DISEASE AND AGEING BIOMARKER IDENTIFICATION FROM MULTIMODAL MAGNETIC RESONANCE IMAGING

Doctoral Dissertation by
Max Korbmacher

Thesis submitted for
the degree of Philosophiae Doctor (PhD)

Department of Health and Functioning,
Radiography

Faculty of Health and Social Sciences

Western Norway University of Applied Sciences

14 November 2023

PREFACE

The author of this thesis has been employed as a Ph.D. research fellow in Neuroscience at the Department of Health and Functioning, Section for Radiography at the Western Norway University of Applied Sciences, Bergen, Norway.

The research presented in this thesis has been accomplished in cooperation with the Norwegian Centre for Mental Disorders Research (NORMENT, University of Oslo) and Mohn Medical Imaging and Visualization Centre Bergen.

The candidate was enrolled in the PhD programme in health, function and participation in Health and Social Sciences, with the specialisation of this thesis, being in Neuroscience.

This thesis is organized in two parts. Part I provides an introduction to brain health and disease biomarkers (as part of the neuroimaging field), embedding the executed studies within this PhD project in the field. Part II consists of a collection of published and peer-reviewed research articles and submitted papers which are part of this PhD project.

Paper A **M. Korbmacher**, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov (2023). Brain-wide associations between white matter and agehighlight the role of fornix microstructure in brain ageing. *Human Brain Mapping*, 44(10), 4101–4119.

Paper B **M. Korbmacher**, T. P. Gurholt, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov (2023). Bio-psycho-social factors' associations with brain age: a large-scale UK Biobank diffusion study of 35,749 participants. *Frontiers in Psychology*, 14, 1117732.

Paper C **M. Korbmacher**, M.-Y. Wang, R. Eikeland, R. Buchert, O. A. Andreassen, T. Espeseth, E. Leonardsen, L. T. Westlye, I. I. Maximov, K. Specht (2023). Considerations on brain age predictions from repeatedly sampled data across time. *Brain and Behavior*.

Paper D **M. Korbmacher**, D. van der Meer, D. Beck, A. M. de Lange, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov (2023). Brain asymmetries from mid- to late life and hemispheric brain age. *bioRxiv*. (under review).

Paper E **M. Korbmacher**, D. van der Meer, D. Beck, D. E. Askeland-Gjerde, A. Lundervold, E. Eikefjord, O. A. Andreassen, L. T. Westlye, I. I. Maximov (2023). Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and Alzheimer's disease in the UK Biobank. *medRxiv*. (submitted).

ACKNOWLEDGMENTS

ABSTRACT

Introduction: Examining brain changes through a lifespan perspective allows for multiple ways of answering fundamental questions about the brain's development and to identify biomarkers of ageing and disease. One opportunity is to establish a connection between the brain's architecture and chronological age using brain magnetic resonance imaging (MRI) features. The resulting metric, often called *brain age*, can be used as a proxy of a person's general health status when comparing the prediction to the chronological age, and can hence be used as a general identifier of disease or symptom load. An adjacent goal was to establish a better understanding of the brain's development and ageing by examining MRI-features age dependencies, such as white and grey matter. Here, we provide an overview of five studies conducted during this PhD project, examining brain ageing as well as ageing and disease biomarkers, and contextualise these in the embedding of the wider research field.

Methods: We examined brain-wide and hemispheric data from the voxel-to-global-level in large cohorts and precision imaging data of healthy controls with a focus on age-associations. We used several data sets, yet mainly cross-sectional MRI data from the UK Biobank (N $\approx$ 50,000). Additionally, we used longitudinal UK Biobank data incorporating two time-points (N = 2,678), local cross-sectional data (Tematisk Område Psykoser (TOP) and Neurogenetics of cognition in aging (NCNG), N_{total} = 1,065), and repeatedly sampled data (Bergen Breakfast Scanning Club (BBSC) and Frequently Travelling Human Phantom (FTHP)) N_{total} = 4 (*M* = 660 scans). For one paper (Study C) we used minimally processed T_1-weighted 3D volumes, which were aligned in MNI space and empty-border cropped. All other studies used processed and regionally or globally atlas-averaged white matter (WM) and grey matter (GM) metrics (John Hopkins University WM atlas, and Desikan-Killiany atlas for GM and WM thickness, area, and volume, respectively). Moreover, after estimating WM microstructure metrics based on different biophysical models, we used tract-based spatial statistics to obtain diffusion metrics of conventional and advanced diffusion approaches within the fractional anisotropy (WM) skeleton. In brief, these approaches contain radial, axial and mean diffusivity metrics, fractional anisotropy and water fractions, kurtosis metrics, and additional information for these metrics for intra- and extra-axonal space. For brain age estimations, we applied both a tree-based algorithm (extreme gradient boosting) to processed and region-averaged multimodal magnetic resonance imaging metrics (2D), and a convolutional neural network (CNN) to minimally processed structural MRI data (3D).

Statistical analyses included associating the estimated brain ages with each other and chronological age. All regional and global WM and GM metrics and their asymmetries were age-associated for both cross-sectional and longitudinal data. We also associated the resulting WM brain age with various bio-psycho-social phenotypes, and assessed cross-sectional and longitudinal measures WM measures with polygenic risk scores.

Finally, we tested an existing brain age model's applicability in practical, clinical context (a pre-trained CNN using T_1-weighted MRI data) by assessing the models'

prediction error in repeatedly sampled data of few individuals and in a cross-sectional validation sample, as well as associations of brain age predictions with quality control metrics and field strength.

Results: We benchmarked age-associations of various GM and WM metrics in midlife to older ages and examined brain age estimated from varying MRI data in different samples. Global brain ageing was characterised by decreases in GM volume, surface area, and thickness, and WM fractional anisotropy, the intra-axonal water fraction, and kurtosis metrics decreases, an axial, radial, and mean diffusivity metrics as well as free water and extra-axonal water fractions increases. These trends were similar for regional age-associations and were presented as age charts of normative white and grey matter development across mid- and late life.

WM and GM regions presented some variability for age estimations, with multimodal models presenting most accurate predictions. Yet, brain ages estimated on metrics from different biophysical models were similarly associated to different bio-psycho-social factors. Moreover, models including diffusion MRI derived WM metrics provided slightly more accurate brain age estimates than T_1-weighted MRI derived metrics.

Across studies, the fornix repeatedly appeared as one of the most age-sensitive region across WM metrics in the UK Biobank. However also forceps minor was highly age-sensitive in addition to the middle cerebral peduncle, which was strongest related to the polygenic risk of Alzheimer's Disease. The annual rate of change in WM was a magnitude stronger associated to the polygenic risk of Alzheimer's Disease and psychiatric disorders than cross-sectional measures. However, effect sizes were small for both global and regional brain-polygenic risk associations.

We furthermore identified a tendency of higher regional GM and WM asymmetry at higher ages. In that sense, for GM, amygdala, hippocampus, pallidum, ventricle volume, thalamus and accumbens were strongest associated with age. For WM, the cingulate tract, unicate fasciculus, superior longitudinal fasciculus, and cerebral peduncle asymmetries were strongest associated with age. New measures of hemisphere-specific age-predictions were suggested and demonstrated promising results to further investigate asymmetric ageing of the brain or diseases affecting the brain asymmetrically. Predictions were highly similar across hemispheres, modalities and handedness-preference. Yet, important sex-differences applied to the sensitivity of brain age to the MRI modality and hemisphere.

When testing a pre-trained brain age model in repeatedly/densely sampled data, we found low correlations between predicted and chronological age in the repeatedly sampled data. In an attempt of explaining such variability by scan quality, we identified inconsistent associations of brain age with quality control parameters. We also found a stronger associations of brain and chronological ages at a higher field strength for one repeatedly sampled individual (where data was available at different field strengths), which was validated in cross-sectional samples.

Conclusion: Central and deep brain regions including the corpus callosum, the brain stem, the limbic system, and the ventricles were regions which were repeatedly strongly associated with age, ageing, and the polygenic risk of ageing related pathology. Among

these regions, fornix stood out as most prominent region. Fornix characteristics have been identified previously as biomarkers of Alzheimer's Disease (AD) progression. Subsequently, fornix has already been used as a target region for deep brain stimulation for AD treatment. We underlined these findings by showing the region's strong age-association, which might be useful to identify earlier stages of cognitive decline and neurodegenerative disease. Additionally, longitudinal changes unveiled a unique pattern of WM changes not only affecting the limbic system (including fornix) but also presenting an anterior-posterior gradient of WM change of WM loss and de-differentiation in superior frontal regions compared to differentiating and potentially plasticity in occipital, brain stem and cerebellar regions. We furthermore provided a spatially distributed pattern of associations between polygenic risk scores and WM, particularly outlining the cerebral peduncle. These genetically informed risk scores associations with brain WM were stronger for the annual change in WM than time-point specific/cross-sectional scores, emphasising the importance of using longitudinal data. The newly tested hemispheric brain age might hold some promise for precision medicine by assessing differences between left and right brain ages. Concerning the test of the existing model, part of the prediction error seems to be driven by differences in field strength. Yet, additional confounds need to be identified and addressed to move brain age towards higher clinical utility.

SAMMENDRAG

Introduksjon: Å undersøke hjerneforandringer gjennom et livstidsperspektiv gir mulighet for flere måter å svare på grunnleggende spørsmål om hjernens utvikling og blant annet å identifisere biomarkører for aldring og sykdom. Dette gir mulighet til å forstå sammenhenger mellom hjernens arkitektur og kronologisk alder ved hjelp av hjerneavbildninger fra magnetisk resonanstomografi (MRT). Det resulterende estimatet, ofte kalt *hjernealder*, kan brukes som et uttrykk for den generelle helsetilstanden til en person når man sammenligner prediksjonen med personens kronologisk alder. Hjernealder kan på denne måten brukes som en generell indikator for sykdom eller symptombelastning. Et tilstøtende mål var å etablere en bedre forståelse av hjernens utvikling og aldring ved å undersøke aldersavhengigheter i hjernen, som hvit og grå substans. I denne avhandlingen inkluderes fem studier som undersøker hjernealdring og biomarkører for aldring og sykdom. Videre kontekstualiseres disse studiene i det bredere forskningsfeltet.

Metode: Vi undersøkte hjernedata fra voxel til globalt nivå i store kohorter og presisjonsavbildningsdata fra friske kontrollgrupper med fokus på sammenheng med alder. Vi brukte flere datasett, men hovedsakelig tverrsnitts-MR-data fra den britiske kohorten UK Biobank ($N \approx 50\,000$). I tillegg brukte vi longitudinelle UK Biobank data fra to tidspunkter ($N = 2.678$), lokale tverrsnittsdata (TOP og NCNG, $N_{total} = 1.065$), og presisjonsdata med gjentatte opptak (BBSC og FTHP) $N_{total} = 4$ ($M = 660$ skanninger). For en studie (studie C) brukte vi minimalt behandlede T_1-vektede 3D-volumer, justert til MNI-rom og renset av tom rom. Alle andre studier brukte prosesserte og regionale eller globale atlas-gjennomsnitter fra hvit substans (WM) og grå substans (GM; John Hopkins University WM atlas, og Desikan-Killiany atlas for GM og WM tykkelse, areal og volum). For å estimere WM-mikrostrukturmålinger brukte vi forskjellige konvensjonelle og avanserte biofysiske diffusjonsmodeller. Traktbasert spatiell statistikk (TBSS) ble deretter brukt for å begrense beregningene til det fraksjonell anisotropi (WM) skjelettet. Kort forklart inneholder de biofysiske tilnærmingene radiell, aksial og gjennomsnittlig diffusjons, fraksjonell anisotropi og vannfraksjoner, kurtosis og tilleggsinformasjon for disse metrikkene for intra- og ekstra-aksonal rom for multishell-tilnærmingene. For å estimere hjernealder brukte vi både en trebasert algoritme (extreme gradient boosting) på prosesserte og regiongjennomsnittede multimodal magnetisk resonansavbildning (MRI) beregninger (2D), og et konvolusjonelt nevralt nettverk (CNN) på minimalt prosesserte strukturelle MR-data (3D).

Statistiske analyser inkluderte assosiering av estimert hjernealder med hverandre (kryss diffusjonsmodeller) og kronologisk alder. Alle regionale og globale WM- og GM-metrikker og deres asymmetrier ble assosiert med alder for både tverrsnitts- og longitudinelle data. Vi assosierte også de resulterende hvite substansens-baserte hjernealderne med forskjellige bio-psykososiale fenotyper, og vurderte tverrsnitts- og longitudinelle mål WM-mål med polygene risikoscore.

Til slutt testet vi en eksisterende hjernealdermodells anvendelighet i praktisk,

klinisk kontekst (en forhåndstrent CNN som bruker T_1-vektet MR-data) ved å vurdere modellenes prediksjonsfeil i gjentatte samplede data fra få individer og i en kryssseksjonsvalideringsprøve, samt assosiasjoner av hjernealdersprediksjoner med kvalitetskontrollberegninger og feltstyrke.

Resultater: Vi evaluerte aldersassosiasjoner av forskjellige GM og WM målinger fra midten av livet til høyere alder, og undersøkte hjernealder estimert fra multimodale MR-data av forskjellige stikkprøver. Global hjernealdring var kjennetegnet av reduksjoner i GM-volum, overflateareal og tykkelse, og WM-fraksjonell anisotropi, den intra-aksonale vannfraksjonen og kurtosis-metrikkene. Samtidig øker det aksial, radial og gjennomsnittlig diffusivitetsmålinger, samt fritt vann og ekstra -aksonale vannfraksjoner. Disse trendene var like for regionale aldersforandringer og ble presentert som aldersdiagrammer for normativ utvikling av hvit og grå substans på tvers av årene/alder.

WM- og GM-regioner presenterte en viss variasjon for aldersestimater, men multimodale modeller presterer mest nøyaktige. Likevel var hjernealder, som ble estimert på beregninger fra forskjellige biofysiske modeller, assosiert på samme måte med forskjellige bio-psykososiale faktorer. Hjernealdersmodeller som var basert på WM var litt mer nøyaktige enn modeller som var basert på GM/T_1-vektet MR.

På tvers av studier dukket fornix gjentatte ganger opp som den mest alderssensitive regionen på tvers av WM-målinger i UK biobanken. Men også forceps minor var svært alderssensitiv i tillegg til den midtre lillehjernens peduncle, som var sterkest relatert til den polygene risikoen for Alzheimers sykdom. Dessuten var den årlige endringsraten i WM en størrelsesorden sterkere assosiert med den polygene risikoen for Alzheimers sykdom og psykiatriske lidelser enn tverrsnittsmål. Effektstørrelsene var imidlertid små for både globale og regionale hjerne-polygene risikoassosiasjoner.

Vi identifiserte videre en tendens til høyere regional GM og WM asymmetri ved høyere aldre. Sånn sett, for GM, var amygdala, hippocampus, pallidum, ventrikkel volumen, thalamus og accumbens sterkest assosiert med alder. For WM var cingulate tract, unicate fasciculus, og superior longitudinal fasciculus asymmetri sterkest assosiert med alder. En ny måte å beregne hjernealder på (én prediksjon per hemisfere) ble introdusert. Målet viste lovende resultater for å undersøke asymmetrisk aldring av hjernen eller sykdommer som påvirker hjernen asymmetri. Forutsigelser var svært like på tvers av hemisferer, modaliteter og håndpreferanse, mens det finnes viktige kjønnsforskjeller i sammenhengen av hjernealder med hemisfere og modalitet.

Når vi testet en forhåndstrent hjernealdermodell i gjentatte/tett samplede data, fant vi lave korrelasjoner mellom predikert og kronologisk alder i de gjentatte samplede dataene. I et forsøk på å forklare slik variasjon ved skannekvalitet, identifiserte vi inkonklusive assosiasjoner til hjernealder med en rekke kvalitetskontrollparametere. Vi fant også en sterkere assosiasjon av hjerne og kronologiske aldre ved en høyere feltstyrke for ett gjentatte samplet individ (hvor data var tilgjengelig ved forskjellige feltstyrker) som ble validert i tverrsnittsprøver.

Konklusjon: Sterke assosiasjoner med alder, aldring og den polygene risikoen for aldringsrelatert patologi ble observert i corpus callosum, lillehjernen, det limbiske systemet og ventriklene. Blant disse regionene skilte fornix seg ut som den mest fremtre-

dende regionen. Fornix-karakteristikker har tidligere blitt identifisert som biomarkører for progresjon av Alzheimers sykdom (AD). Basert på disse funnene har fornix allerede blitt brukt som en målregion for dyp hjernestimulering for AD-behandling. Vi understreket disse funnene ved å vise regionens sterke aldersassosiasjon, som kan være nyttig for å identifisere tidlige stadier av kognitiv svekkelse og nevrodegenerativ sykdom. I tillegg viste longitudinelle endringer et unikt mønster av WM-endringer som ikke bare påvirker det limbiske systemet (inkludert fornix), men presenterer også en anteriorposterior gradient av WM-endring som inkluderer WM-tap og de-differensiering i superiore frontale regioner sammenlignet med differensiering og potensielt plastisitet i lillehjerne, hjernestamme og occipitale regioner. Vi fant videre et spatielt fordelt mønster av assosiasjoner mellom polygene risikoskårer og WM. Spesielt lillehjernens peduncle var assosiert med den poligenetiske risikoen for AD. Disse genetisk informerte risikoskår-assosiasjonene med hjerne-WM var sterkere for den årlige endringen i WM enn tidspunktspesifikke/tverrsnittsskårer. Funnet understreket videre viktigheten av å bruke longitudinelle data. Den nylig testede hemisfæriske hjernealderen kan være et bidrag til presisjonsmedisin ved å vurdere forskjeller mellom venstre og høyre hjernealder. Testen av den eksisterende hjernealdersmodellen, indikerer at en del av prediksjonsfeilen kunne være drevet av forskjeller i feltstyrke. Likevel må det identifiseres flere confounders og inkorporeres i modellbygging og -bruk for å øke hjernealderens klinisk nytte.

