect cortical areas with deep nuclei, cerebellum, brainstem, and spinal cord[42]. Corticospinal and corona radiata tracts are the two main examples for this category[43]. By subsequently assigning each ROI and tract to the respective FG, summary IDP values could be finally derived by averaging across ROIs and tracts, respectively, in each FG. In this way, the whole set of IDPs were assigned to each FG resulting in a total of 18 IDPs (9 from the ROI-based and 9 from the tract-based analyses).

Cardiovascular feature extraction. Cardiovascular Magnetic Resonance (CMR) data were analysed using an automated pipeline[44]. The extracted cardiovascular indices included measures of LV and RV structure and function. Specifically, the indices derived for the LV were end-diastolic volume (LVEDV), end-systolic volume (LVESV), stroke volume (LVSV) and mass (LVM). The RV indices included stroke volume (RVSV), end-diastolic volume (RVEDV) and end-systolic volume (RVESV) were considered. LV and RV volumes are markers of cardiac remodelling, from these stroke volume may be derived as a measure of ventricular function. LVM is an independent risk predictor in clinical cohorts and an indicator of heart aging in population cohorts. To correct for variation in CMR metrics related to body size, these measures were indexed to body surface area (calculated as per Du Bois formula)[45]. As an additional measure of arterial health in a larger sample, we considered arterial stiffness index (ASI) derived from finger plethysmography[46]. ASI was measured at the baseline UKB visit using the PulseTrace PCA2 (CareFusion, USA) device according to a pre-defined protocol, UK Biobank Arterial Pulse-Wave Velocity (2011) that is available at `https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/Pulsewave.pdf`. Outliers were removed from the ASI variable using a 1·5× interquartile range (IQR) rule. Finally, CRFs included hypertension, diabetes, deprivation (reported in UKB as the Townsend index), body surface area (BSA), BMI and exercise level.

Lifestyle features. Regarding daily life measures, 38 variables were available in the UKB database at baseline. The lifestyle and environment measures included seven categories that are: physical activity (7 measures), sun exposure (2 measures), electronic devise use (2 measures), smoking (2 measures), sleeping habits (5 measures), alcohol (3 measures) and diet (17

measures). All the used variables are available in supplementary table 3.

2.3 Brain Age estimation.

All the analyses performed in our study were carried out using Python 3.8.5 and Scikit-learn version 0.23.2. A tract-based healthy aging model was defined for each of the five FG, using the corresponding 18 dMRI IDPs as neuroimaging predictors and the chronological age as dependent variable. To account for the different measurement scales, the features were normalized to zero mean and unit variance[4]. Sex, education level, height and volumetric scaling from T1-weighted head image to standard space were used as covariates because they could be statistically associated with the outcome variable, as previously reported in similar studies[4,5,47]. A Bayesian ridge regression model was run in combination with a 10-fold cross-validation, where the data samples were randomly assigned into ten equal-sized groups. For each group of left out data, the other 90% of subjects were used to estimate the model parameters which were then applied to this additional group for validation. The performance of each model was assessed using MAE and Coefficient of Determination (R^2).
Several studies have revealed a proportional bias in brain age estimation related to regression model dilution, leading to a significant age-dependency between delta age and chronological age[48,49] that needs to be statistically corrected. In this study, we adopted the method proposed by Beheshti et al.[48] which entailed calculating the regression line between brain-PAD and chronological age in the training set:

$$D = \alpha * \Omega + \beta \tag{1}$$

where D is the brain-PAD from training data, α and β represent the slope and the intercept of the linear regression model, and Ω is the corresponding chronological age. Then, these intercept and slope values were used to correct the predicted brain age in the validation set set as follows:

$$CPBA = Predicted\,Brain\,Age - (\alpha * \Omega + \beta) \tag{2}$$

where CPBA represents the corrected predicted brain age (bias-free). After bias correction, the brain-PAD was calculated as the actual age subtracted from the brain-predicted age. Pearson correlation was calculated between actual and predicted brain age as well as actual age and brain-PAD, both before and after the bias correction steps. An ensemble tract-based aging model was finally defined using the overall 90 dMRI IDPs (18 from each FG), and the same analyses detailed above were implemented.

2.4 Association analysis

In order to highlight the role of the different variables to model brain age, the association between delta values as resulting from the five FG models and a set of imaging/non-imaging variables was assessed using linear regression model. This included the corresponding 18 dMRI IDPs, 14 CRFs/CMR measures, and 38 daily life measures. In addition, the same analyses were performed for the brain-PAD values derived from the ensemble model, with the only difference being represented by the associations with the whole set of 90 IDPs for the dMRI part (rather than 18 only). In all models, the brain-PAD represented the outcome measure, while the feature of interest was the independent variable alongside all the above mentioned covariates plus age[4,5]. The resulting p-values were Bonferroni-corrected for multiple comparisons[50]. Of note, Cook's distance was used to identify potential influential observations before performing the association analyses. In particular, a subject was removed if the Cook's distance was greater than 3 times the mean distance of all the subjects[6]. The association between genetic variants and brain-PAD values as resulting from each model was also conducted. The quality control steps on SNPs included Minor allele frequency (MAF) thresholding at 0.01, missing rate less than 0.02 and Hardy-Weinberg equilibrium p-value >=1E-6. Quality control on samples ensured that all participants had genotyping rate > 0.98, heterozygosity rate within ±3 standard deviation, matched genetic/reported gender and were of European ancestry (according to both genetic ethnicity based on principal component analyses and self-reported ethnicity). Related samples were removed based on kinship coefficient > 0.1. The quality control steps resulted in 574492 autosomal SNPs and 12364 subjects for the GWAS analyses. Thereafter, linear regression was performed using PLINK[51] and adjusted for education, gender, age, volumetric scaling from T1-weighted head image to standard space, and 40 genetic principal components of ancestry. For each GWAS analysis, FUMA[52] was used to map the significant SNPs to genes based on positional mapping and eQTL. Using FUMA and GTEx (https://gtexportal.org/home/), we also identified Expression quantitative trait loci (eQTL) to take advantage of gene expression. Finally, we looked at UKB genetic data (http://big.stats.ox.ac.uk/)[53] to find association between the significant SNPs and other phenotypes.

3 Results

3.1 Brain age estimation.

The impact of aging was separately assessed in terms of MAE and (R^2) values after fitting the five considered multivariate linear FG-based models plus the ensemble one. The mean and standard deviation of such values across a ten-fold cross-validation

were reported in order to probe the reliability of the estimation. Results are summarized in Table 1 where the columns 2 to 6 correspond to the five FG, that is, Association, Brainstem, Commissural, Limbic and Projection fibers, and the last column reports the results for the FG ensemble. In the table, the Pearson correlation coefficient between the actual age and the predicted age before (CAPB) and after (CAPA) correction, the actual age and the brain-PAD before (CADB) and after (CADA) correction are also reported in the last four rows.
As it can be observed, the performance is quite uniform across FG, with the exception of the Brainstem group especially regarding the R^2 value that is the lowest. The best MAE was obtained for the tract ensemble model followed by the Limbic FG, which also corresponds to the highest R^2. The last four rows prove that the age-bias was successfully removed.

Table 1. Performance of the five FG-based models plus the ensemble one to estimate brain age in terms of MAE and R^2. The last four rows provide the CAPB, CAPA, CADB and CADA, respectively. The best performing model is highlighted in green.

Matrices	Association	Brainstem	Commissural	Limbic	Projection	Ensemble
Mean R^2	0.26	0.11	0.26	0.29	0.25	**0.42**
STDV R^2	0.02	0.01	0.01	0.02	0.03	0.015
Mean MAE	5.24	5.86	5.23	5.08	5.28	**4.55**
STDV MAE	0.1	0.09	0.11	0.12	0.13	0.08
CAPB	0.51	0.33	0.51	0.54	0.5	0.65
CAPA	0.91	0.95	0.91	0.9	0.91	0.89
CADB	-0.85	-0.94	-0.86	-0.83	-0.86	-0.75
CADA	-0.001	-0.003	-0.001	-0.001	-0.001	-0.006

3.2 Association studies

IDPs association with brain-PAD. Linear regression results describing the relationships between the bias-adjusted brain-PAD values for the five FG-based models and each microstructural IDPs are illustrated in Figure 2. In addition, results for the ensemble model are also reported, including in this case the associations with 90 dMRI IDPs rather than 18 as in the case of the previous five FG models. A highly similar association pattern is apparent across FG, though higher variability was observed for the Limbic FG. More specifically, all the IDPs were significantly associated with brain-PAD in Association and Commissural groups, while few associations did not reach significance in the other three tract groups, that is: mean L1 and mean ICVF in Brainstem, weighted mean L1 in Limbic and weighted mean MO in Projection fibers. Considering the different imaging variables, anisotropy (FA, ICVF, OD and MO and respective weighted versions) and diffusivity (MD, L1, L2, L3, ISOVF and weighted versions) indices led to associations of opposite direction, as expected, in almost all the cases. More precisely, FA and weighted FA showed a significant negative association with brain-PAD in all five groups, while MD/weighted MD presented the opposite pattern and appeared to more strongly contribute to modelling the outcome in all cases. Similarly, increased L1, L2 and L3 plus their weighted versions were associated with increased the brain-PAD for all groups, except L1 and weighted L1 in Brainstem and Limbic FGs, respectively. Finally, the weaker associations were observed for MO and weighted MO in all models. Consistently with what above, for NODDI-based measures, the diffusivity index ISOVF was positively associated with the brain-PAD in all cases. A slightly different pattern was observed for OD and its weighted version across the FG, that presents a higher variability. OD is positively associated with the brain-PAD, as expected, in Brainstem and Limbic fibers, though not in the weighted version, and has a different pattern in the other three groups, with a prevalence of a negative association of the weighted version. The association between the IDPs and the brain-PAD was also assessed FG-wise relying on the Ensemble model, revealing that the pattern was preserved though with slightly different values. In particular, the association was slightly reduced with respect to the values that were obtained for FG-specific brain-PADs. Check table 4 in the supplementary for more details regarding the association.