Contents

Part I

OVERVIEW

CHAPTER 1

INTRODUCTION

The search for brain biomarkers is a promising field of research with the goal of developing feasible diagnostic tools and treatments. Hence, these efforts aim to identify specific indicators within the brain that can help identify individuals' health status, as well as diagnose, track, and treat various neurological and psychiatric conditions [134, 151, 172, 327]. Definitions for biomarkers vary. Yet, biomarkers can generally be understood in the context of the field of application such as patient care, therapeutics, and (clinical) research [47]. Brain biomarkers are measurable characteristics of the brain serving as reliable indicators of both biological and mental processes, treatment or intervention responses, and/or behavioural indices [304, 327]. These can be retrieved from different modalities such as neuroimaging, genetics, or molecular analyses, such as proteomics and metabolomics, or a combination of modalities.

Neuroimaging techniques, such as magnetic resonance imaging (MRI), provide detailed images of the brain's structure and function. The identified spatially specific anatomic and metabolic characteristics can be linked back to health and disorder expression to establish brain biomarkers. Several brain imaging derived biomarkers are commonly used in clinical contexts. For example, differential patterns of atrophy help tracking and confirming disease progression from mild cognitive impairment to Alzheimer's disease [14, 172, 260] or Parkinson's disease [92, 212], and techniques such as β-amyloid imaging combine MRI and positron emission tomography (PET) into diagnostic tools for Alzheimer's disease [134]. Other examples are the possibility to identify disease progression in stroke [113] and multiple sclerosis [98], with future applications potentially allowing to predict recovery outcomes (e.g., in stroke [181, 338]).

While significant progress has been made in the field, the search for reliable and clinically applicable brain biomarkers is an ongoing endeavour. The development of robust biomarkers has the potential to revolutionise the prognosis, diagnosis, treatment, and monitoring of neurological and psychiatric disorders, with the goal of more personalised and effective approaches. This entails the goal of improving patient functioning and societal participation by improving health, while maintaining a critical perspective of the limitations of biomarkers. In the following, we will outline how MRI biomarkers of health and disease can be understood from a lifespan perspective, which forms the baseline for this thesis. Thereafter, we provide an overview of brain ageing. Then, a summary of the project as a whole will be provided, followed by the project objectives. Finally, a short outline of the thesis will be presented.

1.1   Health and disease from a lifespan perspective

The brain matures and ages together with all other bodily systems as individuals grow older [62, 300, 329]. The changes which present during this process are complex, reaching from the molecular level to behavioural expressions [187], and are yet not fully understood. Focusing on the brain, there are various trends that have been identified comprising changes in brain structure and function, such as growing and differentiating grey and white matter, and changing connectivity during childhood early adulthood development [13, 38, 237, 265], which are then changing during the process of ageing [38, 41, 265]. While the structural and functional brain differentiation during childhood and adolescence are reflective of behavioural and cognitive changes [13, 237], also ageing-related tissue loss and metabolic changes have been associated with the observed decline in different cognitive abilities at a higher age [16, 41, 59, 66, 69, 117, 174, 249, 281, 308, 317, 330, 334].

While there are various similarities in brain tissue changes associated with ageing and/or disease [14, 260], there are also considerable differences, for example, in ageing-related white and grey matter [14, 260] and metabolic [41, 59, 317] changes. Hence, studies often differentiate brain maturation and ageing in clinical groups from control groups, where the control groups are usually ought to represent healthy development and ageing. Yet, the differentiation between healthy and diseased ageing is difficult due to its multivariate nature. Attempts to define healthy maturation and ageing, for example, focus on the general level of well-functioning and well-being, as well as the ability to adapt to environmental bio-psycho-social challenges throughout the ageing process [132, 240]. This conceptualisation of healthy ageing, being described by the absence of (major) disease in addition to behavioural aspects, serves to classify people with differing health status.

Traditionally, neuroscientific investigations on diseases simply assess contrasts in brain metrics between reference groups (often called healthy controls) and various diagnostic groups [196]. Extending this conceptualisation by linking pathology to brain development allows not only to examine neurodegenerative but also various psychiatric disorders in relationship to brain development [195, 196, 239]. In other words, developmental and ageing changes in the brain can reflect disease development and progression [195, 196, 239]. Yet, common group-based statistics comparing clinical to control groups cannot necessarily account for large biological variability within clinical (and non-clinical) groups and limitations to diagnostic labels [196]. This calls for new approaches of analysing data, for example within the normative modelling framework [195, 196, 239], using connectomics and predictive modelling [102, 166], or applications such as brain age prediction [105]. Normative models can be compared to growth charts known from pediatric practice where height is plotted as a function of age. Deviating from the normative growth by age can be diagnostic of different underlying conditions. This principal also applies to brain metrics which can be normatively modelled [195, 196, 239]. Brain age is a different way of conceptualising the idea of normative trends. Instead of using single metrics as a function of age, a marker of biological age is being estimated from various brain features [105]. In other words, brain age refers to the prediction of chronological age from a set of MRI features. These predictions can then be compared to the chronological age – large deviations

can inform about poorer general health [105]. Overall, normative models comprise an understanding of health and disease from a lifespan perspective (see Figure 1.1).
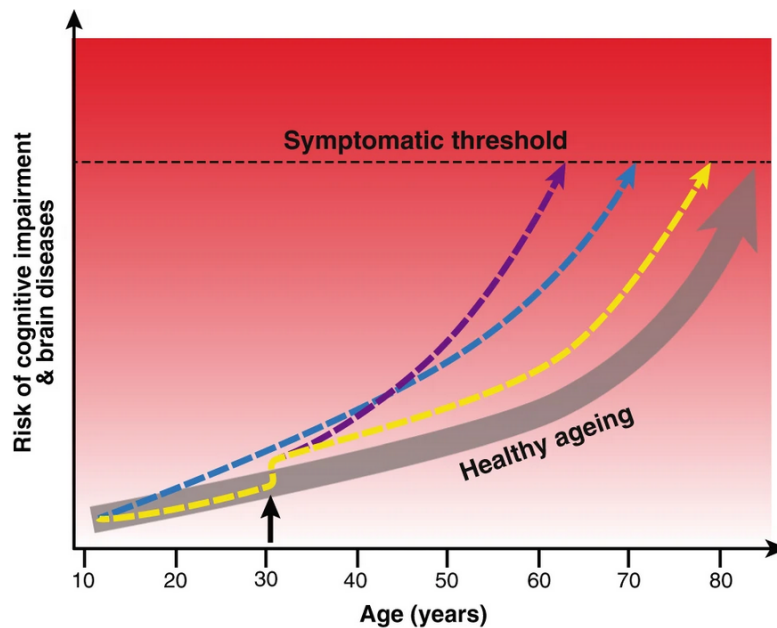
Differential trajectories of brain ageing. Illustration of the concept of ageing trajectories, specifically brain ageing. With increasing age, even healthy people are at higher risk of cognitive impairment and brain diseases, eventually reaching a threshold where symptoms appear. Individuals can differ in their brain-ageing trajectories. For example, a person may have genetic or developmental environmental factors that confer a higher rate of ageing throughout life (blue line). Alternatively, someone may experience a traumatic injury or infection in adulthood (black arrow), which results in them following an accelerated (purple line) or accentuated, but stable (yellow line), trajectory of brain ageing. While the example used here is of brain ageing, the same model can be used to conceptualise biological ageing more generally

Fig. 1.1: Re-printed from Cole et al. [62]

Establishing new models based on these approaches (while obtaining sufficient statistical power) and the models' validation require the usage of large data-sets [44, 184, 194, 327], which is made possible by recently established imaging databases and bio-banks [48, 292, 313] (for an overview see [128]). Leveraging such big data can now give unique information about the development of the human brain throughout the lifespan and how such developments related to health and disease. This can ultimately inform the definition of novel biomarkers.

Hence, understanding the intricacies of brain ageing is crucial to inform clinical practice with the ultimate goal of improving well-being and quality of life. Establishing the required knowledge demands observations on various levels. First, a baseline for healthy ageing needs to be established. This encompasses associations of brain tissue measures or representations of such with phenotypes. Second, suggestive biomarkers identified during this process need to be validated, for example across methods and samples. Third, the biomarkers can then be tested in clinical populations. Fourth, medications can be developed, tailored to the biomarkers. This thesis focused on the first two steps of attempting to establish and validate brain biomarkers.

1.2 Grey and white matter ageing

In order to better understand the process of biomarker identification as well as its limitations, it is important to outline expectable trends in brain tissue maturation and

underlying biological processes. This is particularly relevant when considering the main assumption in this project of generalisable patterns of normative ageing. Hence, this section provides a broad overview of the biological processes connected to grey and white matter ageing.

Ageing refers here to large set of changes in the adult brain reaching from the molecular over the micro- to the macrostructural level[1]. Ageing-dependent microstructural changes involve DNA damages, epigenetic and genomic instability (altered DNA methylation, acetylation, and increased mutation liability), telomere loss due to a lack of telomerase, cellular senescence, altered intercellular communication, loss of proteostasis, stem cell exhaustion, deregulated nutrient sensing, and mitochondrial dysfunction [305]. Microlevel changes include the accumulation of neurofibrillary tangles and amyloid plaques [76, 221, 296], which interrupt the cellular metabolism leading to local hypoactivation and atrophy [305]. Another micro- to mesostructural ageing feature is the shrinkage of blood vessel, which is associated with white matter lesions which can be observed at a higher rate at older ages [305]. Finally, neurotransmitter imbalances (see for review [169]) impact both brain plasticity [46] and electrical oscillations, i.e. neuronal excitability and synchronisation [305]. On the macro scale, both grey and white matter decrease in volume, whereas the proportion of cerebrospinal fluid increases [38], and WM microstructure disintegrates (see Studies A, D, and E, or [31]). Moreover, older adults show lower within but higher between network connectivity (i.e., co-varying functional MRI or blood oxygen level derived signal changes) [37, 268, 291]. Most importantly, these changes can be grossly summarised as cellular and tissue degeneration and death, in addition to vascular changes. Such changes can be captured by neuroimaging techniques such as magnetic resonance imaging (MRI), which was used in this project. Besides ageing-related grey and white matter changes, the following sections will also detail some of the pathology which can occur during tissue maturation and its specificity to these different tissue classes.

1.2.1 Grey matter ageing

The brain's grey matter contains the neuronal cell bodies. At a progressing age, neurons cease, leading to a thinner grey matter (GM) layer, as well as lower GM volume and surface area [194]. However, these changes are highly spatially distributed [260], with certain brain regions' GM being more affected than the GM in other regions.

At a higher age, the ability of neurons to re-generate is impaired, for example, examining the molecular level, due to faulty re-starts of the cell-cycle or oxidative stress, which triggers programmed cell death [165]. Programmed cell death is important for homeostasis by removing cells which have been damaged, are infected, or not loner in use [220]. There are moreover multiple possible abnormalities in the signalling cascade of programmed cell death which are apparent in neurodegenerative disorders such as apoptosis, necroptosis, pyroptosis, ferroptosis, autophagy-associated cell death, and un-programmed necrosis [220]. The result is the loss of neurons and their function on a larger scale that what would be programmed in the absence of the particular

---

[1]Microstructure refers here to all processes and structures from the cellular level to small, potentially even visually observable expressions such as lesions or finer, small region-specific tract characteristics. Macrostructure refers here to larger structures, such as regional or whole-brain morphology.

disorder. Such age-related changes in the GM can be observed in-vivo in the human using MRI. However, not all of the GM shrinkage can be explained by neuronal death. There are multiple other mechanisms involved in observations such as a thinner layer of GM in older compared to younger adults, for example, on the cellular level: cell shrinkage and dendrite degeneration [42]. Furthermore, glial (non-neuronal) cells contribute to multiple crucial functions: the maintenance of homeostasis, the mitigation of neuroinflammation, preserving the brain's immunofunction and the blood-brain barrier [42]. Dysfunction and -regulation of these cells can have systematic neuronal death as a consequence, for example through toxin-releases by astrocytes [42].

Recent technological and normative advances in the in-vivo neuroimaging community allowed to establish large data sets [128, 198, 292, 299, 314] of thousands of brain scans [194]. Such data sets are now allowing for normative modelling of healthy brain ageing, or in the general population by adding disease samples proportional to the expected population stratification, or as a reference sample in comparison to specific disease samples [195, 196, 265, 266].

To date, several studies have evidenced accelerated GM atrophy at higher ages [67, 244]. Moreover, multiple studies presented that such accelerated GM ageing is associated with different psychiatric and neurodegenerative disorders [24, 68, 104, 130], such as mild cognitive impairment and Alzheimer's disease [38, 104, 130], major depression, bipolar disorder, and schizophrenia [24, 45, 68, 103, 168, 326], borderline personality disorder [336], and sub-types and expressions of disorders such as early schizophrenia and Parkinson's disease psychosis [157]. Other conditions which cause GM death are multiple sclerosis, traumatic brain injury/trauma, and stroke, which potentially contribute to accelerated ageing processes [11, 61, 64, 126, 273, 328]. A commonly used approach to compare the age-associations of brain metrics between different groups utilises brain age predicted from GM as a normative proxy to estimate morphological deviations from reference samples, generally showing higher deviations of brain ages from chronological ages in various psychiatric and neurodegenerative disorders [104, 151, 282].

1.2.2 *White matter ageing*

The described process of neuronal death also influences the neurons' axons and hence, the brain's white matter (WM), which is the axonal network extending between neurons. In other words, although WM and GM follow different ageing trajectories [38], these trajectories are intimately connected. For example, WM decline might be detected earlier than GM decline and provide hence useful ageing marker [109].

Just as GM, WM ageing is characterised by atrophy and multiple microstructural and vascular changes, fibre tract disruption, demyelination, and an increased occurrence of inflammation [42] (see Figure 1.2 for a schematic overview of the most common WM ageing mechanisms leading to healthy-ageing associated tissue changes). At a higher age, vascular changes can occur which affect small vessels which can develop to small vessel disease (SVD) [56]. The different types of SVD are observable using MRI as WM hyperintensities, widened perivascular spaces, as well as infarcts and bleedings of different ages [56]. Importantly, SVD contributes to a large proportion of dementias and strokes [56], due to the multiple manifestations of the diseases, including endothelial
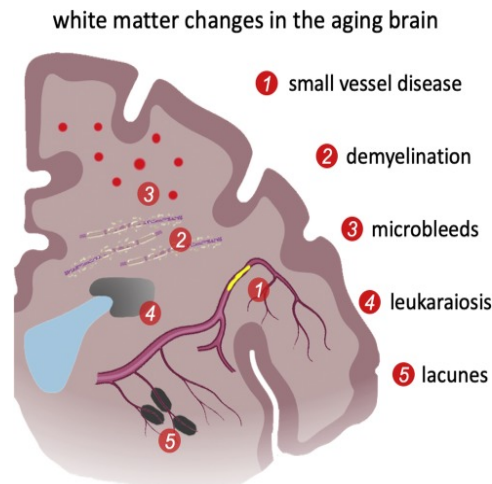
Fig. 1.2: Re-printed from Blinkouskaya et al. [42]

dysfunction (including the dysfunction of the blood–brain barrier), vasodilation impairment, vessel stiffening, impaired blood flow and perivascular fluid drainage (and resulting ischaemia), inflammation, myelin damage and loss, and secondary neurodegeneration [323]. One of the uniting features of WM ageing are cardiometabolic risk factors, which are both associated with WM integrity [32, 162] as well as myelination [303]. For an optimal neuron-to-neuron signalling through the axon, the axon is wrapped in a sheath of myelin, provided by one type of glial cells: oligodendrocytes [125, 202]. The oligodendric myelination process is furthermore regulated by another glial cell, microglia [125, 202]. Factors such as chronic hypoperfusion (reduced blood flow), increased levels of microglial toxic waste products (e.g., due to microglial dysfunction), iron toxicity as well as toxicity after prolonged excitation periods can impair oligodendrocytes and hence their ability to provide an optimal level of myelin [42, 305]. The loss of myelin has been related to systemic neurodegenerative markers such as amyloid-β depositions [76], and can be a risk factor for axonal degeneration [274]. The mentioned cerebrovascular changes, such as stiff and more narrow vessels can lead to blockages, resulting in infarcts and the development of lacunes (3-15 mm cerebrospinal fluid - filled cavities) [42, 111], leukoaraiosis (abnormal change in appearance of white matter near the lateral ventricles) [193], or tearing of the vessel, leading to microbleeds [42, 111].