CRFs and vascular measures association with brain-PAD. Figure 3 reports the results of the linear regression analyses between the bias-adjusted brain-PAD values and the CRFs/CMR measures, revealing consistent patterns across the five FG models. In all conditions, several measures were significantly associated with PAD after multiple comparison correction, in particular increased brain-PAD was associated with a diagnosis of diabetes, hypertension and increased LVM, as well as with reduced LVSV/RVSV and RVEDV/RVESV. Greater BMI was also associated with increased brain-PAD in three out of five models (Projection, Brainstem and Limbic), with the last two fiber groups also showing a positive relationship between delta and BSA. Of note, the model based on Limbic fibers presented the highest number of significant associations and the direction of the relationships was consistent in all the five FC-based models. The same trend was observed for the Ensemble model. These associations followed the same pattern compared to the other five FG models. More precisely, the association results were closer to those found for the the Brainstem and Limbic FG, especially in eight out of 14 measures. Check table 5 in the supplementary for more details regarding the association.

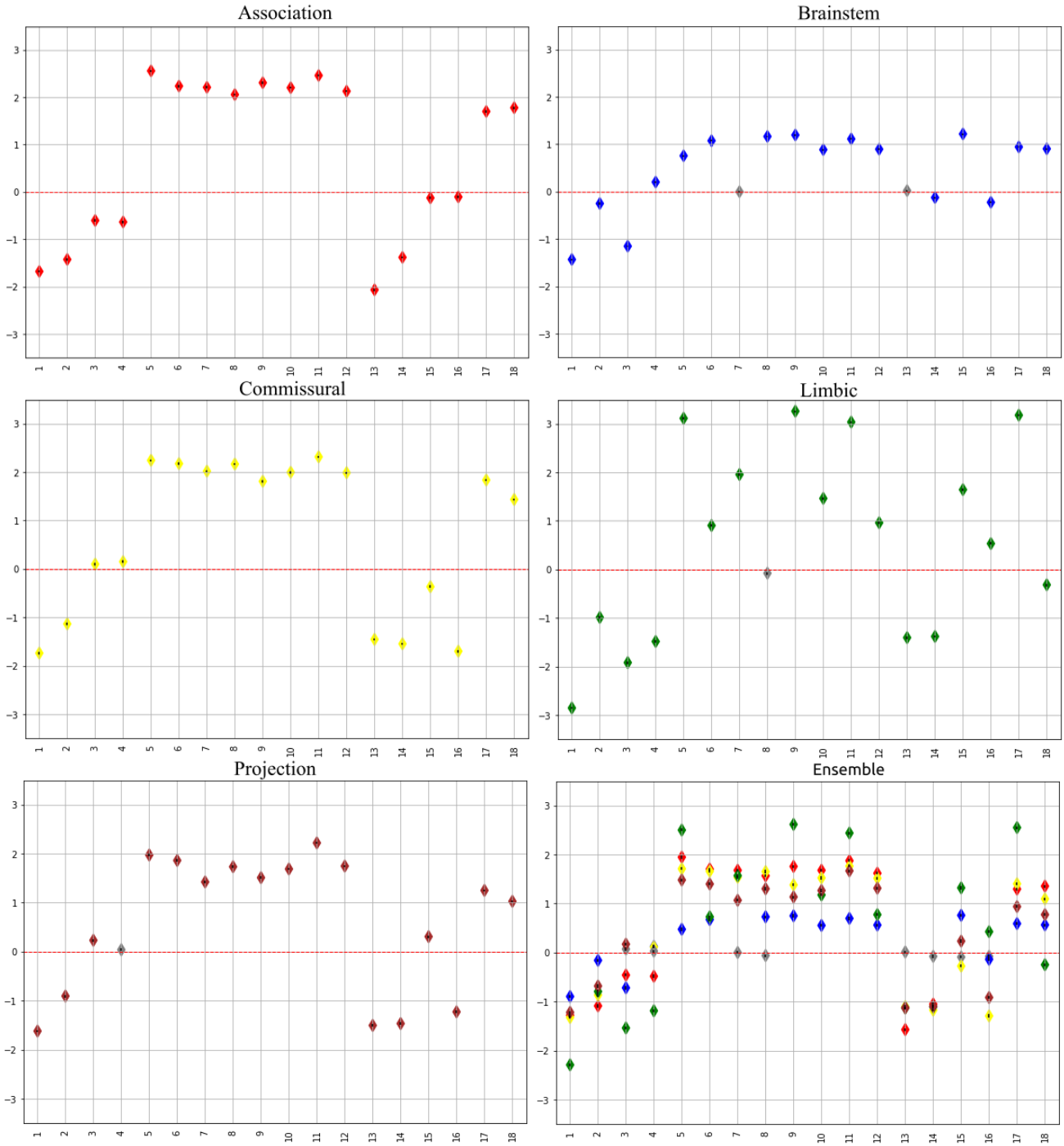


Figure 2. Association of the IDPs and brain-PAD for the different models. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error (the small black dot inside the diamond shape). Grey color indicates non-significant association.

Lifestyle association with brain-PAD. Figure 4 reports the results of the linear regressions between the bias-adjusted the brain-PAD values and the daily life measures in each of the five FG-based models plus the Ensemble one. The majority of

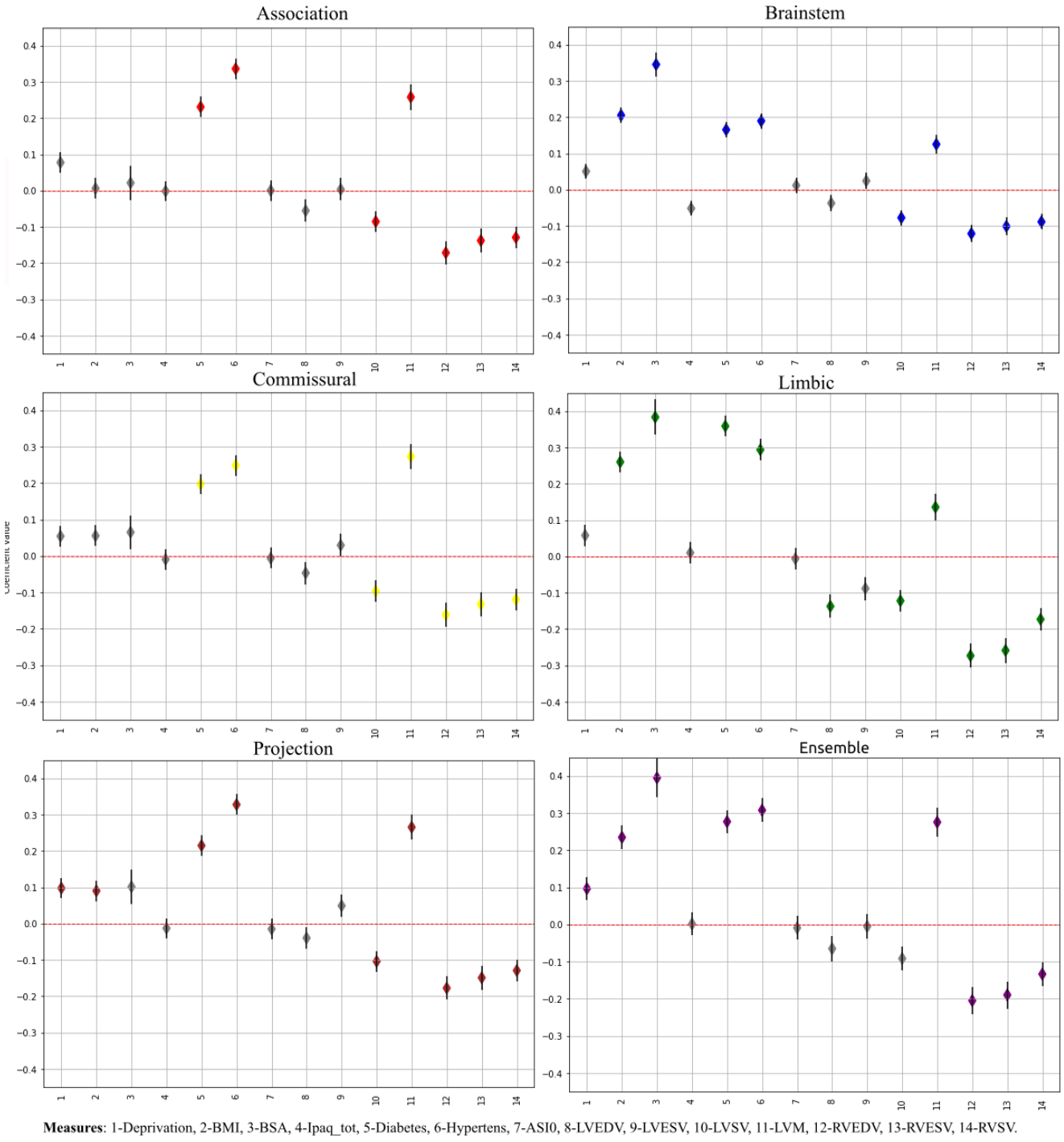


Figure 3. Association of the CMR measures, CRFs and brain-PAD. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error. Grey color indicates non-significant association.