Compared to grey matter investigations, less work has focused on mapping WM ageing, potentially due to the multiple aspects and details which can be examined in WM spanning from the micro to macro structural level and the many different available methods and possibilities of examining WM[2]. Past the age of 50 years, WM volume decline is accelerated, as presented by a large study [194] combining cross-sectional and longitudinal data. In addition, as detailed in Study E, additional ageing processes expressed by WM microstructure changes, which are also accelerate at higher ages. Unsurprisingly, just as GM, accelerated WM ageing (at younger ages)

---

[2]One could also argue that the lack of a unified theory of white matter ageing [259] might contribute to a compartmentalised research landscape. Moreover, the rapid technological and following methodological developments (e.g., [22, 262, 284, 319]) exacerbate establishing standards for the investigation of white matter ageing.

has been associated with various psychiatric disorders, such as schizophrenia, bipolar disorder, autism spectrum disorder, and major depressive disorder [68, 85, 164], and neurodegenerative disorders, such as Alzheimer's or Parkinson's diseases [86, 223].

1.2.3 *Studying brain ageing: A few words of caution*

Multiple factors influence both healthy and pathological brain ageing. Hence, lifespan trajectories as well as ageing processes are highly variable, even in the absence of pathology. Large variability requires large data sets in order to establish meaningful models, and potentially new approaches towards understanding the data, for example, by leveraging longitudinal data [155] or by moving back to designs resembling case-studies which follow a single or few individuals over time [114, 321]. Furthermore, biological processes do not have to present themselves as linearly related to age, and do not have to show linear ageing trajectories. This speaks to the complexity of the biological system of the brain, and the hard problem of studying it. Brain ageing can be studied at different levels, while careful consideration has to be maintained when assessing the strengths and weaknesses of the chosen approach as well as the inferences drawn from the analyses. In this sense, a normative understanding of brain tissue and its development will have its limitations while also being a useful tool to establish fundamental knowledge about what can be expectable in the majority of cases. Another consideration are the data used to model ageing processes. While most studies use cross-sectional data, longitudinal studies might be more suitable to examine ageing processes as they allow to examine the actual changes on individual and group level.

1.3 Project summary

This work has resulted in five first-author publications of the PhD candidate, several conference presentations and popular science communication efforts.

Recent developments in neuroimaging have led to establishing large-scale data bases (e.g., [292]). While data collection effort are ongoing, the current wealth of data offers new, unprecedented possibilities of analysing brain data and to identify biomarkers [51]. As described, a portion of these efforts focuses on brain age estimates [60, 62, 63], yet mostly derived from grey matter. Studies present findings for grey matter more frequently than for white matter for both brain age and more generally speaking. This might be due to the fact that there are several, less streamlined diffusion approaches which can be used to describe white matter microstructure [230]. These diffusion approaches describe radial, axial and mean diffusivity, fractional anisotropy and water fractions, kurtosis metrics, and additional information for these metrics for intra- and extra-axonal space for multi-shell approaches. Considering this as the outset position, we started this project with an investigation into age-associations of white matter microstructure metrics. The goal was to benchmark age associations across diffusion approaches for the first time in large-scale data and comparing brain age estimations between the diffusion approaches.

This lead directly to the second article, which set the white matter derived brain age estimates into the practical context of bio-psycho-social variables. While there are several articles showing associations between bio-psycho-social variables and

brain age, the majority of these articles used brain age based on T1-weighted MRI features [30, 32, 60, 105]. To extend previous studies beyond the examination bi-variate relationships, this second article used groups of variables within the domains of socio-demographics, cognitive scores, life satisfaction, and health and lifestyle to establish models explaining brain age. Additionally, not only a single brain age score was attempted to be explained by the bio-psycho-social variables, but various brain age scores (eight in total) derived from different white matter microstructure features.

While the first two articles presented brain age prediction feasibility in white matter, the practical utility of brain age requires further testing. Hence, the third article set out to test the utility and limitations of T_1-weighted brain age in a practical context. This effort was aided by two unique longitudinal, or deep imaging data sets, [234, 321, 322] following a total of four individuals over time, with N $>$ 25 scans per individual. Instead of using few data points of many individuals, we used many data points of few individuals, for a more accurate understanding of intra-individual variability, described as highly sampled, densely sampled, precision imaging or deep imaging [114, 167, 321]. One of the currently best performing convolutional neural network for brain age predictions from T_1-weighted images [171] was used to estimate brain ages. Beyond practically showing variability in such predictions across participants, we also aimed to identify acquisition and scan quality parameters which could explain such variability.

The fourth article, was motivated by an alternative approach of brain age estimations extending the commonly used single score of global brain age to several regional or local brain ages [170, 248]. However, instead of using a data-driven approach, we built upon the large field of brain asymmetry and laterality which shows various differences between the brain's hemispheres [4, 150, 158, 224, 232, 270], and estimated brain age for each hemisphere, both for metrics derived from T1-weighted, diffusion-weighted, and multimodal MRI. Additionally, the mapping of multimodal brain MRI asymmetries in larger cohorts requires further attention, as there are various potential clinical implications of brain asymmetries [159, 175, 189, 260, 275].

As new repeat-scan data was made available from the UK Biobank [198], for the fifth article (Study E), now, additional longitudinal investigations of white matter changes were possible. The objective of the study was hence to replicate previous age-associations of white matter microstructure metrics (Study A). Namely, a recent study presented evidence that the relationships between age and brain metrics can be expected to be stronger when investigated in longitudinal data [83]. Furthermore, to provide biological underpinnings of the observed longitudinal WM changes, we also assessed the polygenic risk of common psychiatric disorders and Alzheimer's disease with the annual rate of white matter microstructure change to provide additional biological explanations of which metrics in which brain areas (and their change) associate with genetic risk to develop various disorders.

1.4   Objectives and Research Questions

The general objective of the projects was to move closer to identifying bio-markers of health and ageing. This was planned by examining both diffusion-weighted and T_1-weighted MRI data separately, and in combination, with a particular focus on the

relationship between brain and age. The major research question was focused on which patterns could serve as bio-markers of health and disease. The investigations were of exploratory nature, and hence, confirmatory hypothesis testing was held to a minimum. Different statistical approaches were applied to a selection of data sets leading to further specifications of the objectives for the five executed papers:

Paper A. The overall aim of this study was to examine the associations of white matter and age. Specifically, we tested whether associations between age and white matter microstructure were concordant across diffusion approaches using direct associations between age and white matter as well as brain age predictions which were compared across diffusion approaches. The goal of this was to provide a road-map of expectable age-associations across diffusion approaches, and whether and how such associations translate to brain age predictions. A second goal was to assess, chart (age-curves) and discuss typical white matter development. The research question was hence whether there are patterns of associations between age and white matter which translate across diffusion approaches.

Paper B. The overall aim of this study was to further examine white matter derived brain age estimates. The primary investigation aimed at the proportion of brain age variance explained by sets of bio-psycho-social variables. Another question was whether the estimated white matter brain age scores would related concordant to bio-psycho-social factors or whether there would be differences between brain ages based on different diffusion approaches. Finally, the specific nature of the associations was assessed to gain a better understanding between brain age estimated from white matter micro structure and bio-psycho-social variables. The research question of this paper was therefore whether semantic phenotype combinations could explain meaningful proportions of brain age variability, and whether observed associations would be stable when the underlying data for brain age predictions were changed.

Paper C. The general aim of this study was to test brain age predictions in a stringent fashion. The first goal was to observe whether repeated brain age predictions for the same individual but different scans over time would follow the pattern observed in cross-sectional data, particularly considering the model's prediction error (i.e. increasing brain age – with a prediction error somewhat corresponding to root mean squared or mean adjusted error – with age). The second goal was to estimate to which degree the field strength and quality of the scan would influence brain age predictions. Hence, the first research question was whether brain age can be reliably predicted at multiple time-points following the same individual. The second research question was which factors would explain variability in these predictions.

Paper D. The goal of the fourth article was twofold. The first goal was to map brain-wide white and grey matter asymmetries. The second goal was to introduce a new brain age metric which was estimated only on the features from one hemisphere. We called this metric hemispheric brain age. To present the feasibility of the hemispheric brain age metric, we compared this new way of predicting age to common brain age. The presentation of the new metric's practicality was supported with phenotype-association testing in comparison to phenotype associations of regular brain age. The research questions were hence: 1) what are the observable asymmetries in the brain, and 2) is can these be extended by introducing a new brain age marker based on a single hemisphere?

Paper E. The fifth paper's goal was to map short-term longitudinal white matter changes during midlife and older adulthood as well as white matter age-associations. To gain a detailed understanding of these changes, different levels of observation were selected starting with the examination of global/whole-brain average scores, to regional and tract-level average scores, to voxel-level scores. This detailed mapping of white matter changes was supposed to provide a better understanding of white ageing across the brain and within specific regions. Additionally, imaging genetics were mapped for regional and whole brain white matter metrics. The research question was: Which are the observable longitudinal white matter changes and how can these be linked explained by polygenic risk scores of psychiatric disorders and Alzheimer's disease?

1.5   Thesis outline

The thesis consists of two main parts: an overview of the field (Part I) and the produced research articles (Part II). The overview follows the structure of a research article (Introduction, Methods, Results, Discussion).

The introduction section (Chapter 1) comprises a short prelude on brain biomarkers, followed by an account for the conceptualisation of health and disease from a lifespan perspective (Chapter 1.1) and ageing-related tissue changes (Chapter 1.2). Then, we provide a summary of the project as a whole, and, in that context, of the single studies executed in the project (Chapter 1.3). Finally, the project's research objectives are stated (Chapter 1.4).

The following chapters provide an account of the methodology across the studies (Chapter 2), with a focus on multimodal magnetic resonance imaging (Chapter 2.1), specifying T_1-weighted (Chapter 2.1.1) and diffusion-weighted MRI (Chapter 2.1.2), brain age predictions (Chapter 2.2) including the utilised machine and deep learning approaches (Chapters 2.2.1 & 2.2.2), the utilised samples (Chapter 2.3), and the statistical modelling (Chapter 2.4).

The proceeding chapter (Chapter 3) corresponds to a usual "results section" and will provide brief summaries of each of the presented articles, which will be discussed in the light of this project and the research field in general.

Finally, the results will be discussed (Chapter 4), with a focus on methodological considerations (Chapter 4.1), brain age predictions (Chapter 4.1.3), and brain regions, such as the fornix as potential ageing and disease biomarkers (Chapter 4.2.2), and an outlook for future developments (Chapter 4.4).

METHODOLOGY

2.1 Multimodal magnetic resonance imaging

Magnetic Resonance Imaging (MRI) has become one of the primary tools for detailed spatial and diagnostic examinations of the brain. The main reason is that MRI is a non-invasive[1] medical imaging technique which offers rich, high-resolution data, with superior contrast to other medical imaging techniques such as computed tomography (CT) [179]. MRI leverages the magnetic resonance effect requiring a strong magnetic field, usually 1.5-7T in hospitals and, recently, even stronger magnetic fields in research.

The atomic nucleus contains one or several protons which possess an intrinsic angular momentum also known as *spin* [144]. Spin has a clear linear relation with the nuclear magnetic moment and, consequently, it is susceptible to an external magnetic field. In the presence of an external magnetic field, the energy levels of the nucleus are split in accordance with a Zeeman effect. Splitting is dependent on the spin value as the following: $(2S + 1)$, where S is the spin value. For example, the spin of the hydrogen's nucleus (or simply proton) is $1/2$ and energy splitting means there are just two energy levels: a minimum energy level and an excited energy level. The minimum of energy corresponds to the alignment of spin along the external magnetic field, and the excited energy level is presented by spin aligned to an opposite direction, respectively [36, 144, 203].

In the MRI scanner, protons' spins are axially aligned to the external magnetic field which creates a magnetic vector with its orientation along the MRI external magnetic field, while protons still spin at a low-energy state [36]. An external radio frequency pulse with the help of transmitter coils changes the protons' energy state by a transferring spin on the higher energy level. The frequency of absorption energy can be described by the Larmor frequency allowing to manipulate the nuclear magnetic moment as well.

The amount of emitted energy and, hence, the signal depend on the imaged tissue. As the human body consists to 75-80% of water (H_2O) [203], hydrogen (most common isotope: ^1H) is the most abounded nucleus in the body, with all tissue containing hydrogen. However, the hydrogen density depends on the tissue type which, in turn, influences relaxation times. As already known, T_1 relaxation time is defined by a spin-lattice interaction and T_2 relaxation time is a result of spin-spin interaction. Thus, fluids have usually long, water-based tissue and medium long relaxation times,

---

[1]For advantages and safety of MRI see Appendix 2.

whereas fat-based tissue has short T_1 relaxation times [203]. The T_2 relaxation time is shorter than T_1 relaxation time and follows the same order as T_1 relaxation time in respect to tissue type: fluids longest, fat-based tissue medium, water-based tissue shortest, respectively [203]. These differences allow one to describe different tissue types from the recorded signal.

Using multiple RF and magnetic gradient pulses in sequence (i.e., a pulse sequence) allows one to emphasise or suppress certain aspects of the resulting signal by considering tissue relaxation times in addition to the proton density (see Appendix 1 for more information on pulse sequences) [36, 203]. The contrast of an image generally depends on either proton density or relaxation time. Considering several pulse sequences and, hence contrasts is called multimodal MRI. In the following, we will further describe the two acquisition types utilised in this study for multimodal MRI: T_1 and diffusion-weighted MRI, and how the acquired data was used.

2.1.1 T_1-weighted MRI

T_1-weighted brain imaging is one of the most frequently used acquisition types in research acquisitions. This is potentially due to T_1-weighted MRI being part of the standard clinical workflow. The advantage of T_1-weighted MRI is that it highlights anatomical structures and provides a clear contrast between grey and white matter. Additionally, there is a wealth of processing pipelines [94, 95, 99, 141, 242] openly available, allowing standardised low-code, simplified workflows including quality assessments, motion correction, skull stripping, normalisation, segmentation, and the estimation of various statistics, such as grey matter thickness, surface area and volume (see for an example of such pipeline Figure 2.1). Such standardised pipelines increase comparability across studies, as there are many possible degrees of freedom in processing images which ultimately lead to variability in study results [39, 43, 344].
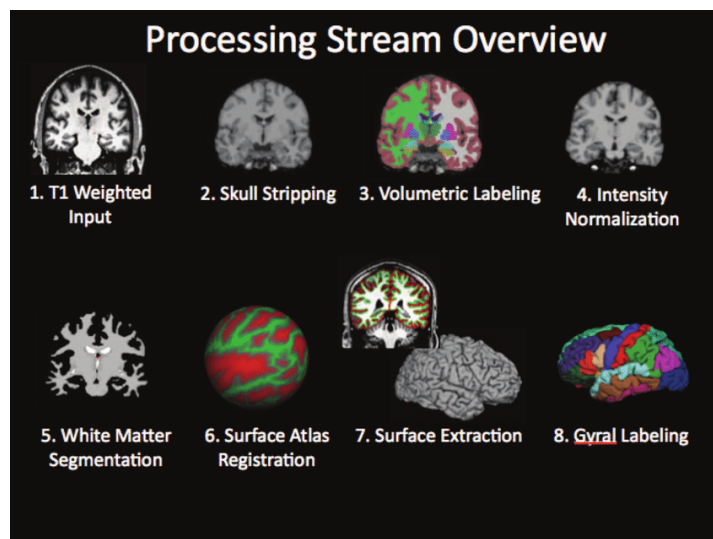
Fig. 2.1: A simplified depiction of the Freesurfer recon-all pipeline. Re-printed from Grossner et al. [115]

First, we used the raw voxel-level data as provided from the scanner and processed it minimally in Study C. A minimal processing approach is useful as it circumvents

relatively long processing times allowing an easier inclusion into clinical workflows than processing pipelines which take hours before predictions are even started. Minimally processing in the case of Study C refers to skull stripping, bringing the brains into the right coordinate space, and cropping the image (which was done to remove unnecessary computation of empty space in the brain images).