the measures were not significantly associated with brain-PAD after multiple comparisons correction, but consistent patterns were visible across the FG. The highest number of significant associations was observed for the Limbic tracts, while only four measures survived for the Brainstem group, though in agreement with the others. In all cases, increased brain-PAD was associated with ever smoked, smoking status, greater oily fish intake, and tea intake (except for Association fibers). In addition, increased brain-PAD values from Association, Commissural and Limbic fibers were associated with greater lamb/mutton intake

and greater frequency of alcohol intake. Duration of walk for pleasure had a positive impact on brain age, being associated with reduced brain-PAD values in both Limbic and Projection fiber models, while increased brain-PAD was associated with water intake in Commisural and Projection FG models. Finally, seven additional daily life measures, including time spent watching TV or using computer and sleep duration, presented only selective associations in one of the models (4 for Limbic, 2 for Association and 1 for Projection). Check table 6 in the supplementary for more details regarding the association. The coefficient value for all those are significantly associated with brain-PAD in FG is small (less than 0.3) indicating small effect.

Association between SNPs and brain-PAD. Two SNPs located on chromosome 6 showed significant associations ($p < 5E-08$) with brain-PAD values in the Projection FG, namely rs1045537 ($p = 2.87E-08$) and rs16891334 ($p = 4.268E-08$). Figure 5 illustrates the Manhattan plot showing the association between the SNPs in all chromosomes and brain-PAD from the Projection FG. As can be observed in Figure 5, another less conservative significance threshold of $10-5$ was considered. Moreover, the Manhattan plots for the other FG and the ensemble model were also reported (Appendix, Figures 5-1 to 5-5). Despite non of the SNPs passed the GWAS threshold, however many SNPs passed the less conservative significance threshold. Many SNPs located in chromosomes 17 in limbic fibers and the ensemble model passed the less conservative threshold. The leading SNP (rs1045537) was mapped to *BTN3A* cluster *(BTN3A1 to BTN3A3)*, *SCGN, SLC17A* cluster *(SLC17A1 to SLC17A4)*, *HIST1H1A* group of genes based on FUMA results using positional mapping and eQTL based on GTEx database (version 8 54 and 8 30) and general tissue types. In addition, it is significantly associated significantly with forced vital capacity, BMI, headache and coeliac disease in UKB cohort.

Discussion

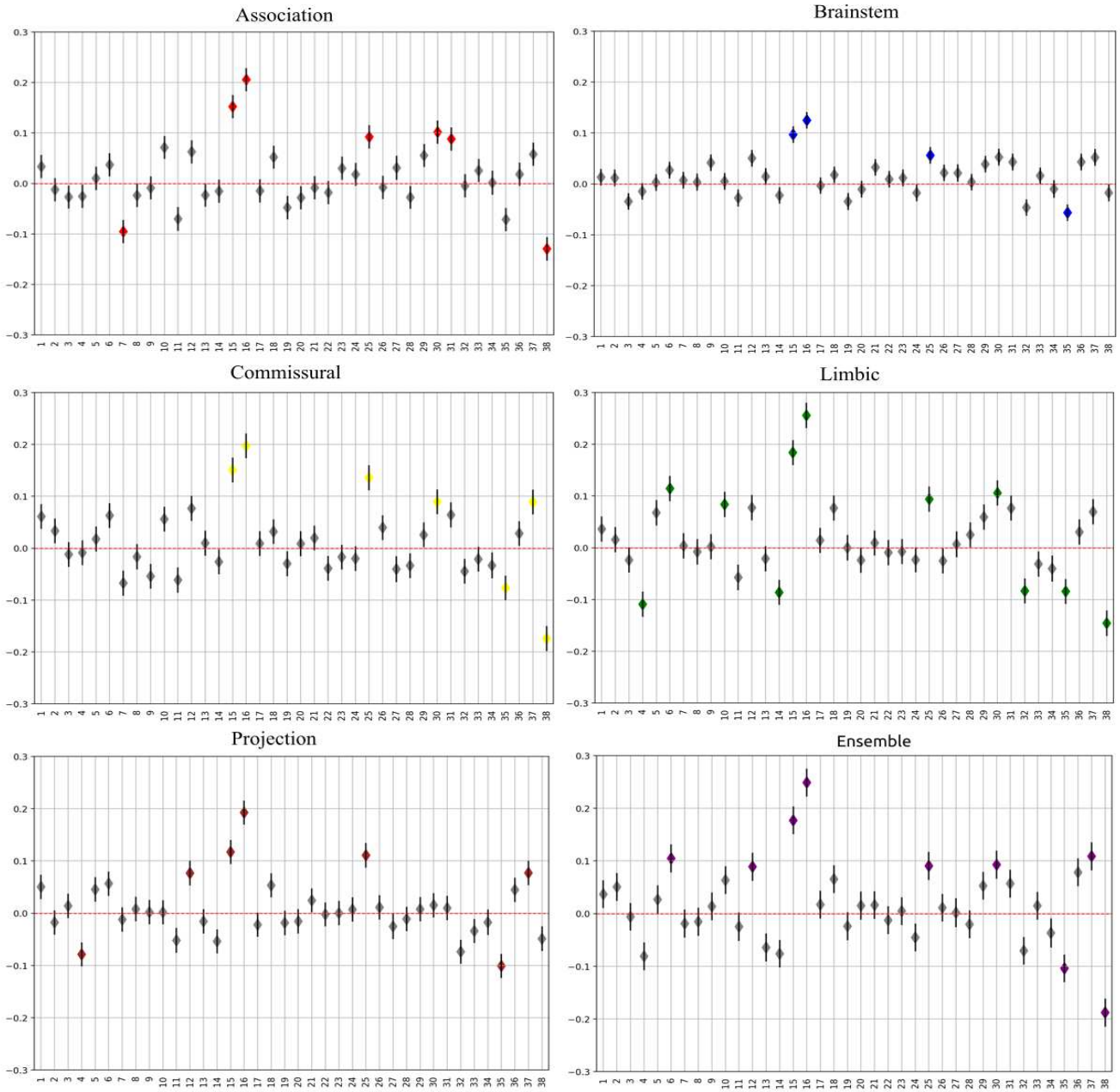
This study focuses on providing a holistic view on the endogenous end exogenous factors shaping brain age as expressed by brain microstructural features of specific WM tracts, providing hints for the multiscale and multifactorial analysis of the system 'human being'. The challenge being to link heterogeneous information living at different scales, this work is a step in that direction by linking microscopic (genes), mesoscopic (dMRI IDPs), macroscopic (cardiovascular IDPs) and behavioral (lifestyle) measures through their respective association to the brain age picture provided by dMRI. After investigating the potential of microstructural measures derived from dMRI in estimating brain-PAD relying on five different FG, the associations of neuroimaging, genetic and cardiovascular IDPs with brain-PAD were assessed and, as a final step, lifestyle and behavioral measures were also considered. The rest of this section will be dedicated to the discussion of the results as well as of the potential consistency of the observed associations across scales while referring to the existing literature.

The estimated brain-PAD was minimized by the Ensemble model, gathering the whole set of 18×5 microstructural features, leading to the minimum MAE (4.55 years) and the maximum R^2 (0.42). On the other end, Brainstem FG led to worst performance, with the highest value for the MAE (5.86 years with std =0.01) and the minimum for R^2 (0.11). The Brainstem fibers, including the midbrain, pons, and medulla, involves structures with complex WM pathways and GM nuclei that are concentrated in a small area. Intricate Brainstem circuitries are difficult to capture using conventional dMRI measures such as DTI, with the consequence that both the tractography and the estimation of microstructural indices are prone to errors[54].

Among the other single FG-based models, the Limbic one provided the best MAE (5.08 with std = 0.02) and R^2 (0.29), closely followed by the others (Association, Commissural and Projection) showing a similar pattern. The Ensemble model combining all available IDPs provided the best results compared to single FG-based models, suggesting that the inclusion of multiple features from different WM FG could depict better the modulations related to brain aging and therefore led to more accurate estimates.