For Study D we used a more "common" approach of cortical reconstruction and estimation of grey matter thickness, volume, and surface area using the Freesurfer [99] recon-all pipeline (see Figure 2.1). Outliers were identified by using Euler numbers, a measure of the topological complexity of reconstructed cortical surfaces [264]. Deviations from the mean Euler numbers were used to assess the quality of the reconstruction, leading to exclusions for deviations above three standard deviations from the mean.
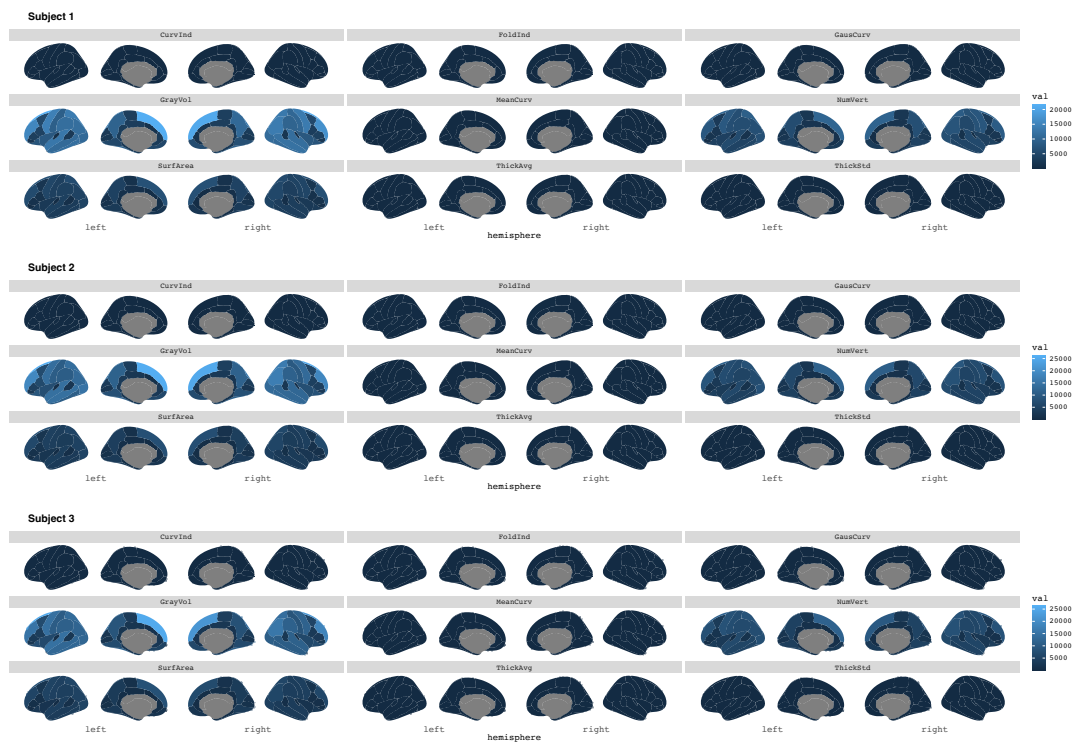
Fig. 2.2: An example of cortical parcellation using the Desikan-Killiany Atlas in Freesurfer, here applied to average across sessions in the BBSC data set to extract various gray matter metrics (curvature, folding, volume, surface area, and thickness) [321, 322]

The next step included a common approach to reduce data complexity and increase biological meaning: parcellation. As the brain is defined by heterogeneity in structure and functional networks, defining distinct partitions aids to not only to establish a logical topography, but also to execute topology-informed research [89]. We used the Desikan-Killiany Atlas [77] as a parcellation scheme to average over grey matter features. For an example of a parcellation scheme on three subjects see the average cortical values of the three BBSC subjects (Study C) across their sessions in Figure 2.2.

2.1.2   *Diffusion-weighted MRI*

Diffusion-weighted MRI is another commonly used MRI technique both in clinical and research settings. The thermal random motion of some small insoluble particles in different media can be described by Brownian motion [219]. Importantly, particles are limited in the distance they are able to travel in a certain time interval which allows to describes particle diffusivity by a probability density function [90].

Due to the described high prevalence of water molecules across the human body (Section 2.1), diffusion MRI captures the random *movement* of water molecules throughout tissue. In order to being able to image diffusion, the magnetic field strength needs to be varied linearly by a pulsed field gradient. Diffusion contrasts are similar to inverse T_2-weighted contrasts, as watery tissues have more mobile molecules which produces a lower signal intensity. Opposing, the signal is stronger in tissue characterised by more solid and static tissue [203].

Most commonly used diffusion sequences are pulsed gradient spin echo (PGSE) [203]. PGSE consists of a pair of spin-echo RF pulses with large, equal gradients on either side of the 180° RF pulse (i.e. opposing directions) [203]. Manipulating the timing between gradient pulses (Δ), and pulse characteristics (amplitude (G) and length/duration (δ)) allows to control the degree of the *b*-factor weighting in PGSE sequences (with *b*-factor units in s·mm^{-2}):

$$b = \gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3}). \tag{2.1}$$

Water diffusivity in white matter is restricted by cellular structures such as neuronal fibres or tracts. These restrictions are reflected in the signal. Using biophysical models/approaches allows one to identify characteristics of cellular structures making diffusion MRI extremely useful for white matter examinations. [229]. Hence, the work within this project focused on white matter when using diffusion MRI (except from Study C using T_1-weighted MRI).

2.1.2.1   *Biophysical models*

Biophysical approaches of diffusion refer to the modelling of diffusion along recalling fibres mathematically, by making different assumptions about the water diffusion in the axons and the surrounding space [229]. The resulting metrics allow then for further inference on the structural integrity and potential health of examined tissue. The most commonly applied model is diffusion tensor imaging (DTI) [27].

Biophysical models can be grouped into two proposed approaches: signal representations and tissue models [138]. Each approach allows to extract differential information about tissue microstructure from diffusion MRI. Signal representation (or statistical models) describe the behaviour of the signal in a given voxel without making assumptions about the underlying tissue. Hence, such models' estimations lack specificity, rendering the estimates as indirect measures of microstructure characteristics. This limitation is filled by tissue models, which make assumptions about tissue geometry. Hence, tissue model parameters might be also be more specific and biologically-relevant, assuming that the modelled tissue features (e.g., geometry) are

accurately captured [138].

As the different diffusion approaches highlight different features of the signal or geometry, and hence underlying biology, it is yet important to gain a better and comparative understanding of such approaches. In this project, we presented comprehensive comparisons of both signal representation approaches DTI and DKI, as well as biophysical models: (1) the spherical mean technique (SMT) [149], (2) SMT's multicompartment extension SMTmc, (3) the DKI multicompartment extension White Matter Tract Integrity (WMTI) [97] , and (4) the Bayesian Rotationally Invariant Approach (BRIA; see Appendix 3 for all included metrics) in the context of age and phenotype relationships.

See Table 2.1 for an overview of key features of the utilised diffusion approaches, and Appendix 3 for the metrics included in each approach. While the observed multi-shell[2] diffusion models (BRIA, SMTmc, WMTI) differentiate intra-axonal from extra-axonal space [97, 148, 253], BRIA delivers the most detailed account by providing metrics on water and cerobrospinal fluid fractions [253].

Diffusion approach	Fundamental assumption
Bayesian Rotationally Invariant Approach (BRIA)[253]	Bayesian estimation of diffusion in several compartments: intra & extra axonal & water
Diffusion Kurtosis Imaging (DKI)[142]	DTI & kurtosis tensor (non-Gaussian)
Diffusion Tensor Imaging (DTI) [27]	tensor* fitted across each voxel
Spherical Mean Technique (SMT)[149]	Spherical mean value of the diffusion signal over gradient directions
Multi-compartment Spherical Mean Technique (SMTmc)[148]	Double b-shell extension of SMT
White Matter Tract Integrity (WMTI)[97]	Based on kurtosis model, modelling axon collections as a Gaussian compartment in contrast to extra axonal diffusivity
	*comprised of vectors describing diffusivity into each 3 directions

Table 2.1: The utilised diffusion approaches and their fundamental assumptions.

In the following, we will first describe signal representation approaches and then tissue models, both in general and then briefly describing the utilised diffusion approaches (BRIA, DKI, DTI, SMT, SMTmc, WMTI).

2.1.2.2 *Signal representation models*

The most common signal representation is the expansion of the logarithm of the signal in polynomials up to a given order in b (see Eq. 2.1) [138]:

$$\ln(\frac{S}{S_0}) = -bD + \frac{1}{6}(bD)^2 K + ...,\tag{2.2}$$

where D is the diffusion coefficient and K indicates the kurtosis (of the diffusion density probability function), whereas S is the signal at echo time and S_0 the signal at zero echo time.

Diffusion tensor imaging (DTI) entails the expansion of this function up to the first order of b, valid for $b \ll \frac{1}{DK}$. When this assumption is met, DTI assumes tissues' diffusion to be *near*-identical to a Gaussian probability function [138]. Diffusion kurtosis imaging (DKI) [142] can be used to model higher b-values which then allows for a more accurate tissue characterisation [138].

---

[2]The number of shell refers to the number of unique b values which are not b = 0.

DTI refers to fitting a second order tensor for each voxel of the MRI volume. These tensors can be described by a 3-by-3 matrix indicating the 3 directions of possible diffusion or eigenvectors and their eigenvalues λ (Figure 2.3). The tensor is often described with the analogy of a ball and stick, where the ball describes isotropic diffusion, and the stick anisotropic diffusion. Isotropic refers to unrestricted diffusion observable in water (i.e., no cellular boundaries) and anisotropic diffusion the opposite, i.e., diffusion observed in clearly restricted cellular space, such as white matter fibres. The code for the DKI estimations can be found at `https://github.com/NYU-DiffusionMRI/DESIGNER`, and DTI metrics can be estimated during TBSS procedure in FSL (see Section 2.1.2.4).

Isotropic voxel
$\lambda_1 \simeq \lambda_2 \simeq \lambda_3$

Anisotropic voxel
$\lambda_1 > \lambda_2 , \lambda_3$
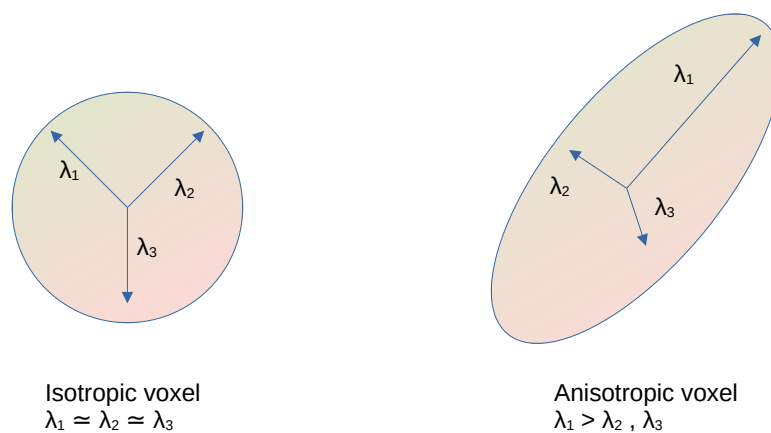
Fig. 2.3: Example of an isotropic tensor (left) and an anisotropy tensor (right).

The DTI eigenvalues can be presented in a form of rotational invariants and used as scalar metrics. For example, the highest eigenvalue describes the axial diffusivity (AD), the average of the width of the tensor as radial diffusivity (RD), the mean (or trace of matrix) of all three eigenvalues as mean diffusivity, and the squared eigenvalue differences by their squared sum as fractional anisotropy, which refers to the level of (an)isotropy (Figure 2.4).

While DTI provides useful metrics of anisotropy (FA) and diffusivity (MD, RD, AD), DTI assumes that diffusivity is distributed nearly perfectly in accordance with the Gaussian density function. DKI does not assume near-Gaussian diffusivity as expressed in axial, radial and mean kurtosis (AK, RK, MK). Both approaches (DTI and DKI) do not separate intra and extra-axonal spaces. Hence, they rather map mesostructural features than microstructure. Furthermore, the approaches cannot account for crossing fibres, which are a commonly occurring phenomenon in the brain's white matter [31, 229].

Such shortcomings have been addressed by different more advanced and potentially biophysically more informative models. These models make different assumptions about the tissue compartments (hence tissue models) and are dependent on the underlying data. Diffusion MRI data can be acquired using a single or multiple shells.
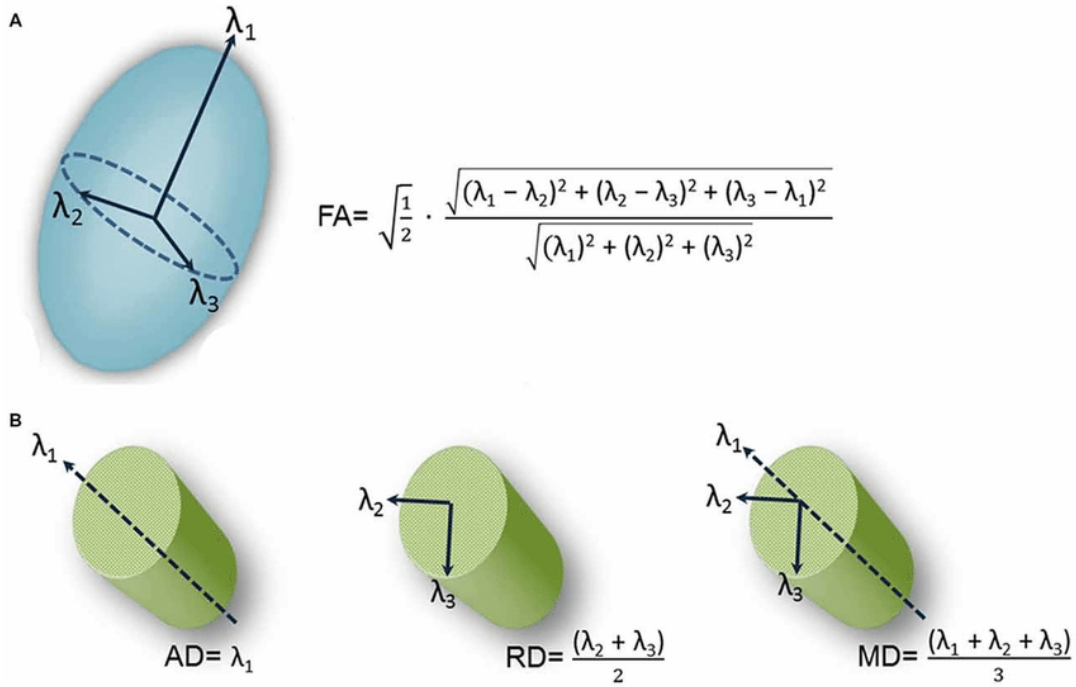
Fig. 2.4: Diffusion Tensor Imaging (DTI) metrics. Reprint from DeSouza et al. [78].

Single shell refers to the acquisition with a single unique b-value (see Equation 2.1), whereas multi-shell acquisition refers to the acquisitions with several unique b-values, for example $b_0 = 0\frac{ms}{\mu m^2}$, $b_1 = 0.5\frac{ms}{\mu m^2}$, $b_2 = 1\frac{ms}{\mu m^2}$, $b_3 = 2\frac{ms}{\mu m^2}$. Multi-shell acquisitions allow for more complex models taking into account multiple compartments (e.g., tissue and water fractions) and directions of diffusion in a voxel.

2.1.2.3 Tissue models

Biophysical models usually make use of three compartments: (1) intra-axonal space described by infinitely long sticks modelled by the orientation distribution function (ODF), (2) extra-axonal space modeled as Gaussian anisotropic, and (3) freely diffusing water modelled as Gaussian isotropic [138].

The spherical mean technique (SMT) and **SMTmc** (multicompartment SMT) address model degeneration problems of the standard diffusion model (while using standard diffusion sequence protocols [231, 254]) by applying a powder averaging technique [149]. Such technique can be applied to both single and mulicompartment models [148]. While such averaging technique has been used in Neurite Orientation Dispersion and Density Imaging (NODDI), compared to SMT and SMTmc, NODDI is sensitive to noise due to the effects of local minima and limited by poor precision of its non-linear fitting choices [140], affected by the numeric algorithm used [70, 341], as well as post-processing [200]. This renders the resulting NODDI metrics as unstable. Hence, for SMT (1) the assumptions of the tortuosity model about diffusivity were added [294], as well as the assumption that (2) intra- and extra-axonal axial diffusivity is equal, and that (3) radial inta-axonal diffusion is equal to zero, representing intra-axonal diffusivity as a stick. For SMT and SMTmc output metrics see Supplement 3.

The original code for the SMT estimations can be found at https://github.com/

`ekaden/smt`

White matter tract integrity (WMTI) utilizes the standard diffusion model but differentiates intra-axonal from extra-axonal space [97]. As in SMT, the intra-axonal space is assumed to be isotropic and hence modelled as stick, with a radial diffusivity equal to zero. The extra-axonal space assumes anisotropic diffusion (ball-shaped diffusion, equal towards all directions). Moreover, the intra- and extra-axonal compartments are assumed to be impermeable. Additional biophysical assumptions contained in the model are that (1) the intra-axonal space consists of myelinated axons, where the signal is not influenced by this myelin proportion due to the quicker relaxation time in myelin/fat. (2) As in DTI, Gaussian diffusion tensors are applied, however separating between intra- and extra-axonal spaces [97, 139]. (3) Diffusivity in the extra-axonal space (glial cells) is stated to be faster than in the intra-axonal space (neurons) to avoid statistical degeneracy. While this assumption aids simplicity, it does not correspond to the underlying biology [200]. WMTI estimations allow for dispersion of up to 30°, enabling WMTI estimates to account for coherent or parallel fibres. However, the approach is not valid for crossing fibres or high orientation dispersion [200], and differences in diffusivity pace have not been confirmed [81, 143].

The original code for the WMTI estimations can be found at `https://github.com/NYU-DiffusionMRI/DESIGNER`. For WMTI output metrics see Supplement 3.

The Bayesian rotationally invariant approach (BRIA) assumes different Gaussian compartments separating intra- and extra-axonal space. Extending the approches presented above, beyond diffusivity and FA, BRIA allows also to estimate additional metrics informing about intra-axonal and extra-axonal cerebrospinal fluid and water fractions [253]. Additionally, BRIA was informed by both scalar and tensor parameters to identify potential degeneracies in the scalar metrics [231]. An attempt to address model degeneracies is to not only account for local minima, but to select the non-degenerated branch of linear estimated moments when model fitting [230]. However, model degeneracy and hence precision remain challenges [231]. Yet, BRIA is a step towards better teasing apart meso-structure from micro-structure [231, 253]. Practically, Reisert and collegues [253] used supervised machine learning, informed by a Bayesian estimator, to estimate microstructural properties. The Bayesian estimator was established by using uniform distributions as priors which were then updated in the posterior distribution (according to Bayes' theorem [29]). From the posterior distributions, expectation values (not the most probable value) were used, using Monte Carlo integration [258], resulting in the training data [253].

The original Matlab code for the BRIA estimations can be found at `https://bitbucket.org/reisert/baydiff/src/master/`. For BRIA output metrics see Supplement 3.

Research validating the presented approaches is still limited and requires further attention. Additionally, the interpretation of diffusion MRI derived metrics is not necessarily straight forward and limited by modelling assumption [138, 230]. Hence, it is useful to validate the different approaches against each other and connect them to biological processes such as ageing [31].

2.1.2.4 Tract-based spatial statistics (TBSS)

After voxel-level scalar metrics are obtained from each diffusion approach, tract-based spatial statistics (TBSS) were used in order to prepare the data for and conducting statistical analysis. TBSS are a suite of tools for analysing diffusion data [288] which can be used following the subsequent steps for data preparation. First, volumes are being aligned based on a template. Here, we use the FMRI58_FA template from FSL [289] using non-linear transformation as implement in FNIRT [289]. Second, using the described tensor-fitting approach from DTI, an FA image is established for each data set/participant, averaged across subjects, and finally thinned to create a mean FA skeleton. Third, each participant's FA values are projected onto the mean skeleton by filling the skeleton with FA values from the nearest relevant tract centre. This skeleton-based analysis minimises confounding effects due to partial volume effects and residual misalignments, originating from non-linear spatial transformations (step two). Fourth, the skeleton can be used as a mask to average the metrics obtained from the different biophysical approaches (a) over the entire skeleton, (b) single hemispheres, or (c) pre-defined regions by parcellation schemes, such as defined by the Johns Hopkins white matter atlas [129, 217].

Prior to analyses of the obtained voxel-level as well as region-wise and global averages, we used an optimised quality control pipeline [199], which contains the following general steps: (1) artefacts, eddy current, and motion corrections, (2) creation of the fractional anisotropy mean skeleton across subjects (where fractional anisotropy is a common DTI metric indicative of the isotropy), (3) registration of the subjects to the FA skeleton, (4) application of biophysical models and statistics on voxel-level or region-averaged data [199].

2.2 Brain age predictions

Brain age refers to the biological age estimated by predicting a person's chronological age off a set of brain features. Brain age involves usually a machine learning model trained on different MRI brain features to determine the biological age [105]. The set of features can incorporate different levels of analysis (pixel- or voxel-level values [63, 171, 248, 297] (see also Study C or atlas-defined regional and/or whole-brain averaged metrics [31]) (see also Studies A and D), and modalities, such as diffusion, T_1-weighted and functional MRI [60, 73, 261]. Models trained on region averaged diffusion features outperform models trained on T_1-weighted region averaged features in predicting age, which in turn outperform models trained on functional MRI features; with the yet best predictions obtained from multimodal models [60, 73, 261, 269] (also shown in Study D).

There are various possibilities of choosing levels of analysis – from raw pixel- or voxel-level data to values averaged over atlas-defined regions – which also make different analysis tools more or less appropriate. For example, previous findings showed promising model performance on raw or minimally processed 2D or 3D data are obtained by deep learning methods convolutional neural networks (CNNs) [63, 171, 248, 297], whereas tabular data is potentially better predicted by linear and ensemble algorithms [120, 145, 285]. For an example comparison of three commonly

Fig. 2.5: Relationship of T$_1$-weighted MRI brain age and chronological age across algorithms, including extreme gradient boosting (XGBoost) from region-averaged data, Gaussian Processes regression (brainageR), and a convolutional neural network (DeepBrainNet). brainageR and DeepBrainNet utilised 2D minimally processed MRI data. Reprint from Bacas et al. [18].

used algorithms see Figure 2.5.
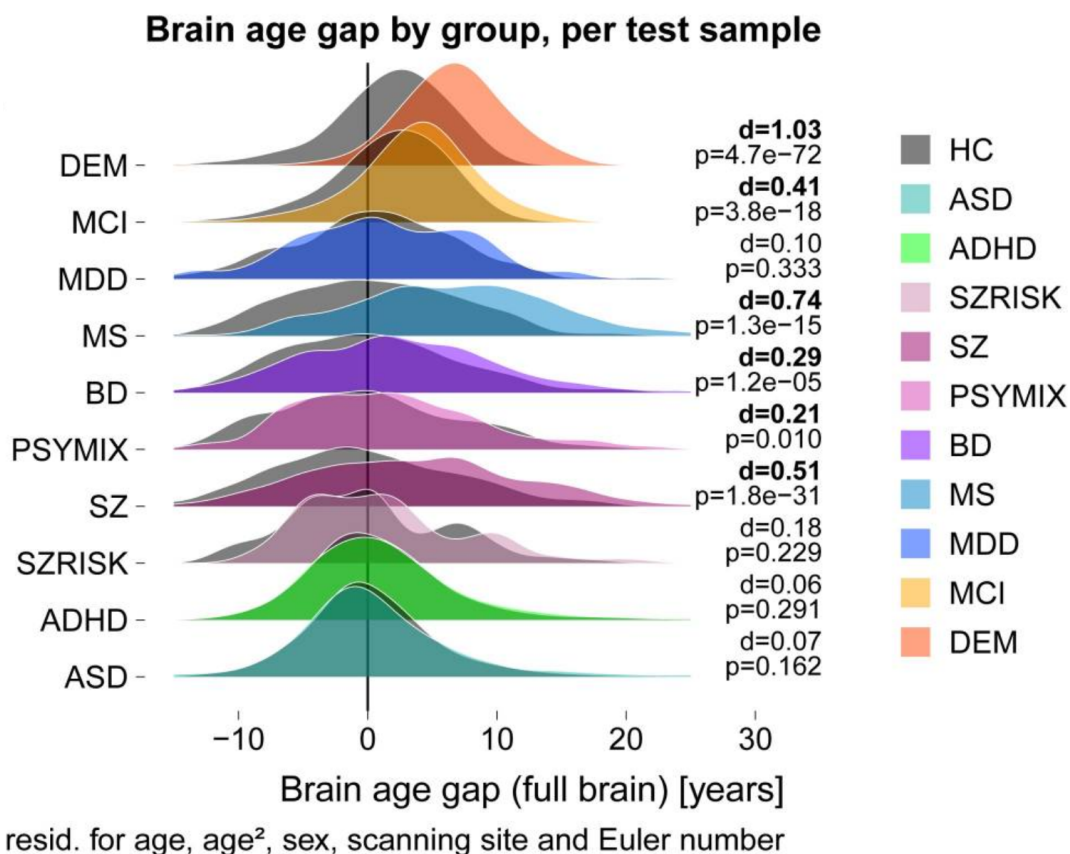
Fig. 2.6: Brain age (adjusted for covariates of no interest) across common brain disorders in comparison to healthy controls. Reprint from Kaufmann et al. [151].

While the choice of algorithms still needs benchmarking, one study [145] lines out some trade-offs connected to the choice of brain age prediction algorithm: XGBoost's predictions were more accurate overall and more reliable predictions when using lower quality scans, while a 2D convolutional neural network's predicted age was

more sensitive to associations with cognitive functioning in several samples aged 8-22. Another study [18] presents highly correlated predictions from XGBoost and CNNs (with XGB predictions being slightly less accurate), but XGBoost brain age being more sensitive to detecting cognitive impairment in a sample aged 19-100. Hence, further investigations applying different algorithms is required for a better understanding of brain age in general and it's estimation [145].

Furthermore, brain age differences can be expected for various groups, especially when manifestations within the brain are strong. For example when comparing dementia patients' brain age to healthy controls' brain ages (see Figure 2.6).

The following sub-sections will briefly explain the utilised algorithms across the studies included in this thesis: (a) eXtreme Gradient Boosting (XGBoost; as an example of ensemble algorithms), and (b) convolutional neural networks (CNNs; as an example of deep learning techniques), and present previous research on their utilisation for brain age prediction. Finally, another crucial topic when conducting brain age will be introduced: (c) age-bias correction procedures.

2.2.1 *Brain age predictions using eXtreme Gradient Boosting*

EXtreme Gradient Boosting (XGBoost) is a machine learning algorithm combining decision trees with regularisation, and can be used for regression and classification problems [52]. Focusing on regression problems, such as brain age predictions, different from other tree-based algorithms, XGBoost uses a special form of regression trees [226]. The tree starts with a regularisation function of the similarity score (SS), calculated from the sum of the residuals (R), being the difference between observed and predicted values, and number of the residuals N, in addition to the regularisation parameter λ (intending to reduce the predictions sensitivity to individual observations):

$$SS = \frac{\Sigma R^2}{N + \lambda} \tag{2.3}$$

Based on SS, the sample is then split into different groups (i.e., two leafs per split), serving to calculate the gain G from each leaf's similarity score SS_L and SS_R using Eq. 2.3:

$$G = SS_R + SS_L - \sqrt[2]{SS} \tag{2.4}$$

The G is then used to decide which further threshold to use to divide the sample, and hence, which characteristics following leaves should have. G is also used to prune the tree by subtracting the pruning parameter γ from the respective G value. Branches with negative results are pruned/removed.

The overall equation for numeric Predictions includes the starting or smallest value A, the Learning Rate (η) as another hyperparameter and an outcome value estimated similar to the SS, but using only the sum of residuals without squaring them:

$$Prediction = A + \eta \times \frac{\Sigma R}{N + \lambda} \tag{2.5}$$

Predictions can be made for (1) binary nominal outcomes (classifications) by restricting the values for example to an interval between 0 and 1 and rounding up or down to

obtain classifications, and (2) for continuous values such as brain age, directly using the obtained `Prediciton`.

XGBoost has been found to be a highly efficient and precise alternative for brain age predictions [12, 30, 31, 75] (see also Studies A, B, D). Although predictions depend rather on the underlying data (e.g., acquisition differences, age-distribution) and training procedure than the algorithm [20, 145], XGBoost can outperform deep learning methods on tabular data problems [285], and performs well on brain age predictions compared to other machine learning algorithms [214]. Yet, the findings of different studies comparing brain age predictions of various algorithms show variations in results, which does not allow to answer the question which algorithm could be the "best" to predict age. For example, one study [20] examined the prediction performance of different algorithms (support vector regression, relevance vector regression, Gaussian vector regression) revealing only small differences. This suggests that the training data have greater impact on model performance than the selected algorithm. Similarly, when using relative small sets of volumetric and surface area features (N = 152), the performance of a wide range of algorithms converges [120], and deep neural networks potentially performing slightly better than random forest, support vector machine, and least absolute shrinkage and selection operator (LASSO) regression [186]. When increasing the number of features (N = 800), using a large training sample size (N = 2105), eXtreme Gradient Boosting was among the best performing algorithms for brain age predictions, in another study [214].

2.2.2 *Brain age predictions using Convolutional Neural Networks*

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm allowing to fit complicated curves to complex data [6]. A CNN consists of nodes and connections between notes, which are analogues to intercept and slope values in simple regression problems (see Figure 2.7 for a simple depiction). The type of lines fitted to the data are dependent of the used activation function(s) producing e.g. curved lines, such as softplus [342] or ReLu [88]. CNNs usually have several input nodes (in Figure 2.7 only a single input node on the left) and output nodes (in Figure 2.7 only a single output node to the right) [177]. Additionally, there are several nodes in-between input and output nodes called hidden layers. CNNs fit curves to the data based on estimated parameters which are both multiplied (weights) and added to input values (called bias) across nodes (Figure 2.7). New curves are estimated by combining the various curves across layers to finally solve regression or classification problems [177].

Back-propagation is the central method to estimate weights and biases [343]. In short, back-propagation uses the chain rule [271] to estimate derivatives (see Eq. 2.6 for Leibnitz' notation), on which gradient descent is applied [10, 343].

The chain rule expresses the derivative of the composition of two differentiable functions. For example, if a variable z depends on the variable y, it is also dependent on x, in case y depends on x:

$$\frac{dz}{dt} = \frac{dz}{dx} \times \frac{dx}{dt}. \tag{2.6}$$

Gradient descent refers to taking repeated steps in the opposite direction of the gradient in order to reach a local minimum as fast as possible to minimize the loss function [10].
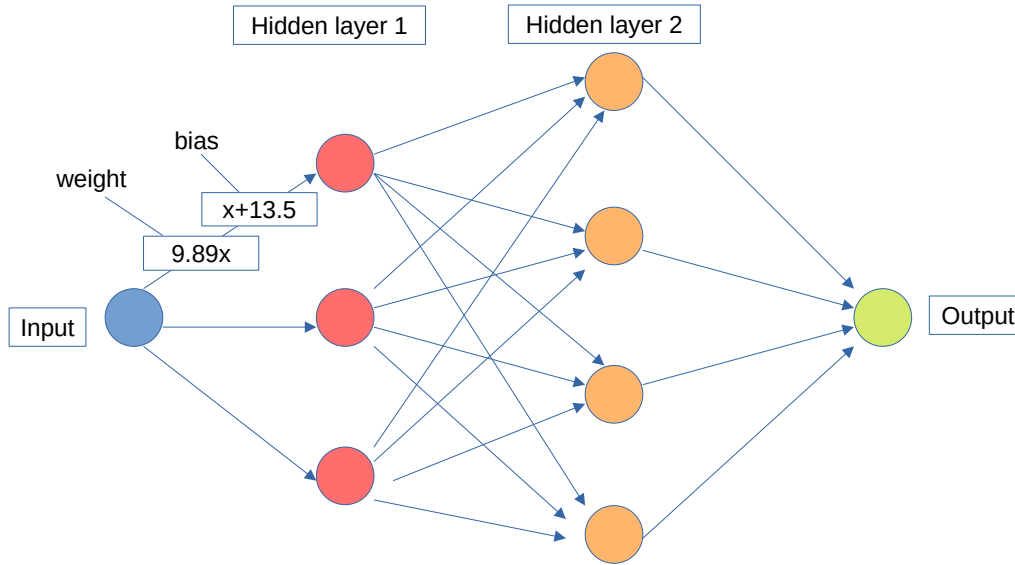
Fig. 2.7: The architecture of a simple convolutional neural network with two hidden layers showing the weight and bias of a single connection between nodes.

In other words, backpropagation refers to gradient descent based on the chain rule, by finding the derivative of the sum of squared residuals (SSR) of n values with an index of i observations with respect to the bias and weights. SSR is hence a simple function of the difference of observed (O) and predicted (P) values.

$$SSR = \sum_{i}^{n} (O_i - P_i)^2. \tag{2.7}$$

This results in the following equation (adding the derivative of the Bias (B)):

$$\frac{dSSR}{dB} = \frac{dSSR}{dP} \times \frac{dP}{dB}, \tag{2.8}$$

for which the resulting values can be tuned with the learning rate η:

$$StepSize = \frac{SSR}{d} \times \eta, \tag{2.9}$$

until the optimal value for the bias (or weight) is identified. A feature which is specific to CNNs are convolutions, which are filters applied to an input, such as blurring, sharpening or edge-detection filters. Filters are iteratively applied using different configurations (e.g., kernel size and position), which results in activation or feature maps. These maps are indicative of certain features of the input, such as cerebrospinal fluid in 3D MRI data, and can be used to inform weights and biases.

There are several studies showing good performance of CNNs on brain age predictions, using either 2D [145, 297] or 3D MRI data [63, 171, 248, 248, 297], with particularly good performance shown in [171]. This CNN [171] was used for brain age predictions in Study C, yet using only the regression output (Figure 2.8), as this algorithm was found to perform best across utilised methods (see [171]). As
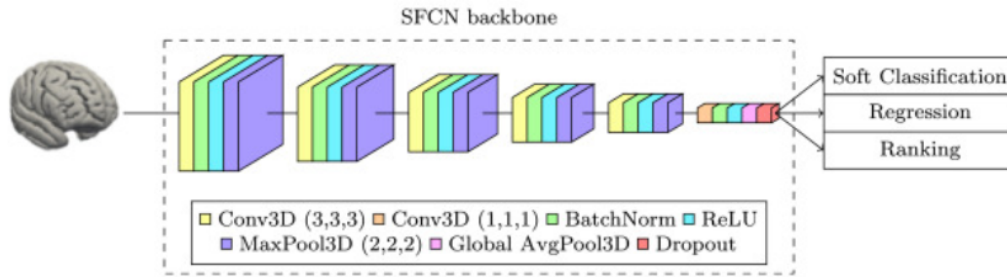
Fig. 2.8: The architecture of the utilised convolutional neural network for brain age predictions in Study C. Adapted from Leonardsen et al. [171].

visible in Figure 2.8, the model is composed of various compositions of layers with multiple activation functions. While this model performed well in cross-sectional cohort data [171], also tested independently and compared to other trained models [84] there are various steps necessary to improve CNNs (for brain age predictions and in general) [297]. Potential next steps include setting benchmarks for deep learning architectures, integrating more multimodal data, transfer learning, and *especially* increasing explainability [297].

2.2.3 *Age-bias corrections*

Once the predictions are obtained, usually the difference between the prediction and age, called the brain age gap or brain age delta is the variable of interest, for example delineating group differences between clinical and control groups [21, 151]. However, this metric is dependent on several characteristics of the training data [145], calling for adjusting brain age models or predictions. The most apparent adjustment is age adjustment, as the distribution of the used training sample influences the brain age predictions [74]. This affects particularly predictions made at the extreme ends of the age distribution, i.e. the youngest and oldest participants in the sample. Yet, there are several ways to account for such bias [74].

Age bias corrections used in the research within this project (Eq. 2.10-2.11) were only executed when directly relating predicted brain age and chronological age (i.e., relationships between the label and different predictions) or the brain age gaps which were based on different predictions. For the age-bias corrections we followed the procedure suggested in de Lange et al., [75] which is equivalent to other commonly used brain age correction methods [74].