The association of brain age delta with the dMRI IDPs used to estimate brain age revealed the path these measures follow in brain aging. Based on the corrected p-value, their effects were significant in all tract groups apart from very few cases. Fractional anisotropy and ICVF were reduced in all tract groups while, L1, L2, L3, MD, ISOVF are increasing with advancing age. The contribution of OD and MO was relatively inconsistent among tract groups, in some groups they are increasing, while they experience reduction in other tract groups. The association was more similar in some groups than others. For example, the results were more similar in the association, commissural, projection tracts groups and also when all IDPs were used. The results of limbic tract showed different pattern compared to other tract groups. Our results are inline with what have been published before in terms of the direction these IDPs follow in brain aging. Smith, et.al[4] showed similar results regarding the contribution of the dMRI in brain aging. Their results also indicate that FA and ICVF are reducing in brain aging while L1, L2, L3, MD and ISOVF experience increasing. In addition, they showed that the dMRI matrices in Fornix are among of those most significant features to estimate brain age in both male and female. Our results followed similar pattern as the IDPs from limbic tracts which are dominant by Fornix tracts were on the top of the list of significant features to model brain age when all the features used in the model. It seems white matter tract alterations during aging process are significantly driving by the changes in limbic tracts.

The association of brain-PAD in different tract groups and SNPs led to the identification of one significant locus with leading



Measures: 1-Duration of heavy DIY, 2-Duration of light DIY, 3-Duration of walk, 4-Duration of walk for pleasure, 5-Time spent driving, 6-Time spent using computer, 7-Time spent watching TV, 8-Length of mobile phone use, 9-Plays computer games, 10-Sleep duration, 11-Getting up in morning, 12-Nap during day, 13-Sleeplessness / insomnia, 14-Snoring, 15-Ever smoked, 16-Smoking status, 17-Alcohol drinker status, 18-Former alcohol drinker, 19-Time spend outdoors in summer, 20-Time spent outdoors in winter, 21-Cooked vegetable intake, 22-Salad / raw vegetable intake, 23-Fresh fruit intake, 24-Dried fruit intake, 25-Oily fish intake, 26-Non-oily fish intake, 27-Processed meat intake, 28-Poultry intake, 29-Beef intake, 30-Lamb/mutton intake, 31-Pork intake, 32-Never eat eggs, dairy,wheat, sugar, 33-Cheese intake, 34-Bread intake, 35-Tea intake, 36-Coffee intake, 37-Water intake, 38-Alcohol frequency intake

Figure 4. Association of daily lifestyle measures and brain-PAD. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error. Grey color indicates non-significant association. A unique color was assigned to each group measures(e.g physical activity).

SNP rs1045537 ($p < 5 \times 10^{-8}$) in projection fibers. There was association of rs1045537 SNP and an eQTL of *BTN3A2* in heart left ventricle, basal ganglia, frontal Cortex and cortex. *BTN3A2* gene has been identified as a potential risk gene

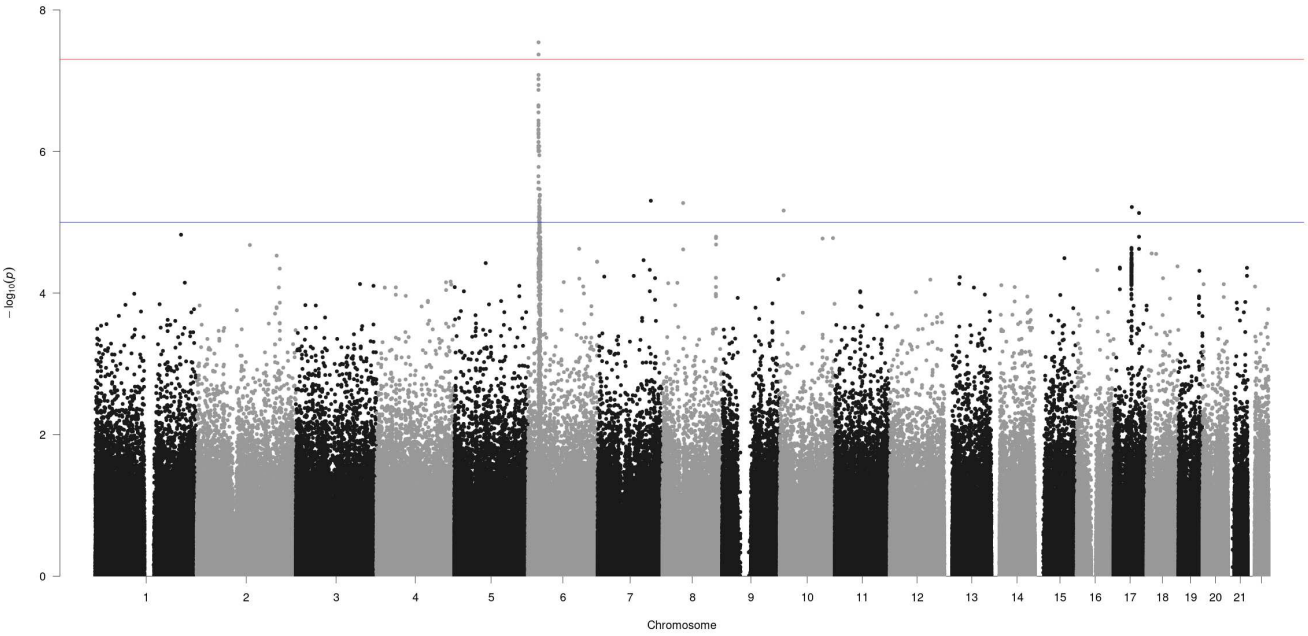


Figure 5. Manhattan plot reporting the association results between SNPs and brain-PAD in Projection FG. The red line indicates the GWAS threshold on p-value (i.e.,5E-8), while the blue line indicates the suggestive threshold of p=5E-5.

for schizophrenia[55,56]. The SNP is also significantly associated with malabsorption/coeliac disease, body mass index and headache. *HIST1H1A* gene was associated with DNA methylation at early AD stages[57]. *SCGN* gene was identified as one of the most common psychostimulants in brain-wide targets[58]. *SLC17A2* is one of the solute carrier family that is membrane protein and transporter. It was associated with neurodegenerative disorders because of its important role in the recovery of neurotransmitters[59].

Despite the fact that the associations of SNPs and brain-PAD in other tract groups did not pass the significance threshold, many of them did pass the threshold for suggestive significant association ($p < 5E-5$). In addition, the pattern of the association was more similar in some tract groups than others. For instance, the GWAS analysis in the limbic and ensemble models is similar, that is many SNPs passed the suggestive significant association threshold in chromosome 17 and many other variants in other chromosomes. Moreover, the pattern of the association was more close in association and projection fibers which many SNPs in same locus in chromosome 6 passing the suggestive significant threshold. The GWAS analysis in the commissural model followed a different pattern while very few SNPs passed the suggestive significant threshold in brainstem fibers. The results of GWAS analysis is consistent with the other analysis: we could observe more similarity in limbic and ensemble models, followed by the projection and association models while less SNPs passed the significant threshold in the brainstem model which was the case in the other analysis.

Most of CRFs and CMR measures led to significant association with brain-PAD in different tract groups. The direction of the association was shared by all tract groups. Brain-PAD in limbic tract groups associated significantly (5 positively and 5 negatively) with most of these measures comparing with other tract groups. Brainstem brain-PAD associated with 4 measures negatively and 5 positively. Positive associations was two times compared to negative with brain-PAD in the case when all IDPs were used to model brain age, while association and commissural seem less affected. Among these measures, diabetes, hypertension and LVM associated positively with brain-PAD in all tract groups while RVEDV, RVESV and RVSV associated negatively with brain-PAD in all tract groups. The other measures showed inconsistent association with brain-PAD across different tract groups. Body mass index and diabetes are positively associated with brain-PAD which indicates acceleration in brain aging. Based on the results, people who suffer from diabetes might accelerate brain aging by about half year, while in[5] it was estimated to cause acceleration about 2 years. The difference might be related to the features preprocessing and normalization steps. Body mass index has been associated with risk of developing a neurodegeneration or cognitive decline. Increasing in adiposity in overweight and obese individuals alter the white matter volume that causes faster brain aging up to 10 years[60]. Cardiac index has significant association with brain aging even for healthy people. People who presents a lower cardiac index or least pumping blood from heart to their body size appeared almost 2 years more brain aging from those who

have highest cardiac index[15]. Moreover, low cardiac index might increase risk of brain disorders. In[61] they have concluded that low cardiac index increase the risk of incident Dementia and AD.

Regarding daily lifestyle factors and activities, limbic tracts appeared more affected by these measures as 11 of them were significantly associated with brain-PAD derived from limbic tracts. Among them, 6 causes acceleration in brain aging and 5 push in the opposite direction. On the contrary, the results showed that brainstem tracts are less affected by the measures with only 4 measures being significantly associated after p-value was corrected. Three of them increase brain aging and only one has negative effect. Association, commissural, projection tracts showed close results of 7 significant associations for each one of them. In the case of using all fibers, 9 measures were significantly associated with brain-PAD, and mostly positively associated. Among all these measures, smoking statues and alcohol drinker status was significant in all cases. Alcohol frequency intake negatively associated with brain-PAD which indicate acceleration in brain aging. This has been confirmed before in other studies. Of note, alcohol frequency intake was coded as lower value means higher intake. In[62], alcohol intake history was negatively associated with white matter volume specially in corpus callosum. In addition, alcohol frequency intake was associated with deleterious in white matter tracts cause atrophy in ensemble and regional brain[63]. Our findings are inline with findings from previous studies and this was observed in most tract groups. Smoking is associated positively with brain-PAD suggesting that smoking has a negative impact on brain aging. It should be noted that smoking habits is being associated with different alterations in both white and grey matter. Moreover, smoking duration linked with reduced total volume of white matter. It is also associated with reduction in fractional anisotropy in cingulate gyrus[64]. Lamb/mutton intake was also positively associated with brain-PAD in some tract groups. Low red meat intake has been associated with better cognitive function[65]. In addition, limited consumption of red meat might reduce risk of AD, slow cognitive decline and reduce AD biomarker such as atrophy[66]. Duration of walk for pleasure associated negatively with brain-PAD which lead to a healthy brain aging. This might be due to walking stimulate blood circulations and expose to the sun light.

To conclude, one of the main focus of the current study is the interpretation of the results due to the use of highly heterogenous data from different domains. Overall, consistent results were obtained regarding the associations in different tract groups. In addition, some tract groups showed similar pattern of the different performed associations. One of the main contributions of the study shows which tract groups are more affected by the tested measures. Furthermore, we discovered that the limbic tracts have valuable role to drive aging in brain white matter tracts. In addition, brainstem observed to age faster and lest affected by the tested measures, precisely with daily life styles and activities. This could be explained that brainstem might age faster as it is more involved in many body functions. Benarroch[67] revealed that brainstem involves tracts that are critically associated with control of cardiovascular function, respiration, arousal and wake-sleep cycle. In that matter, brainstem tracts are more prone to alterations due to direct association with body organs. These findings suggest that further research is required to obtain a more comprehensive understanding of the role of limbic and brainstem tracts in brain aging and their association with both body functions and environmental exposures.

References

1. Niu, X., Zhang, F., Kounios, J. & Liang, H. Improved prediction of brain age using multimodal neuroimaging data. *Hum. Brain Mapp.* **41**, 1626–1643, DOI: https://doi.org/10.1002/hbm.24899 (2020).
2. Boyle, R. *et al.* Brain-predicted age difference score is related to specific cognitive functions: A multi-site replication analysis. *Brain imaging behavior* **15**, 327–345 (2021).
3. Jónsson, B. A. *et al.* Brain age prediction using deep learning uncovers associated sequence variants. *Nat. communications* **10**, 1–10, DOI: https://doi.org/10.1038/s41467-019-13163-9 (2019).
4. Smith, S. M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T. E. & Miller, K. L. Estimation of brain age delta from brain imaging. *Neuroimage* **200**, 528–539, DOI: https://doi.org/10.1016/j.neuroimage.2019.06.017 (2019).
5. Cole, J. H. Multi-modality neuroimaging brain-age in uk biobank: relationship to biomedical, lifestyle and cognitive factors. *Neurobiol. Aging* DOI: https://doi.org/10.1016/j.neurobiolaging.2020.03.014 (2020).
6. Ning, K., Zhao, L., Matloff, W., Sun, F. & Toga, A. W. Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants. *Sci. reports* **10**, 1–10, DOI: https://doi.org/10.1038/s41598-019-56089-4 (2020).
7. Feng, T., Lacrampe, A. & Hu, F. Physiological and pathological functions of tmem106b: a gene associated with brain aging and multiple brain disorders. *Acta neuropathologica* 1–13, DOI: https://doi.org/10.1007/s00401-020-02246-3 (2021).
8. Harding, S. R. *et al.* The tmem106b risk allele is associated with lower cortical volumes in a clinically diagnosed frontotemporal dementia cohort. *J. Neurol. Neurosurg. & Psychiatry* **88**, 997–998, DOI: http://dx.doi.org/10.1136/jnnp-2017-315641 (2017).
9. Pascale, E. *et al.* Genetic architecture of mapt gene region in parkinson disease subtypes. *Front. cellular neuroscience* **10**, 96, DOI: https://doi.org/10.3389/fncel.2016.00096 (2016).

10. Smith, S. M. *et al.* Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *Elife* **9**, e52677, DOI: https://doi.org/10.7554/eLife.52677 (2020).

11. van der Velpen, I. F., Yancy, C. W., Sorond, F. A. & Sabayan, B. Impaired cardiac function and cognitive brain aging. *Can. J. Cardiol.* **33**, 1587–1596, DOI: https://doi.org/10.1016/j.cjca.2017.07.008 (2017).

12. Jefferson, A. L. *et al.* Cardiac index is associated with brain aging: the framingham heart study. *Circulation* **122**, 690, DOI: 10.1161/CIRCULATIONAHA.109.905091 (2010).

13. Tahsili-Fahadan, P. & Geocadin, R. G. Heart–brain axis: effects of neurologic injury on cardiovascular function. *Circ. research* **120**, 559–572, DOI: https://doi.org/10.1161/CIRCRESAHA.116.308446 (2017).

14. Takeda, J. R. T., Matos, T. M. & Souza-Talarico, J. N. d. Cardiovascular risk factors and cognitive performance in aging. *Dementia & neuropsychologia* **11**, 442–448, DOI: https://doi.org/10.1590/1980-57642016dn11-040015 (2017).

15. Jefferson, A. L. Cardiac output as a potential risk factor for abnormal brain aging. *J. Alzheimer's Dis.* **20**, 813–821, DOI: https://doi.org/10.3233/JAD-2010-100081 (2010).

16. de Lange, A.-M. G. *et al.* Multimodal brain-age prediction and cardiovascular risk: The whitehall ii mri sub-study. *NeuroImage* **222**, 117292, DOI: https://doi.org/10.1016/j.neuroimage.2020.117292 (2020).

17. Franke, K. & Gaser, C. Ten years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained? *Front. neurology* **10**, 789 (2019).

18. Miller, K. L. *et al.* Multimodal population brain imaging in the uk biobank prospective epidemiological study. *Nat. neuroscience* **19**, 1523–1536, DOI: https://doi.org/10.1038/nn.4393 (2016).

19. Jiang, H. *et al.* Predicting brain age of healthy adults based on structural mri parcellation using convolutional neural networks. *Front. neurology* **10**, 1346, DOI: https://doi.org/10.3389/fneur.2019.01346 (2020).

20. Rokicki, J. *et al.* Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Hum. brain mapping* **42**, 1714–1726, DOI: https://doi.org/10.1002/hbm.25323 (2021).

21. Liem, F. *et al.* Predicting brain-age from multimodal imaging data captures cognitive impairment. *Neuroimage* **148**, 179–188, DOI: https://doi.org/10.1016/j.neuroimage.2016.11.005 (2017).

22. Soares, J., Marques, P., Alves, V. & Sousa, N. A hitchhiker's guide to diffusion tensor imaging. *Front. neuroscience* **7**, 31, DOI: https://doi.org/10.3389/fnins.2013.00031 (2013).

23. Nemanich, S. T., Mueller, B. A. & Gillick, B. T. Neurite orientation dispersion and density imaging quantifies corticospinal tract microstructural organization in children with unilateral cerebral palsy. *Hum. brain mapping* **40**, 4888–4900, DOI: https://doi.org/10.1002/hbm.24744 (2019).

24. Sone, D. Neurite orientation and dispersion density imaging: clinical utility, efficacy, and role in therapy. *Reports Med. Imaging* **12**, 17, DOI: https://doi.org/10.2147/RMI.S194083 (2019).

25. Merluzzi, A. P. *et al.* Age-dependent differences in brain tissue microstructure assessed with neurite orientation dispersion and density imaging. *Neurobiol. aging* **43**, 79–88, DOI: https://doi.org/10.1016/j.neurobiolaging.2016.03.026 (2016).

26. Beck, D. *et al.* White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. *NeuroImage* **224**, 117441, DOI: https://doi.org/10.1016/j.neuroimage.2020.117441 (2021).

27. Yang, A. C., Tsai, S.-J., Liu, M.-E., Huang, C.-C. & Lin, C.-P. The association of aging with white matter integrity and functional connectivity hubs. *Front. aging neuroscience* **8**, 143, DOI: https://doi.org/10.3389/fnagi.2016.00143 (2016).

28. Bender, A. R., Völkle, M. C. & Raz, N. Differential aging of cerebral white matter in middle-aged and older adults: a seven-year follow-up. *Neuroimage* **125**, 74–83, DOI: https://doi.org/10.1016/j.neuroimage.2015.10.030 (2016).

29. Huang, H. *et al.* Distinctive disruption patterns of white matter tracts in alzheimer's disease with full diffusion tensor characterization. *Neurobiol. aging* **33**, 2029–2045, DOI: https://doi.org/10.1016/j.neurobiolaging.2011.06.027 (2012).

30. Petersen, S. E. *et al.* Uk biobank's cardiovascular magnetic resonance protocol. *J. cardiovascular magnetic resonance* **18**, 8, DOI: https://doi.org/10.1186/s12968-016-0227-4 (2015).