The procedure uses the intercept α and slope β when regressing age from brain age (BA) using a linear model:

$$\hat{BA} = \alpha + \beta \times Age, \tag{2.10}$$

to correct the brain age estimates (BAc) with the model coefficients and age:

$$BAc = BA + (Age - (\alpha + \beta \times Age)). \tag{2.11}$$

Otherwise, age could simply be regressed from brain age (Eq. 2.12), allowing more interpretable results of associations when examining un-standardised slopes (β)

compared to using corrected brain age gap values. Equation 2.12 shows an example of examining the association between brain age (BA) and a phenotype of choice (P) while adjusting for age, sex, the age-sex interaction and using scanner site (Site) as a random intercept (RI).

$$\hat{BA} = \beta_0 + \beta_1 \times P + \beta_2 \times Age + \beta_3 \times Sex + \beta_4 \times Sex \times Age + RI(Site). \qquad (2.12)$$

2.3 Examined samples and data

We used five different samples in the different studies (Table 2.2) containing nearly $N \approx 50,000$ MRI acquisitions of diffusion MRI (dMRI) and/or T_1-weigthed (T_1). Most of the executed studies contained analyses of UK Biobank (UKB) data only (Studies A, B, D, E). Study C used a combination of densely sampled / longitudinal data from the Bergen Breakfast Scanning Club (BBSC) and the Frequently Travelling Human Phantom (FTHP), and cross-sectional validation data from the Norwegian Cognitive NeuroGenetics sample (NCNG) and local Oslo sample (TOP).

Table 2.2: Overview of the samples used across studies

Study	Design	Modality	Sample
Study A	cross-sectional	dMRI	UKB: $N = 35,749$
Study B	cross-sectional	dMRI	UKB: $N = 35,749$
Study C	longitudinal	T_1	BBSC & FTHP: $N_{total} = 4$
Study C	cross-sectional	T_1	NCNG & TOP: $N_{total} = 1,065$
Study D	cross-sectional	dMRI & T_1	UKB: $N = 48,040$
Study E	longitudinal	dMRI	UKB: $N = 2,944$

2.3.1 UK Biobank

The UK Biobank is a unique data source comprising the data of over 500,000 people [198, 292]. This prospective study focuses on midlife to old ages (inclusion age range: 40-69 years) and contains a large range of phenotype and genotype information. Among the imaging modalities, there are currently brain MRI scans available for around 10% of the total sample ($N \approx 50,000$), with the goal of scanning around 20% of the total sample ($N = 100,000$) [7, 213].

Additionally, follow-up scans are already available for around 1% of the total sample ($N \approx 5,000$) [7, 198, 213, 213]. MRI data were collected at 3T Siemens Skyra machines (software platform VD13, 32-channel head coil) using a standardised protocol across scanner sites which included T_1-weighted, T_2-weighted (FLAIR), susceptibility-weighted T_2*, diffusion, resting state, and task-based functional MRI [7, 198, 213]. In the current project we used T_1-weighted and diffusion MRI data. The T_1-weighted scans were acquired using a 3D MPRAGE (sagittal) sequence, R=2, TI/TR=880/2000ms, with a 1.0x1.0x1.0mm voxel size and 208x256x256 matrix. Diffusion acquisition parameters were MB=3, R=1, fat sat, b=0 (5x + 3x phase-encoding-reversed), 1000 (50x), 2000 (50x), using a voxel size of 2.0x2.0x2.0mm and a 104x104x72 matrix [7, 213].

2.3.2 *Bergen Breakfast Scanning Club*

The Bergen Breakfast Scanning Club (BBSC) dataset is a deep imaging dataset included repeatedly acquired MRI scan and phenotype data of N = 3 individuals over a period of circa one year with a summer break in the middle of the scanning period [321, 322]. This resulted in a total number of N = 103 scans, which were relatively equally distributed across subjects ($N_1 = 38$, $N_2 = 40$, $N_3 = 25$).

T_1-weighted volumes of the three BBSC participants were acquired with a spin echo sequence (TE = 2.95ms, TR = 6.88ms, FA = 12°, TI = 450, 188 slices, slice thickness = 1mm, in-plane resolution = 1mm×1mm, FOV = 256mm, isotropic voxel size = 1mm^3) at a 3T GE system with 32-channel head coil [321, 322].

2.3.3 *Frequently Traveling Human Phantom*

The frequently traveling human phantom (FTHP) MRI dataset [234] contains the data of one male subject with N = 157 independent imaging sessions. The sessions were conducted at N = 116 different locations, totaling N = 557 MRI volumes.

T_1-weighted volumes of the FTHP participant were acquired at different scanners with various scanning parameters [234] (see also https://www.kaggle.com/datasets/ukeppendorf/frequently-traveling-human-phantom-fthp-dataset).

The imaging sites where the FTHP study was executed were informed about the goal of using the acquired scans for volumetry, and about a standardized protocol agreeing with the ADNI [135]recommendations for magnetization prepared rapid gradient-echo (MP-RAGE) MRI for volumetric analyses. While there was some variation in the acquisition parameters, the acquisitions are representative of clinical routine MRI scans with the goal of volumetry, yet offering a higher quality based on the low level of motion artefacts [234].

2.3.4 *Norwegian Cognitive NeuroGenetics and TOP samples*

As validation data to the two densely sampled datasets (BBSC and FTHP), we used healthy controls from the Norwegian Cognitive NeuroGenetics (NCNG) sample [93] and the local TOP data [301]. Together these data-sets include a total of N = 209 people's scans at 1.5T, and N = 856 at 3T, respectively.

NCNG data [93] were acquired at a 1.5T Siemens Avanto scanner using two 3D MP-RAGE T1-weighted sequences (TR = 2400 ms, TE = 3.61 ms, TI = 1000 ms, FA = 8°, with 160 sagittal slices (1.3 x 1.3 x 1.2 mm)).

TOP data were acquired at 3T on a GE 3T Signa HDxT (TR = 7.8ms, TE = 2.956ms, FA = 12°; one subset with HNS coil, one subset with 8HRBRAIN coil), and a GE 3T Discovery GE750 (TR = 8.16ms, TE = 3.18ms, FA = 12°).

2.4 Statistical Modelling

This section describes the statistical modelling done throughout the paper but not the process of brain age predictions (see sections 2.2.1 and 2.2.2 for these details). However, brain age was used as outcome variable in Study B. Throughout the five papers, linear

mixed models were used with the random effect(s) being limited to random intercepts. Where applicable, these random intercepts were the scanner site, field strength, slice thickness for cross-sectional data, and an additional random intercept on the level of the participant for longitudinal and repeat-scan data. In case of model convergence issues when applying random effects, standard linear models (containing only fixed effects) were chosen. Additionally, outcome measures were adjusted for sex and the sex-age interaction before being associated with age. In detail, sex, sex-age-interaction, and site corrections in Study A were conducted using fixed effects only to adjust white matter metrics (M):

$$\hat{M} = \beta_0 + \beta_1 \times Age \times Sex + \beta_2 \times Sex + \beta_3 \times Site. \tag{2.13}$$

For study B, we used site as a random effect and modelled brain age from the variables used in Eq. 2.13. Yet, we included also age as covariate in order to correct age-bias in brain age (BA;i.e., the equivalent of the corrected brain age gap, see also section 2.2.3), for an adjusted association with various phenotypes P:

$$\hat{BA} = \beta_0 + \beta_1 \times Age + \beta_2 \times Sex + \beta_3 \times Age \times Sex + \beta_4 \times P + RI(Site). \tag{2.14}$$

For study C, longitudinal/densely sampled data (with constant sex, and constant scanner site for the BBSC sample) were used, and the random intercept at the participant level (ID) was estimated:

$$\hat{BA} = \beta_0 + \beta_1 \times Age + RI(ID). \tag{2.15}$$

However, although data were only available for a single subject, for the FTHP, there were additional variables available in the data-set, allowing to model also field strength (FS), and slice thickness (ST) as random intercepts:

$$\hat{BA} = \beta_0 + \beta_1 \times Age + RI(ID) + RI(FS) + RI(ST). \tag{2.16}$$

For study D, the same modelling was applied as in study B or Eq. 2.14, but regressing age and covariates of no interest on white matter metrics M instead of brain age, or the annual rate of change in white matter metrics ΔM:

$$\hat{M} \wedge \Delta\hat{M} = \beta_0 + \beta_1 \times Age + \beta_2 \times Sex + \beta_3 \times Age \times Sex + RI(Site). \tag{2.17}$$

For the voxel-level analysis in paper E, a random intercept model was deployed with only one random intercept on the level of the participant, similar to Eq. 2.16 but regressing on white matter metrics M instead of brain age (using randomize with 10,000 permutations in FSL [289], and threshold-free cluster enhancement [290]):

$$\hat{M} = \beta_0 + \beta_1 \times Age + RI(ID). \tag{2.18}$$

This was extended to account for both random effects at the level of the participant *and* scanner sites in the analyses of regional and global white matter metrics in Study E:

$$\hat{M} = \beta_0 + \beta_1 \times Age + RI(Site) + RI(ID). \tag{2.19}$$

Finally, for the modelling of both the effect of PGRS on the annual rate of change in white matter microstructure as well as cross-sectional white matter microstructure we used simple linear models accounting for sex, age, site, and the sex-age interaction. Simple linear models were used instead of mixed effects models for model convergence:

$$\hat{M} = \beta_0 + \beta_1 \times Age + \beta_2 Sex + \beta_3 Age \times Sex + \beta_4 \times \times Site. \tag{2.20}$$

Across papers, β-values were standardized to be compared across predictors and models, and p-values were adjusted conservatively using Bonferroni correction to correct for multiple comparisons.

SUMMARY OF ARTICLES

3.1 Paper A: Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing

Background: Unveiling the details of white matter maturation throughout ageing is a fundamental question for understanding the ageing brain which still requires further investigation. Particularly for white matter, age-dependencies are still to be investigated using representative samples. There are various biophysical models / diffusion approaches which allow for the extraction of detailed information about white matter microstructure. Neither the correspondence of such models with each other has not been systematically tested in representative samples nor their age relationships. Beyond direct age associations of white matter metrics, brain age, a suggested bio-marker of health, can be utilized to improve our understanding of white matter ageing by providing information on feature contributions and by providing predictions which can be compared across diffusion approaches. Benchmarking age-associations in a largely healthy sample is the first step towards normative models describing associations between white matter and age. These benchmarks can inform further investigations into disease by examining deviations from the suggested ageing patterns.

Methods: Tract-based spatial statistics were applied to estimate region and tract level averages across both the John Hopkins University tract and region atlases of multiple diffusion approaches. The diffusion approaches included Diffusion Tensor Imaging, Diffusion Kurtosis Imaging, the Spherical Mean Technique and its multi-compartment extension the multi-compartment Spherical Mean Technique, the Bayesian Rotationally Invariant Approach, and White Matter Tract Imaging, an extension of Diffusion Kurtosis Imaging.

We first probed optional hyperparameter tuning versus train and test sample splits, and estimated brain ages for several diffusion approaches and their combination in a cross-sectional sample of $N = 35,749$ healthy UK Biobank participants. Brain age model performance was compared across diffusion approaches, and features ranked by their contribution to the models. Additionally, raw and adjusted correlations between white matter and age were presented. We also estimated the correlations between white matter metrics.

Results: White matter microstructure estimated from conventional and advanced diffusion MRI approaches' predicted brain age consistently, yet best when combining approaches. Similarly, when estimating correlations between the metrics from the

different diffusion approaches, we find two major association clusters where metrics were either strongly positively associated with each other or strongly negatively associated, respectively. Fornix and forceps minor features explained most variance in age across diffusion approaches in the brain age models. Similarly, these two regions were strongest associated with age when using crude correlations.

Additionally, we presented general patterns of white matter deterioration for higher ages in fornix, forceps minor, and across the brain in age charts, which generally showed a pattern of lower fractional anisotropy, extra axonal water fractions, and kurtosis metrics at higher ages. Opposing, diffusivity (axial, radial, and mean diffusivity in intra-axonal and extra-axonal space) showed higher values at higher ages, together with intra-axonal water fractions.

Conclusions: There are general patterns of age-associations of white matter, independent of the assumptions made by the biophysical models the different diffusion metrics were estimated on. These age-association patterns outline, among others, forceps minor, which has frequently been highlighted as age-sensitive regions, as well as fornix, which has even been suggested as a deep brain stimulation target, as most age-sensitive. We provided validations for the fact that different advanced biophysical models measure the same underlying concept of white matter degeneration captured by brain age and age associations. Overall, we encourage the application of multiple diffusion MRI approaches for detailed insights into white matter, and the further investigation of fornix and forceps as potential bio-markers of brain age and ageing.

3.2 Paper B: Bio-psycho-social factors' associations with brain age: a large-scale UK Biobank diffusion study of 35,749 participants

Background: This paper was a direct follow-up of the first paper using the estimated brain age scores to relate them to bio-psycho-social factors. While different dMRI-derived parameters provide valuable information about white matter architecture, which can be associated with macroscopic outcomes, such as bio-psycho-social factors, investigations into white matter brain age-phenotype associations are still lacking. It is not only important to close this gap in general, but also to further investigate whether brain ages estimated from white matter metrics derived from different diffusion approaches are consistently related to bio-psycho-social variables. Only by relating brain age to real-world outcomes its general and clinical utility can be assessed.

Methods: Bio-psycho-social factors were used in addition to brain ages based on cross-sectional dMRI data from $N = 35,749$ healthy UK Biobank participants. The statistical analyses comprised establishing linear mixed models of bio-psycho-social factors explaining variability in brain age; namely, a baseline model as comparison, and models of socio-demographics, cognitive scores, life-satisfaction, and health and lifestyle factors. The outcome variable, brain age was varied, to gain an understanding of the stability of associations between phenotypes and brain ages estimated from differently estimated white matter microstructure features. The resulting models were then compared in their relative contribution to explaining brain age variability, and their single predictors were examined with the goal of identifying phenotype-brain age associations which were concordant across diffusion approaches.

Results: Bio-psycho-social models of cognitive scores, life satisfaction, health and lifestyle factors explained a small proportion of the brain age variance. Single bio-psycho-social factors that were significantly associated with brain age across the different brain age estimates were found for waist-to-hip ratio, diabetes, hypertension, smoking, matrix puzzles solving, and job and health satisfaction and perception. Furthermore, we found large variability in brain age differences between sexes and ethnicity groups.

Conclusions: The observed associations were consistent with previous research outlining particularly health and lifestyle associations with brain age derived from T_1-weighted MRI. These phenotype-brain age association findings highlighted white matter-derived brain age as a similar concept to grey matter-derived brain age. Importantly, independent of the diffusion approach the particular brain age was calculated on, associations between brain age and bio-psycho-social variables were consistent. This allowed the conclusion that characteristics of both white matter meso and microstructure reflect brain age similarly, and potentially describe similar underlying white matter degeneration processes which can be directly linked back to bio-psycho-social variables. Novel findings included the apparent associations between life satisfaction factors and brain age. While, overall, only small proportions of brain age could be explained, future brain age associations studies should consider sex, ethnicity, cognitive factors, as well as health and lifestyle factors as well as their interaction as possible confounders.

3.3 Paper C: Considerations on brain age predictions from repeatedly sampled data across time

Background: Brain age is a promising biomarker of brain or general health. The metric has been extensively related to different diagnoses but to a lesser extent been tested and validated in longitudinal data. Hence, to extend the metric's clinical applications, the large intra-individual variability in age predictions needs addressing. Such intra-individual variability becomes particularly important when the aim is to tailor individual-specific treatments. A step towards validating whether brain age predictions can be used for such "precision medicine" approach is to validate group level models on densely sampled individual level data. Densely sampled data refers to repeated acquisition of data from a single individual or a small group of individuals over time to reach clearer insights into individual-specific trajectories of brain development as compared to group average statistics which do not allow for such detailed individual level investigation. Hence, here, we showed how brain age predictions perform for the same individual at various time points and validated our findings with age-matched healthy controls. We used an established brain age model producing brain age estimates from minimally processed T_1-weighted magnetic resonance imaging data.

Methods: We used the pre-trained deep neural network *pyment* to predict brain ages from densely sampled T_1-weighted magnetic resonance imaging data from three individuals (BBSC1-3, N = 3) scanned in total $M_{BBSC} = 103$ times over a one-year interval, and an independent data set including one individual (FTHP1, N = 1) scanned $M_{FTHP} = 557$ times over a three-year interval. To assess scan quality, we use (a) scanner

field strength and (b) quality control parameters derived from MRIQC, including (1) noise measures: contrast-to-noise ratio (CNR), signal-to-noise ratio (SNR), coefficient of joint variation of grey and white matter (CJV), (2) measures based on information theory entropy-focus criterion (EFC) and foreground-background energy ratio (FBER), (3) white-matter to maximum intensity (WM2MAX), and (4) other measures: full-width half-maximum (FWHM). These parameters were then associated with the derived brain ages to examine potential sources of brain age variability. We validated the findings of an effect of field strength in FTHP1 using locally collected cross-sectional data of healthy controls from the TOP and the Norwegian Cognitive NeuroGenetics samples, together $N_{1.5T} = 209$ and $N_{3T} = 856$.