31. Raisi-Estabragh, Z., Harvey, N. C., Neubauer, S. & Petersen, S. E. Cardiovascular magnetic resonance imaging in the uk biobank: a major international health research resource. *Eur. Hear. Journal-Cardiovascular Imaging* DOI: https://doi.org/10.1093/ehjci/jeaa297 (2020).

32. Bycroft, C. *et al.* The uk biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209, DOI: https://doi.org/10.1038/s41586-018-0579-z (2018).

33. Alfaro-Almagro, F. *et al.* Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. *Neuroimage* **166**, 400–424, DOI: https://doi.org/10.1016/j.neuroimage.2017.10.034 (2018).

34. Hua, K. *et al.* Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage* **39**, 336–347, DOI: https://doi.org/10.1016/j.neuroimage.2007.07.053 (2008).

35. Wakana, S., Jiang, H., Nagae-Poetscher, L. M., Van Zijl, P. C. & Mori, S. Fiber tract–based atlas of human white matter anatomy. *Radiology* **230**, 77–87, DOI: https://doi.org/10.1148/radiol.2301021640 (2004).

36. Standring, S. *et al.* Gray's anatomy: the anatomical basis of clinical practice. *Am. journal neuroradiology* **26**, 2703, DOI: https://doi.org/10.5860/choice.43-1300 (2005).

37. Haines, D. E. & Mihailoff, G. A. *Fundamental Neuroscience for Basic and Clinical Applications E-Book* (Elsevier Health Sciences, 2017).

38. Basinger, H. & Hogg, J. P. Neuroanatomy, brainstem. *StatPearls [Internet]* (2020).

39. Pascalau, R., Stănilă, R. P., Sfrângeu, S. & Szabo, B. Anatomy of the limbic white matter tracts as revealed by fiber dissection and tractography. *World neurosurgery* **113**, e672–e689, DOI: https://doi.org/10.1016/j.wneu.2018.02.121 (2018).

40. Catani, M., Dell'Acqua, F. & De Schotten, M. T. A revised limbic system model for memory, emotion and behaviour. *Neurosci. & Biobehav. Rev.* **37**, 1724–1737, DOI: https://doi.org/10.1016/j.neubiorev.2013.07.001 (2013).

41. Wu, Y., Sun, D., Wang, Y., Wang, Y. & Ou, S. Segmentation of the cingulum bundle in the human brain: a new perspective based on dsi tractography and fiber dissection study. *Front. neuroanatomy* **10**, 84, DOI: https://doi.org/10.3389/fnana.2016.00084 (2016).

42. Daroff, R. B. & Aminoff, M. J. *Encyclopedia of the neurological sciences* (Academic press, 2014).

43. Kulkarni, N. V. *Clinical anatomy (a problem solving approach)* (JP Medical Ltd, 2011).

44. Attar, R. *et al.* Quantitative cmr population imaging on 20,000 subjects of the uk biobank imaging study: Lv/rv quantification pipeline and its evaluation. *Med. image analysis* **56**, 26–42, DOI: https://doi.org/10.1016/j.media.2019.05.006 (2019).

45. Du Bois, D. A formula to estimate the approximate surface area if height and weight be known. *Nutrition* **5**, 303–313, DOI: https://doi.org/10.1001/archinte.1916.00080130010002 (1989).

46. Laurent, S. *et al.* Expert consensus document on arterial stiffness: methodological issues and clinical applications. *Eur. heart journal* **27**, 2588–2605, DOI: https://doi.org/10.1093/eurheartj/ehl254 (2006).

47. Alfaro-Almagro, F. *et al.* Confound modelling in uk biobank brain imaging. *NeuroImage* **224**, 117002, DOI: https://doi.org/10.1016/j.neuroimage.2020.117002 (2021).

48. Beheshti, I., Nugent, S., Potvin, O. & Duchesne, S. Bias-adjustment in neuroimaging-based brain age frameworks: A robust scheme. *NeuroImage: Clin.* **24**, 102063, DOI: https://doi.org/10.1016/j.nicl.2019.102063 (2019).

49. Le, T. T. *et al.* A nonlinear simulation framework supports adjusting for age when analyzing brainage. *Front. aging neuroscience* **10**, 317, DOI: https://doi.org/10.3389/fnagi.2018.00317 (2018).

50. Bonferroni, C. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Super. di Scienze Econ. e Commericiali di Firenze* **8**, 3–62 (1936).

51. Purcell, S. *et al.* Plink: a tool set for whole-genome association and population-based linkage analyses. *The Am. journal human genetics* **81**, 559–575, DOI: https://doi.org/10.1086/519795 (2007).

52. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with fuma. *Nat. communications* **8**, 1–11, DOI: http://dx.doi.org/10.1038/s41467-017-01261-5 (2017).

53. Elliott, L. T. *et al.* Genome-wide association studies of brain imaging phenotypes in uk biobank. *Nature* **562**, 210–216, DOI: https://doi.org/10.1016/10.1038/s41586-018-0571-7 (2018).

54. Zhang, Y. *et al.* Diffusion tensor tractography of brainstem fibers and its application in pain. *PloS one* **15**, e0213952 (2020).

55. Wu, Y. *et al.* Identification of the primate-specific gene btn3a2 as an additional schizophrenia risk gene in the mhc loci. *EBioMedicine* **44**, 530–541, DOI: https://doi.org/10.1016/j.ebiom.2019.05.006 (2019).

56. Bhalala, O. G. *et al.* Identification of expression quantitative trait loci associated with schizophrenia and affective disorders in normal brain tissue. *PLoS genetics* **14**, e1007607, DOI: https://doi.org/10.1371/journal.pgen.1007607 (2018).

57. Blanco-Luquin, I. *et al.* Early epigenetic changes of alzheimer's disease in the human hippocampus. *Epigenetics* **15**, 1083–1092, DOI: https://doi.org/10.1080/15592294.2020.1748917 (2020).

58. Fuzik, J. *et al.* Brain-wide genetic mapping identifies the indusium griseum as a prenatal target of pharmacologically unrelated psychostimulants. *Proc. Natl. Acad. Sci.* **116**, 25958–25967 (2019).

59. Aykaç, A. & Sehirli, A. O. The role of the slc transporters protein in the neurodegenerative disorders. *Clin. Psychopharmacol. Neurosci.* **18**, 174 (2020).

60. Ronan, L. *et al.* Obesity associated with increased brain age from midlife. *Neurobiol. aging* **47**, 63–70, DOI: https://doi.org/10.1016/j.neurobiolaging.2016.07.010 (2016).

61. Jefferson, A. L. *et al.* Low cardiac index is associated with incident dementia and alzheimer disease: the framingham heart study. *Circulation* **131**, 1333–1339, DOI: https://doi.org/10.1161/CIRCULATIONAHA.114.012438 (2015).

62. Ruiz, S. M. *et al.* Drinking history associations with regional white matter volumes in alcoholic men and women. *Alcohol. Clin. Exp. Res.* **37**, 110–122, DOI: https://doi.org/10.1111/j.1530-0277.2012.01862.x (2013).

63. McEvoy, L. K. *et al.* Alcohol intake and brain white matter in middle aged men: Microscopic and macroscopic differences. *NeuroImage: Clin.* **18**, 390–398, DOI: https://doi.org/10.1016/j.nicl.2018.02.006 (2018).

64. Gray, J. C. *et al.* Associations of cigarette smoking with gray and white matter in the uk biobank. *Neuropsychopharmacology* **45**, 1215–1222, DOI: https://doi.org/10.31234/osf.io/wyvnm (2020).

65. Corley, J. *et al.* Dietary patterns, cognitive function, and structural neuroimaging measures of brain aging. *Exp. Gerontol.* 111117, DOI: https://doi.org/10.1016/j.exger.2020.111117 (2020).

66. Walters, M., Hackett, K., Caesar, E., Isaacson, R. & Mosconi, L. Role of nutrition to promote healthy brain aging and reduce risk of alzheimer's disease. *Curr. Nutr. Reports* **6**, 63–71, DOI: https://doi.org/10.1007/s13668-017-0199-5 (2017).

67. Benarroch, E. E. Brainstem integration of arousal, sleep, cardiovascular, and respiratory control. *Neurology* **91**, 958–966, DOI: https://doi.org/10.1212/WNL.0000000000006537 (2018).

Acknowledgements

Data access was through UK Biobank access application 2964. Ahmed is supported by INVITE program co-financed by the European Union within the Horizon 2020 Programme and by the Regione del Veneto. ZRE is supported by a British Heart Foundation Clinical Research Training Fellowship (FS/17/81/33318). SEP acknowledges support from the Barts Biomedical Research Centre funded by the National Institute for Health Research (NIHR).

Author contributions statement

A.S conceived and conducted the experiments. Also writing the first version of the manuscript.
I.B provided feedback and guide the experiments. She also guided in writing the manuscript to finalize it.
Z.R helped in interpreting the results clinically. She also helped in writing the first manuscript.
E.R helped in interpreting the results clinically.
P.G helped in providing feedback.
S.P and K.L provide access to UK biobank data.
A.A helped to guided the genetic analysis.
P.R helped in advising.
G.M is the principal investigator.
All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Figures

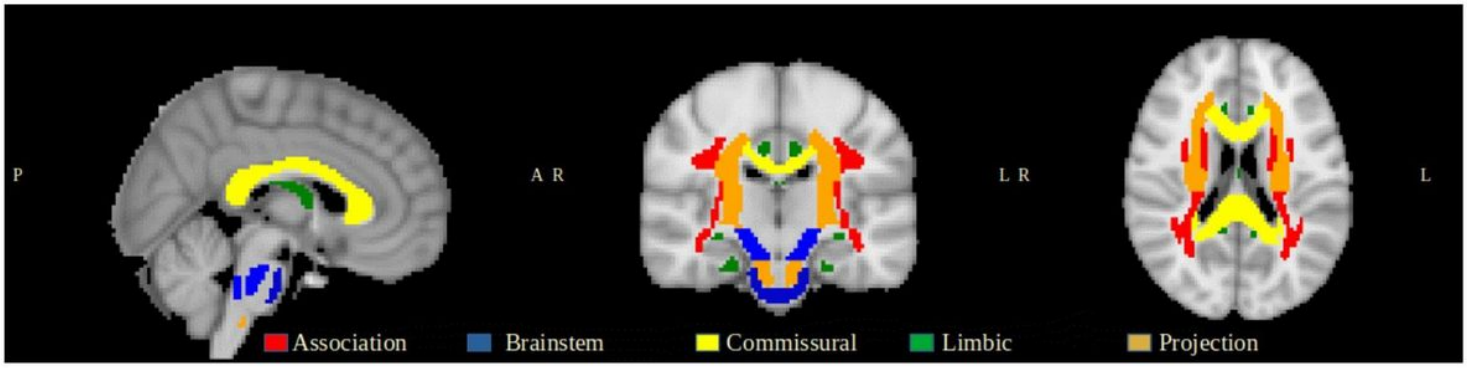


Figure 1

White matter tract groups

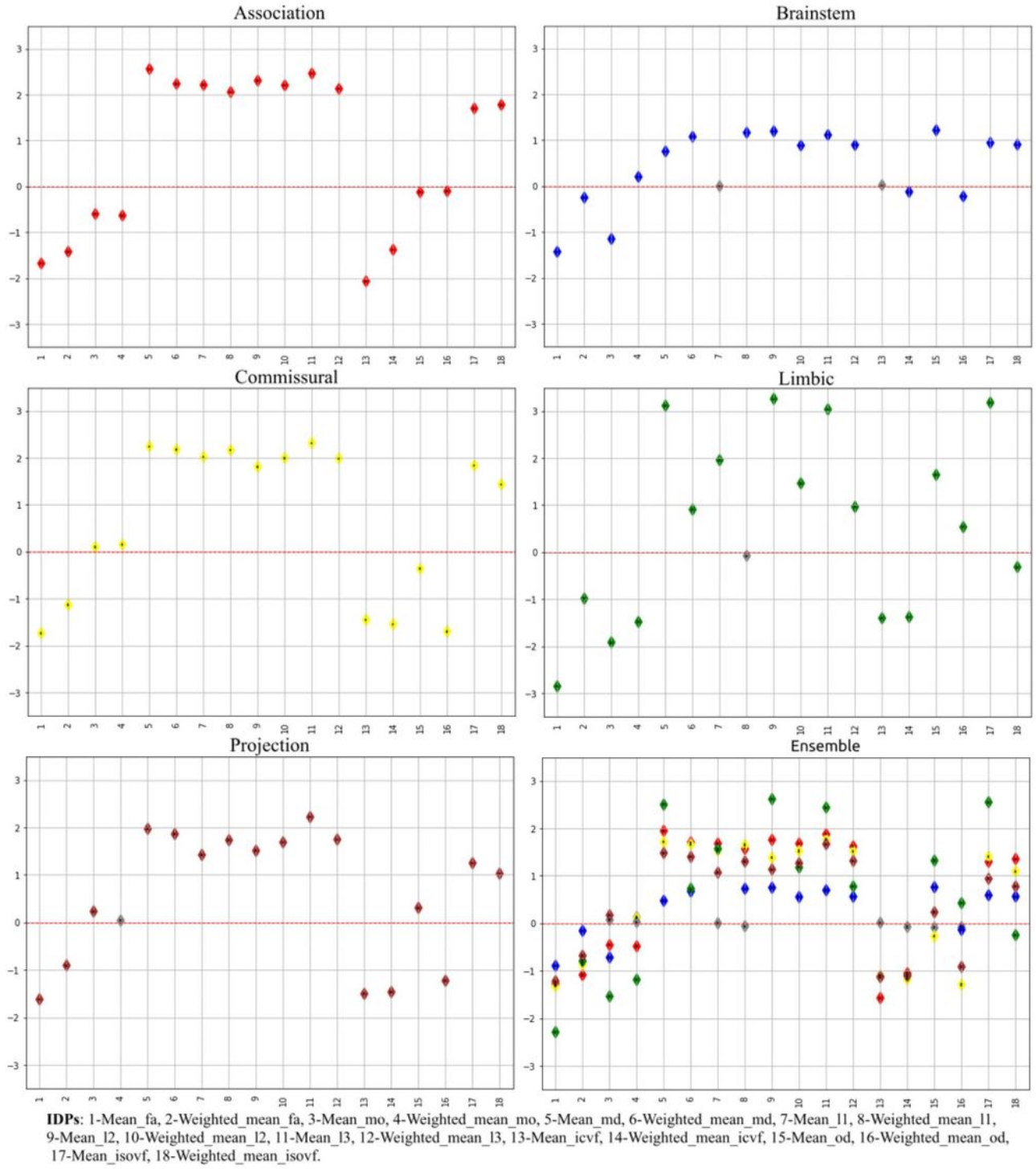


Figure 2

Association of the IDPs and brain-PAD for the different models. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error (the small black dot inside the diamond shape). Grey color indicates non-significant association.