Results: We found small within-subject correlations between age and brain age. These correlations were stronger when more data were sampled over a longer period of time. However, the magnitude of deviations between chronological ages and brain ages were relatively stable within subjects, but showed strong discrepancies between individuals. This was accompanied by different QC metrics being more and less predictive of brain age, both dependent on the individual and the sample: EFC and FBER were significant predictors of brain age in the BBSC sample, yet none of the QC metrics predicted brain age in the FTHP suject. Overall, there were no replicable patterns of QC metrics associations with brain age, rendering QC effects as inconclusive. However, we found evidence for the influence of field strength on brain age in FTHP1 which replicated in the cross-sectional validation data. The findings showed lower prediction errors at 1.5T compared to 3T, whereas predictions were characterised by generally higher brain age estimates at 1.5T compared to 3T.

Conclusions: Brain age estimates are potentially influenced by acquisition parameters and scan quality. Particularly the field strength was shown to influence brain age estimates using the 3D convolutional neural network *pyment*. While the influence of acquisition protocol was presented previously, the identified confounding effects of field strength were a novel finding. An avenue for future brain age modelling could be to employ multiple, more specific models, tuned to developmental and individual differences and acquisition parameters. For example, dependent on the available data, one could train models for certain age ranges and field strengths.

3.4 Paper D: Brain asymmetries from mid- to late life and hemispheric brain age

Background: The brain demonstrates various age-sensitive asymmetries. Additionally, there are several differences in brain asymmetry between healthy controls and disease groups, including neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease, and psychiatric disease such as obsessive–compulsive disorder. This outlines the clinical utility of assessments of brain asymmetry. Yet, a systematic mapping of grey and white matter asymmetries (i.e., across metrics) from midlife to old adulthood during healthy ageing is still missing. Hence, we set out to map brain asymmetries throughout midlife to older ages and the relationship of such asymmetries with age. Furthermore, we extended the concept of brain age, the estimation of chronological age from sets of neuroimaging features, by differentiating between left

and right brain ages estimated from diffusion-weighted, T_1-weighted, and multimodal magnetic resonance imaging (MRI). Such extension was investigated in order to provide brain age estimates which not only reflect brain asymmetries but also allow for proxies of these asymmetries by comparing left to right brain ages.

Methods: After quality control (including corrections for noise, Gibbs-ringing, susceptibility-induced and motion distortions, and eddy current artefacts), and isotopic smoothing, we estimated conventional and advanced diffusion approaches. The diffusion approaches included Diffusion Tensor Imaging, Diffusion Kurtosis Imaging, the Spherical Mean Technique and its multi-compartment extension the multi-compartment Spherical Mean Technique, the Bayesian Rotationally Invariant Approach, and White Matter Tract Imaging, an extension of Diffusion Kurtosis Imaging. Tract-based spatial statistics were applied to estimate region and tract level averages across both the John Hopkins University tract and region atlases of multiple diffusion approaches. For T_1-weighted images, we select the Desikan-Killiany atlas as a parcellation scheme after using the standardized Freesurfer cortical reconstruction (recon-all) pipeline, which estimates metrics for volume, surface area and cortical thickness. From these estimated regional features, we select regions specific to either left and right hemisphere (discarding regions and tracts which cross hemispheres).

We presented brain asymmetries from multimodal magnetic resonance imaging (MRI) UK Biobank ($N > 39,500$) data using the laterality index for region-averaged and global grey and white matter microstructure metrics. We furthermore showed how to leverage brain asymmetries by estimating hemispheric brain age from the left and right hemisphere separately instead of from the whole brain. Finally, we assessed whether the laterality index of hemispheric brain age is similarly age-associated as the laterality index of brain features.

Results: Left, right, and whole-brain age predictions are strongly correlated across modalities and show similar prediction errors. We found no significant influence of hemisphere, modality or handedness on hemispheric brain age, but age-sensitivity of the hemispheric brain age asymmetry. Moreover, we showed that various cardiometabolic risk factors concordantly related to hemispheric brain age.

Most grey and white matter features were asymmetric, with these regional asymmetries presenting themselves with moderate to high effect sizes. Interestingly, asymmetries could even be observed when observing hemisphere-wide averages by comparing age-curves for the examined multimodal MRI metrics between left and right hemispheres. The presented regional metrics were also age-sensitive, yet only around 50% of the T_1-weighted and 80% of the diffusion-weighted features' asymmetries were age-sensitive. The regional asymmetries presented a pattern of larger asymmetry at a higher age across grey and white matter, but lower brain age asymmetry, whether brain age was calculated from T_1-weighted, diffusion-weighted or multimodal MRI data. Our findings highlighted several brain regions' asymmetry as particularly age-sensitive. These include WM asymmetries in the cingulate which were strongest in their positive age-associations, and cerebral peduncle and superior longitudinal fasciculus metrics' asymmetries which were consistently strongly associated with age. For T_1-weighted metrics, largest negative age-associations were observed for asymmetries in the lateral ventricles, putamen, hemispheric white matter, and cerebellum volumes, as well as rostro-middle thickness. Largest positive associations with asymmetries included

accumbens volume, WM surface area, amygdala volume, caudal anterior cingulate thickness, and WM volume.

Conclusions: Our findings emphasise the presence of brain asymmetries in both grey and white matter. These metrics and their asymmetries are largely age-dependent, and appearing higher later in life. These benchmarks can inform further research both investigating fundamental as well as clinical questions. Hemispheric brain age can be used to assess brain health specific to a single hemisphere, and asymmetries in hemispheric brain age capture the general trend of decreasing asymmetry in white matter.

3.5 Paper E: Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and Alzheimer's disease in the UK Biobank

Background: While there are multiple studies examining white matter microstructure in cross-sectional data, well-powered studies examining white matter ageing are still lacking. Additionally, the different studies usually analyse either region-averaged or voxel-level changes, supplemented by global average scores. Studies giving a systematic overview across these levels of analysis, or beyond diffusion tensor imaging are still lacking. Finally, while the field of imaging genetics is experiencing an immense growth, polygenic risk associations in longitudinal data are also a rarity. Thus, associating the polygenic risk of a given person with the trajectories of their brain changes is a valuable step of adding a practical component to identified patterns of changes. Additionally, at this point there is no studies giving a broad overview of white matter associations with polygenic risk scores of common brain and psychiatric disorders.

Methods: We use the UK Biobank repeated measures diffusion MRI data of $N = 2,678$ participants (53.36% females), which was taken at two time-points with a mean inter-scan interval of 2.43±0.73 years. Additionally, we use the portion of the UK Biobank cross-sectional sample which was not included in the longitudinal sampling $N = 31,056$ as a validation sample for cross-sectional investigations of polygenic risk score associations with white matter metrics. After quality control exclusions, tract-based spatial statistics (TBSS) were estimated for conventional and advanced diffusion models, and statistics run within the confines of the fractional anisotropy white matter skeleton. We estimated polygenic risk scores for Alzheimer's disease and common psychiatric disorders, including Major Depressive Disorder, Bipolar Disorder, Anxiety Disorder, Autism Spectrum Disorder, Schizophrenia, Attention Deficit Hyperactivity Disorder, and Obsessive-Compulsive Disorder. These polygenic scores were informed by previous large-scale genome-wide association studies.

Results: Our findings demonstrated lower global and regional fractional anisotropy, the intra-axonal water fraction, and kurtosis metrics at higher ages. In contrast, axial, radial, and mean diffusivity metrics as well as free water and extra-axonal water fractions appeared higher at higher age. These findings highlighted fornix and cingulate which presented strongest age-associations over time. Additionally, the annual rate of white matter change accelerated at higher ages for most examined diffusion metrics, indicating naturally occurring accelerated ageing at higher ages. Voxel-level analyses

showed general trends of decreasing anisotropy, and variable spatial patterns for other diffusion metrics, indicating differential changes in frontal compared to other brain regions. By assessing associations of global and regional white matter change with polygenic risk scores, we provided practical implications of the presented age-related changes. We identified strongest and most consistently associated with Alzheimer's Disease polygenic risk scores with the annual change in inferior cerebral peduncle. Fornix and cingulate annual change presented some of the strongest polygenic risk associations with different psychiatric disorders.

Conversely, considering cross-sectional measures of white matter, global averages of multiple diffusion metrics were particularly associated with Alzheimer's polygenic risk in the observed longitudinal sample, which was however not the case in the validation sample, which showed stronger associations with autism spectrum disorder. Regional cross-sectional white matter metrics showed spatially distributed associations with the polygenic risk of several disorders, consistently presenting strongest and most significant associations in the limbic system and brain stem. Finally and importantly, polygenic risk scores for Alzheimer's Disease and psychiatric disorders were stronger related to the annual rate of change than cross-sectional WM measures.

Conclusions: The presented results indicate that demyelination processes and WM disintegration are heterogeneously distributed across the brain during midlife and older ages. Differences in water compartments' diffusivity suggest weaker plasticity in the frontal regions during this short time. Importantly, longitudinal changes of white matter reflected the genetic risk for disorder development better than the utilised cross-sectional measures. Moreover, the spatial pattern of these brain-gene associations is distributed, indicating that certain brain regions correspond more to one or a set of polygenic risk scores than others. For example, such as the Our results underline the importance of longitudinal data analyses in understanding of the ageing brain and its' genetic underpinnings.

DISCUSSION

In this this project we investigated the question: How is the human brain architecture affected by healthy ageing? The goal was to identify concrete brain tissue biomarkers of healthy ageing / brain maturation, including white matter microstructure features as a main feature, brain age as well as grey (thickness, volume, surface area). Such biomarkers can be potentially explored as clinical markers of different disorders.

The identification of brain biomarkers has been done from a conceptual and methodological perspectives by validating the various biophysical diffusion approaches of white matter, and by using brain age models on both grey and white matter features. From a practical point of view, the white matter age-association and brain age analyses outlined central and ventricle-near areas, such as the fornix and the forceps minor, as central age-associated regions. These findings are, hence, not only part of the biomarker or brain age literature, but also relevant in the broader context of white matter organisation and brain ageing pathways. Overall, our results contribute to a better understanding of the ageing brain and white matter mapping.

In the following Discussion section, the revealed findings from the five publications will be discussed in relation to each other and in the context of the research field. The focus of the discussion will be methodological choices and considerations before turning to brain age predictions, the identified brain regions which provide markers of ageing and potentially of disease development, as might be identified by future research, and, finally, ethical considerations.

4.1 Methodological considerations

In the last decade, problems to replicate MRI findings were increasingly outlined in a series of publications, underlying the importance of the issue [35, 91, 131, 210]. Limitations to the ability to replicate scientific findings are not unique to the field of neuroscience, and exhibit varying manifestations through scientific practice [160]. Recent discussions in the human brain mapping community examine the revealed scientific findings in terms of the statistical power for brain-behaviour associations [49, 80, 160, 184, 192, 194]. In order to increase the statistical power of brain-wide associations, one needs to either acquire sufficiently large data (e.g., from many individuals) to detect small effects [184, 194], or, in the case of smaller samples, simply observe phenomena which express large effect sizes, make use of within-subject designs (i.e., repeated or longitudinal measures), or adapt analytical choices towards reducing

the number of tests, for example, by diverging from voxel-by-voxel analyses [49, 80, 192]. The movement towards large-scale population representative data ("going big"), such as the UK Biobank [292] allows one to obtain more detailed and generalisable insights into brain ageing and, thereby to isolate brain biomarkers.

Another trend is the examination of densely sampled data from few subjects [114, 247, 321]. Instead of sampling one or few data points from many participants, many data points are being sampled from a reduced number of individuals. Such "going small" approach allows one to focus on individual differences as well as a better understanding of factors influencing measurement reliability [321]. Here, we made use of both trends, by establishing models first using population-level data and attempted to validate a brain age model in densely sampled data.

4.1.1 Samples

Although the UK Biobank contains an enormous imaging sample compared to the average samples reported in the literature [295], there are various biases which might limit the generalisability of the obtained findings. First of all, the UK Biobank is dominated by Caucasian individuals, with individuals holding another ethnicity being underrepresented. To this end, the UK Biobank could be seen as representative of parts of the North European and North American population, also known as WEIRD (Western, Educated, Industrial, Rich, Democratic) population [123]. Yet, a potentially even larger problem is the bias introduced by the health status of the sampled *volunteers* [307]. This volunteer-based sampling strategy seems to have resulted in more healthy participants compared to the rest of the population, rendering the UK Biobank as potentially not even representative of the UK population [106, 190, 277].

The sampling bias in densely sampled data sets is even more apparent. Here, only a very small selection of participants is observed, usually convenience sampled from the same lab [114, 247, 322]. This comes very clear in Study C, where the individual differences are highlighted when examining brain age prediction. Here, particularly one individual shows consistently strong deviations of the prediction from the chronological age of on average nearly 10 years. While it is not possible to generalise to the broader population from such case-study like "low N investigations", important lessons can be learnt about the measurements, which can directly inform clinical use cases. For example, the case of the large brain age gap of the one participant in Study C illustrates the limitations of brain age models applied on the individual level compared to group statistics as usually done when validating machine learning models: group-level statistics of model performance do not necessarily correspond with longitudinal/repeatedly sampled data from the same individual. Yet, such tests are an invaluable part of moving potential biomarkers such as brain age towards clinical use. Finally, and equally valid across the utilised longitudinal data sets, longer inter-scan intervals would likely be fruitful and a better source of information. However, multi-wave longitudinal data with population-level samples and long inter-scan intervals are yet to be established.

4.1.2   Brain Metrics

Across studies we used T$_1$-weighted MRI but mainly diffusion derived images and presented age associations (using voxel-values, region-averages, and whole brain averages) as well as brain age predictions based on region- and whole-brain-averaged metrics. Brain metrics are usually averaged for simplicity, explainability, but at times also to maximise statistical power by reducing the number of tests run. The reduction in the number of statistical tests of the same hypothesis increases statistical power compared to the scenario where a larger number of features is being tested on. Across studies we noticed, however, an information loss when averaging across the brain. On the one hand, region-averaged features provided specific and explainable age-associations. For example, we observed lower fractional anisotropy at higher ages in most white matter regions, giving an intuition about white matter degeneration as part of the healthy ageing process. Additionally, averaging over populations of voxels provides more stable estimates than voxel-level estimates, for example, from the standard model of white matter diffusion [57]. On the other hand, voxel-based analyses might be a more sensitive method to detect age-related changes, as they provide finer-grained spatial details. Averaging across voxels might mask such details. However, voxel-level analyses require large enough samples to be sufficiently powered, and might otherwise lack specificity [91] and sensitivity [227]. When using more recent cluster-based corrections to increase specificity, such as Threshold-Free Cluster Enhancement (TFCE) [290], sensitivity might still be limited (and lower than for uncorrected analyses), especially when effect sizes (e.g., of age or time point differences) are small [227]. However, large samples, as used in this work, provide sufficient power and hence large enough sensitivity for voxel-level analyses [194].

Furthermore, the selection of parameter settings for voxel-level analysis might be arbitrary, resulting in low replicability or even reproducibility (e.g., using tract-based spatial statistics analysing diffusion MRI [19]). After the results are obtained, a remaining step is the inference from these findings, often involving a subjective level of highlighting results [298] and a discrepancy between verbal expressions and statistical observations [335]. This gap between statistical outcomes and their presentation and interpretation induces bias and limits reproducibility as well as generalisability [19, 298, 335].

In the hope to reduce the outcomes' dimensionality and to deliver concrete (and easy-to-interpret) markers, machine and deep learning models can be employed. Machine learning models such as tree boosting algorithms show excellent performance on tabular data [285] and offer explainable outputs, such as feature importance rankings (as an alternative to scale-less feature gain). Deep learning models show better performance on more complex data, both in 2D [145, 297] and 3D data [63, 171, 248, 248, 297] data. However, to establish explainability of predictions is more difficult than for simpler models [118]. With other words, certain models of different complexities will fit better based on the characteristics of the data, which ultimately impacts model performance (e.g., in case of brain age predictions using 3D convolutional neural networks (CNNs) [63, 171, 248, 248, 297]) . On top of this consideration, the features which are feeding into a model can also be used to design additional features. This feature design practice can involve various different methods and steps to ultimately improve the

model's performance. For example, engineering features and their dependencies using patch-wise grading based on image similarity metrics leads to high brain age model performance [33], which is longitudinally reliable [34]. At the same time, heavily processed features might decrease the explainability component of a given model, meaning that even when implementing steps to make model choices transparent, the design of the features might still not allow for clear conclusions from the model choices such as weightings of features, for example, due to highly abstract features [118]. Hence, both data and models are important in any inferential statistical context [50]. Therefore, we present in the following first different considerations on feature extraction from T_1-weighted and diffusion MRI which (form the data and hence) lay the foundation for modelling choices. We then discuss the concept of brain age, its potential, its limitations, both in general and in the light of how the concept is commonly used, and finally future directions for brain age research.

4.1.2.1   *Diffusion MRI*

Diffusion MRI allows us to reformulate water diffusion, based on the assumptions of geometrical models, and extract a wide range of tissue parameters. These parameters could be used as biomarkers, in the framework of this project via age-associations. There is a standard diffusion model containing a series of white matter metrics [57, 229]. However, the standard diffusion model is ill-posed and possesses degeneracies [229]. While diffusion tensor imaging has been established as the predominant approach for white matter modelling, due to the relative simplicity of diffusion tensor imaging (DTI), it is limited in various aspects. A gold standard for which white matter microstructure approaches to preferably use is still to be established and requires thorough systematic examinations of the approaches in clinical and non-clinical settings. Hence, while the white matter metrics extracted by different diffusion approaches are informative for a better understanding of white matter ageing and a range of diseases, there is also no unified theory of white matter microstructure ageing [259].