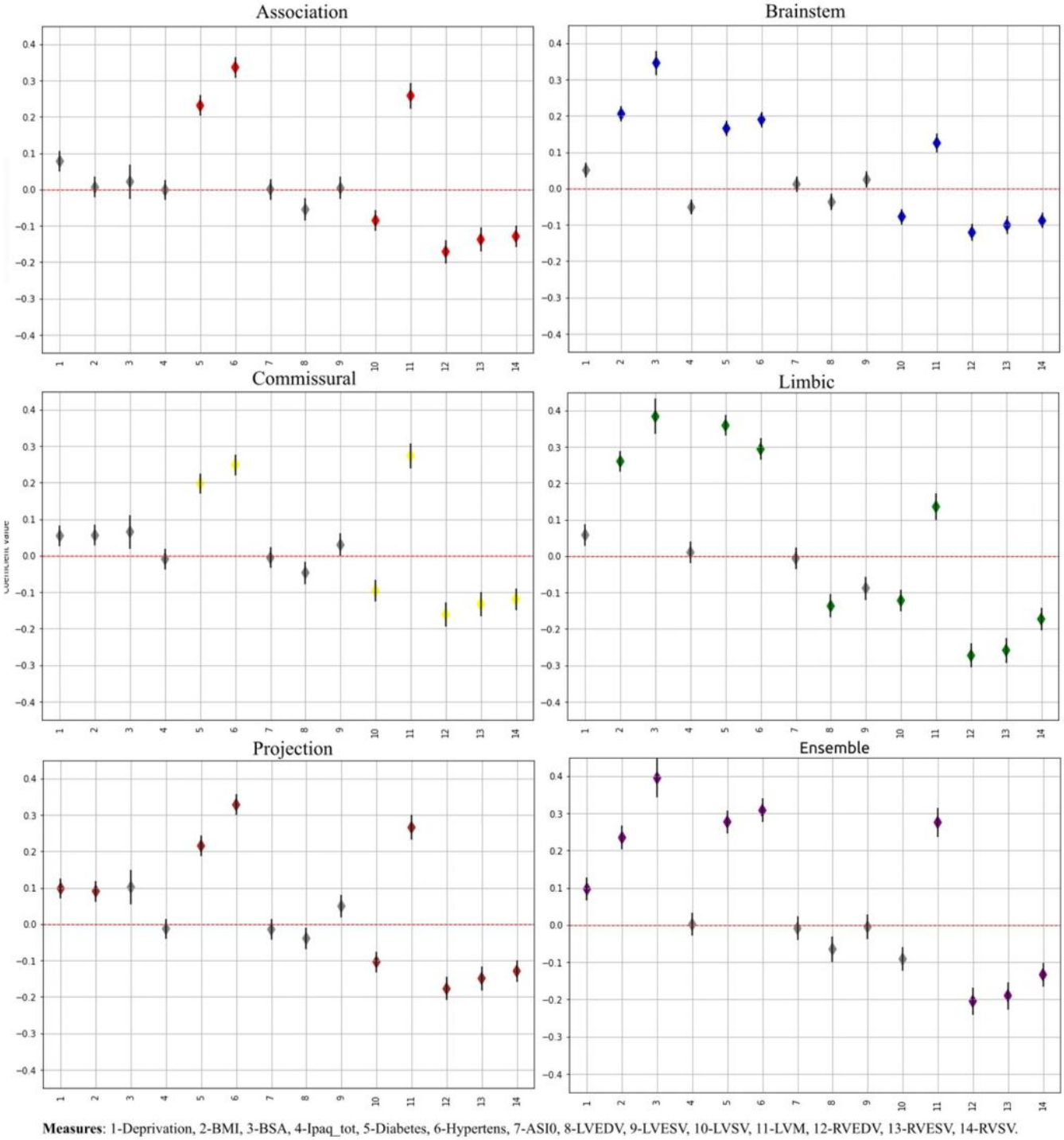
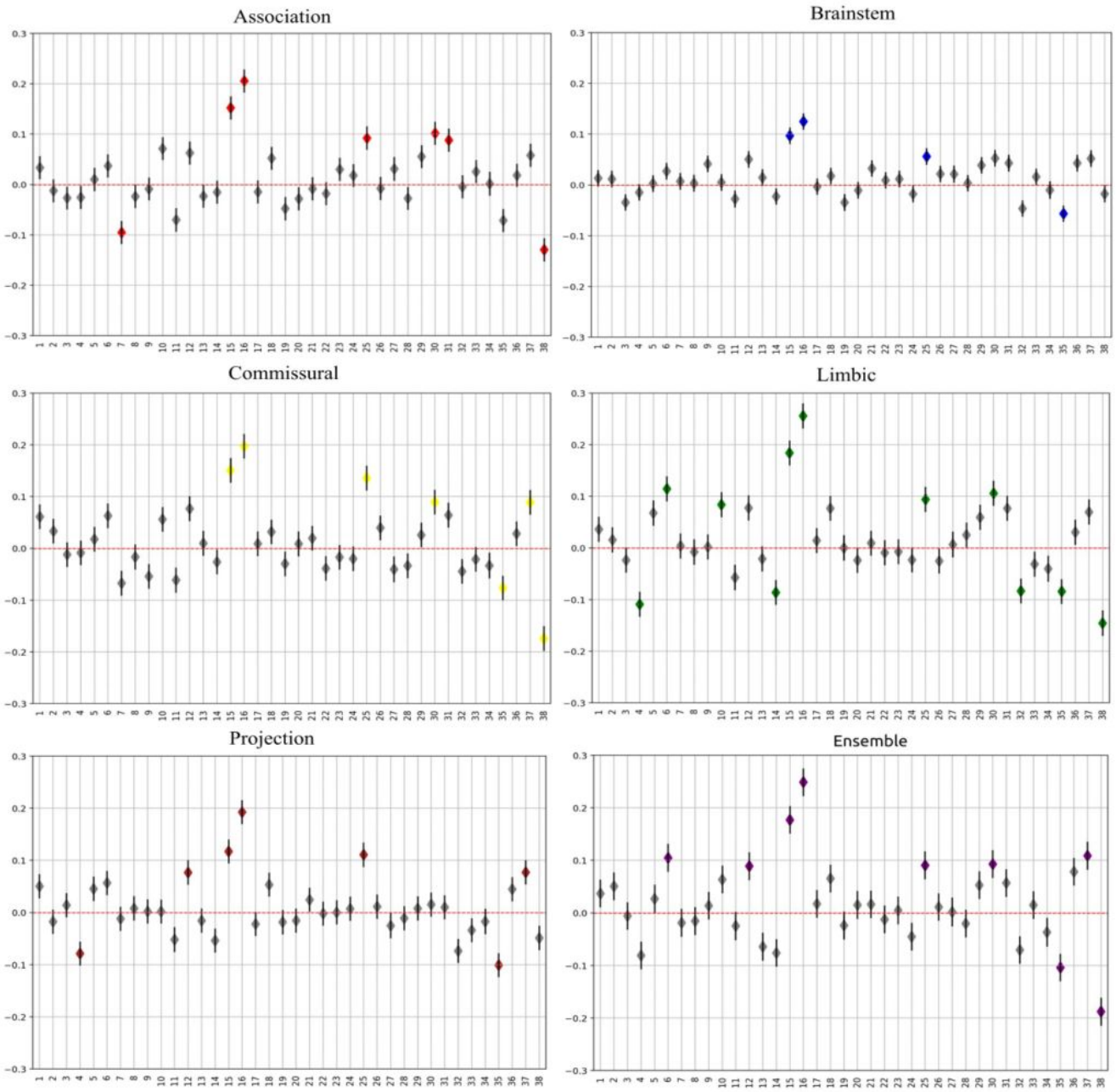


Figure 3

Association of the CMR measures, CRFs and brain-PAD. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error. Grey color indicates non-significant association.



Measures: 1-Duration of heavy DIY, 2-Duration of light DIY, 3-Duration of walk, 4-Duration of walk for pleasure, 5-Time spent driving, 6-Time spent using computer, 7-Time spent watching TV, 8-Length of mobile phone use, 9-Plays computer games, 10-Sleep duration, 11-Getting up in morning, 12-Nap during day, 13-Sleeplessness / insomnia, 14-Snoring, 15-Ever smoked, 16-Smoking status, 17-Alcohol drinker status, 18-Former alcohol drinker, 19-Time spend outdoors in summer, 20-Time spent outdoors in winter, 21-Cooked vegetable intake, 22-Salad / raw vegetable intake, 23-Fresh fruit intake, 24-Dried fruit intake, 25-Oily fish intake, 26-Non-oily fish intake, 27-Processed meat intake, 28-Poultry intake, 29-Beef intake, 30-Lamb/mutton intake, 31-Pork intake, 32-Never eat eggs, dairy,wheat, sugar, 33-Cheese intake, 34-Bread intake, 35-Tea intake, 36-Coffee intake, 37-Water intake, 38-Alcohol frequency intake

Figure 4

Association of daily lifestyle measures and brain-PAD. For each model, the x-axis represents the different IDPs summarised in the legend, while the regression coefficient values are reported in the y-axis along with their standard error. Grey color indicates non-significant association. A unique color was assigned to each group measures(e.g physical activity).

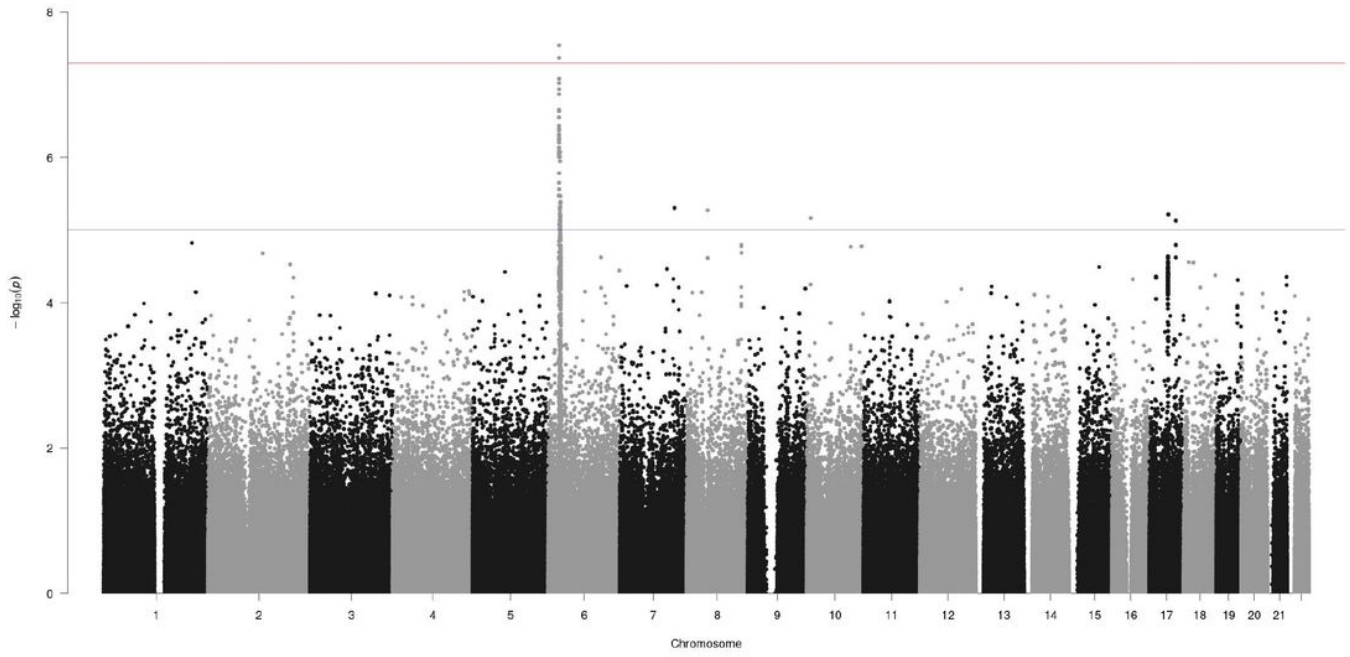


Figure 5

Manhattan plot reporting the association results between SNPs and brain-PAD in Projection FG. The red line indicates the GWAS threshold on p-value (i.e.,5E-8), while the blue line indicates the suggestive threshold of p=5E-5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.zip