In this project, we performed a comprehensive comparisons of multiple state-of-the-art diffusion approaches in relationship to age. We have revealed differences in the age-associations of the single approaches' metrics in parallel to relatively coherent pattern of associations outlining the stability of the standard diffusion model independent on the applied diffusion approach (such as DTI, DKI, BRIA, SMT, SMTmc, WMTI). In that sense, one could argue that diffusion models such as the Bayesian Rotationally Invariant Approach (BRIA) might be a preferable choice, as multiple intra and extra-axonal metrics on both diffusion metrics, fractional anisotropy, and cerebrospinal fluid are provided. However, due to the different biophysical assumptions of the different diffusion approaches and the resulting differences in the models sensitivity to different bio-physical phenomena, multiple models can act as complementary to each other. Furthermore, different analytic strategies will consequently influence the results, and limit their comparability. For example, in the presented studies here, we chose to conduct analyses using the FA skeleton. The FA skeleton-based approach minimises confounders such as partial volume and normalisation-resulting misalignment effects [200], leading to exact estimations of voxel-level and whole-brain metrics.

Yet, TBSS and its FA projection procedure have also various shortcoming based

on the procedures' underlying assumptions [19]. Only the highest FA values serve to establish the FA skeleton, as these values are thought to represent the centre of the tracts. This procedure can result in the loss of information [337] and the introduction of artefacts [279], usually produced by misalignments [152]. The alignment depends on how homogeneous the sample is and is influenced by noise [19]. In case of large between subject differences, the skeleton estimation will show inaccuracies, for example, when examining disorders affecting the white matter. This is also true for smaller or thinner tracts, such as the fornix [19]. These issues are not resolved by increasing the resolution [19]. However, previous studies have shown excellent reproducibility [2] and test retest reliability [146, 191, 205, 207, 263] of FA skeleton derived values in different age groups. Furthermore, the utilised sample was relatively homogeneous and large enough to reliably identify main tracts (see Appendix 4 for global and Appendix 5 for region-averaged test-retest reliability of the utilised metrics in the longitudinal sample).

Moreover, when considering region-averaged metrics, there are several issues with proposed atlases. On the one hand, there will not be any one atlas which perfectly fits all brains. On the other hand, the FSL implementation of the most commonly used (John Hopkins University (JHU)) white matter atlas shows considerable problems. Namely, the region labelled tapetum in the JHU is a region that does not overlap with the actual white matter contained in the mean FA skeleton in the UK Biobank (see Figure 4.1). This is consistent with a previous study [235] showing that this is a persistent problem across FSL versions.
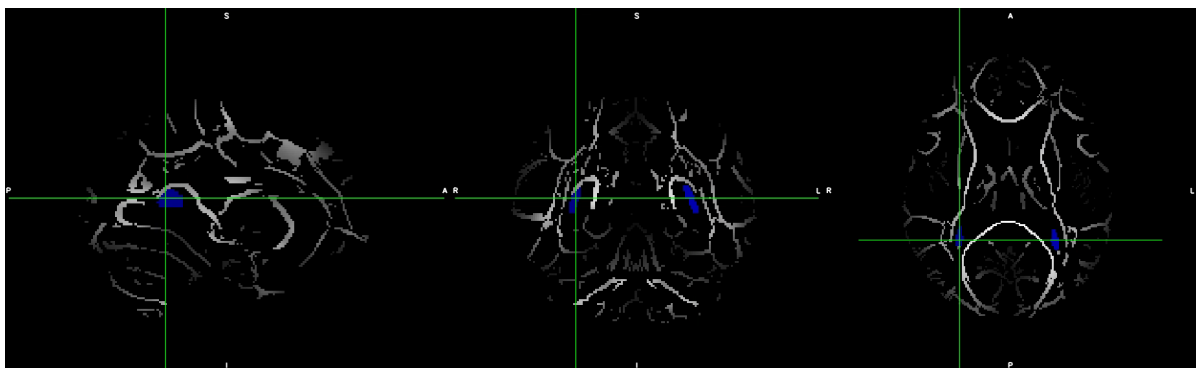
Fig. 4.1: The figure presents the tapetum white matter according to the John Hopkins University white matter atlas label (blue) overlayed by the mean fractional anisotropy skeleton (grey). The cross is located at the following MNI coordinates: x=60, y=78, z=86.

Moreover the JHU regions labelled fornix stria terminalis overlap to a minimal extend with the mean FA skeleton in the UK Biobank (see Figure 4.2). Such missing overlap in a parcellation scheme is problematic as it leads to averaging over voxels not containing any true signal but only noise. Hence, region-averages for tapetum and fornix-stria-terminalis can be assumed to be measures of random noise in the mjority of cases in the FSL releases (up to the end of 2023, the date of this work).

Other regions approximate the WM skeleton better. However, research is currently ongoing examining the implementation of different atlases across commonly software packages. While such mismatches between atlases (or their software implementations) and the true skeletal data are limitations, the TBSS WM skeleton approach is otherwise
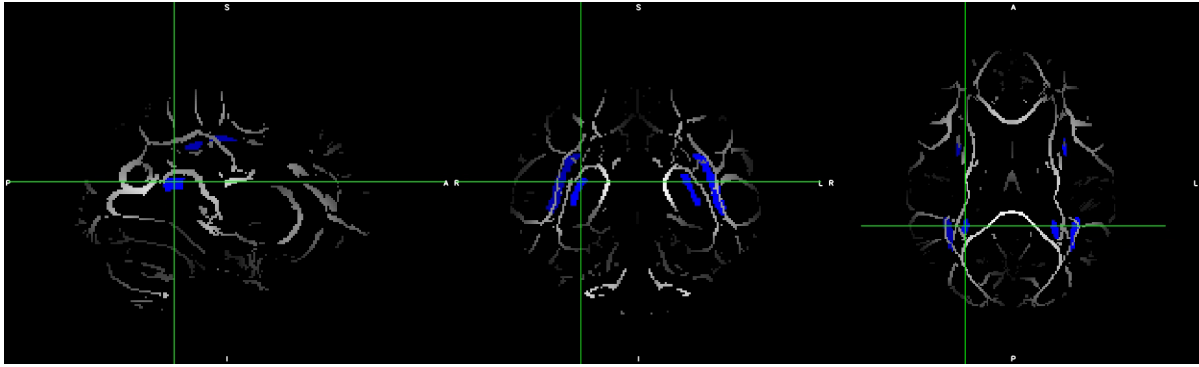
Fig. 4.2: The figure presents the fornix stria termininalis white matter according to the John Hopkins University white matter atlas label (blue) overlayed by the mean fractional anisotropy skeleton (grey). The cross is located at the following MNI coordinates: x=62, y=82, z=90.

robust to various confounders as introduced for example by partial voluming or misalignments resulting from spatial transformations [200].

While there are multiple other approaches, for example, using supervised or semi-supervised tract reconstruction algorithms, such as TRActs Constrained by UnderLying Anatomy (TRACULA) in Freesurfer [99], such approaches deliver results which are not directly comparable to the results obtained from TBSS, and cannot account for the described microstructure modelling. Furthermore, these methodological advances are an active field of research and their yet low accuracy (as suggested in [276]) is to be improved. Fixel-based analyses [82], another approach of analysing diffusion MRI data, might hold promise for further microstructure and especially fibre-specific explorations. Yet, the framework still requires additional validation and translation, as well as (and similar to the other approaches of analysing diffusion data) work on the interpretability of produced metrics [82].

Overall, more work is required to verify the different diffusion approaches and their practicality against each other. This does not only include working towards a theory of white matter microstructure ageing, but also a more basic understanding of the different metrics to allow clearer connections between the metrics and biophysical processes such as axonal decay.

4.1.2.2 T_1-weighted MRI

The most commonly used metrics which can be extracted from T_1-weighted MRI are the volume of cerebrospinal fluid, grey and white matter, the surface area of both grey and white matter, as well as grey matter thickness. These metrics are based on probability values for the respective tissue classes [99]. Their estimates were found to be increasingly more reliable at more recent Freesurfer versions [122]. However, different processing pipelines will influence the estimates [39, 122], just as software versions [122, 344], and scanner/acquisition site [107].

Hence, standardised protocols, such as the Human Connectome Project protocol [314], and analysis pipelines, such as sMRIPrep (https://www.nipreps.org/smriprep/), might add to standardised protocols and less variability due to design,

processing, and analysis choices. However, using standardised pipelines also restricts the flexibility when exploring new methods. While this was not true for T_1-weighted data used here, in the presented studies on white matter microstructure, we used various methods which are not implemented in standardised pipelines, requiring in-house code.

Finally, region averages are dependent on the selected parcellation scheme, which will consequently also influence the results of analyses on the obtained average scores. We used a relatively fine-grained parcellation scheme, for a higher number of features to be fed into brain age models. Other atlas solutions which are even more fine-grained use post-mortem data [121], allowing for the gene mapping of tissue properties [228, 238], or offer mappings of brain metabolism, including neurotransmitter as well as molecular systems [87, 120, 154, 185, 280, 324]. More recent atlases include information from multiple modalities, for example, using multiple MRI modalities [110]. A new frontier is to combine atlases to incorporate anatomical, metabolic, and immune system features applying normative models [188], which can in the future be extended to more advanced multi-modal atlases.

4.1.3 Brain age predictions

What does brain age actually measure? This question seems obvious at the first look. Nevertheless, the answer to this question is not so straightforward. Instead, additional details need to be addressed and additional definitions formulated. From the literature (see review [105], and the executed studies B & D in the present project), brain age appears as a general marker of health. While pathology and associated tissue changes lead to higher brain age values, brain age has a low level of specificity to any particular disorder. There are some examples with extremely clear difference in terms of brain age such as Alzheimer's Disease (and neurodegenerative disorders in general) compared to other brain and psychiatric disorders [105, 151].

Brain age limitations. While such findings are intriguing, there are several deeply rooted and unanswered questions attached to brain age. The potentially most prominent problem is the definition of brain age itself. First, while the chronological age provides a ground truth for predicting brain age, there is no ground truth for the brain age gap (the difference between brain age prediction and chronological age), which is meant to describe biological processes connected to age. Other machine learning problems might have hard ground truths such as disease labels, which is, however, not the case for brain age prediction models. Hence, the brain age gap on its own is limited and might rather be supplemental than a stand-alone marker. Second, brain age is dependent on age, and hence, the brain age gap is not necessarily an objective criterion for certain biological (e.g., ageing) processes. Instead, as brain age is usually based on cross-sectional data, both brain age and the brain age gap might simply represent a sample's (feature's) age-dependence, but not be telling for age*ing* processes [316].

Previously, it has been discussed that the sample-specific bias in brain age can be corrected after brain age calculations [74], optimally using training data age and brain age characteristics applied to the test set [72]. However, corrections for age-dependencies should potentially be executed on each feature *before* starting the training process, instead of executing post-hoc corrections on the already executed predictions

Discussion

(based on the test set). This might tease apart dependencies between age and brain age from dependencies between the brain and brain age. Such experiments are currently ongoing. The reason for this is that brain features are inherently age-dependent and hence biased [63, 72, 74]. This goes beyond the problems any post-hoc brain age bias-correction method can address, as the features used to train a model contain sample-specific age dependencies [63].

Another consideration, besides residualising brain features prior to the model training, is to treat training features completely differently: Namely, by subjecting the features to previously established normative models, which results in a set of expectancy or Z-scores. These new features, which are informed on large population-level trained models, can establish another training basis than commonly used features for brain age predictions. Such approach would transfer the age-dependency problem to another level of consideration, increasing the complexity of the modelling process and, hence, potentially decreasing the replication of such studies. The reason is that, now, a previously trained normative model needs to be employed to establish the brain age model training data. The training data of the normative model must fit the data which is to be used for the brain age modelling. This entire process increases researchers' degrees of freedom [286], characterised by, among others, arbitrary and subjective decisions. Questions such as which normative model to employ and how to employ the model must now be addressed, and the complexity of the research process is generally increased.

Other options are to change the label altogether and rather predict more clinically relevant scores [267], or their latent classes, allowing the prediction of meaningful units of multiple labels at the same time. Transfer [302] and composite learning approaches (i.e., combining algorithms [153, 320]) might help to assemble more complex models, which better depict the complex reality of the human body. Finally, the potentially most meaningful approach is to create multimodal data assembles which go beyond imaging-only data, but rather towards combining multiple data sources to establish integrative biomarkers, such as health registry data from different assessments, demographics, and imaging markers [3].

The age dependency of the brain age gap (BAG; the difference between brain age and chronological age) and how brain age can reflect brain ageing processes are not straightforward when examining relatively short-term changes in UK Biobank T_1-weighted data [316]. Based on longitudinal evidence from a stroke intervention [256], BAG shows time-sensitivity. This is contrasted by another study, including patients before and after stroke, which presented no time dependence of the BAG but instead an association of BAG with the extend of neurocognitive disorder / cognitive function [1]. Hence, one could assume that for brain changes to become detectable with brain age, these changes need to be fairly dramatic (either pathology or atrophy over a prolonged period of time).

Finally, brain age can be estimated from various modalities, such as different MRI sequences [60, 73], producing different estimates. Moreover, as we show in Study D, these brain ages are differently associated with age. Furthermore, various acquisition parameters influence brain age [145], which was also shown by the influence of field strength and indications for an influence of scan quality in Study C. Another factor influencing brain age estimates is the selection of features and their usage [145, 216].

There is no unique or strict answer on which algorithm to pick, beyond the indication that deep learning methods are useful for 2D and 3D data, and assemble algorithms, such as XGBoost, for tabular data. However, there is some evidence that XGBoost detects cognitive impairment with higher sensitivity [18]. In the case of brain age predictions from tabular data, the choice of algorithm remains arbitrary due to many researcher degrees of freedom, such as parameter and hyperparameter tuning choices, and the lack of ground truth for the brain age estimates. In other words, the many decisions about how to train models using very different algorithms can make comparisons (even using common metrics [72]) arbitrary. Moreover, beyond age distributions, variations in brain age might be either due to model error or biologically meaningful information. Hence, one could speculate that predictions with slightly higher error might simply indicate a higher sensitivity to the underlying biology, which would be in contrast to the engineering approach of building the best prediction model by reducing prediction error.

While variability in approaches of brain age model building is not a problem, the lack of reporting of such variability might give a wrong impression of homogeneity across brain age models. There are multiple aspects which influence brain age predictions: the underlying data, data processing and exclusions, feature engineering and selection procedures [216], algorithms [145], or brain age bias correction strategies [72, 74], just to mention a few. The result is that brain age studies are difficult to compare [72], and that brain age associations with phenotypes, estimated on the same sample, depend on the combination of pipeline and algorithm [145]. Yet, some brain age studies indicate that the same underlying data subjected to the same feature engineering will produce similar results when changing algorithms [72, 316] or when keeping the algorithm constant but changing the feature processing / engineering (Study A & B). Furthermore, the imaging modality and hence also the tissue influence both brain age predictions and its associations with different phenotypes [60, 73, 261, 331] (see also Study D). Finally, also acquisition parameters such as field strength influence brain age predictions [145] (see also Study C). Therefore, a more realistic depiction of brain age would be to mention the tissue and potentially the scanner and key acquisition and image processing parameters, such as "Siemens Skyra 3T Stejskal-Tanner pulse sequence FSL-TBSS skeleton average UK Biobank (WM) brain age", as obtained in multiple studies of this project. Other suggestions would involve the sharing of materials, code, models, and data as far as possible, in order to encourage and maximise reproducibility of the results. Only recently, repeated efforts focused on validating brain age models against each other [145, 216, 331]. Yet, there many more opportunities in order to improve brain age predictions and to build better models which reflect biological processes such as changing features, varying age ranges, or using recent machine learning approaches of confederate and ensemble learning or re-trainable and transfer-learning models.

Steps ahead for brain age. The outlined limitations of the concept of brain age do also restrict the process of translating the marker into clinical practice. However, multiple avenues can be taken to improve brain age models. One approach could be to move away from a single and general predictive model and rather towards a cluster of models. More specific models on a more narrow age-range might be more meaningful considering heterogeneous and non-linear ageing.

A complementary approach would be to establish brain age models based on

previously available information about certain disease states and their developments. For instance, if it is clear that only a set of features is relevant in a disorder group, one could train a model on healthy controls' data on the relevant features (e.g., metrics restricted to a set of brain regions). While there were no clinical samples included, an example of such approach was presented in Study D, where the conventional concept of brain age was extended by estimating brain age separately for each hemisphere. Such approaches can also be more specific and estimated on smaller brain regions, as done by Kaufmann and colleagues [151]. However, approaches such as the examined hemispheric brain age need to be further validated, tested for different disorders (such as unilateral stroke or epilepsy), on different datasets, and, potentially, using different levels of analysis (voxel-based and region averaged).

Bain age predictions have relied heavily on T_1-weighted MRI data, with less work focusing on dMRI. A specific gap to be filled in this context is the focus on voxel-based brain age predictions utilising dMRI data. One approach could be using CNNs on FA skeleton data to build brain WM-specific age models. New models should also include measures of uncertainty such as confidence intervals, as proposed by Hahn and colleagues [119], and focus on higher levels of model explainability (of the predictions) instead of only reducing prediction errors.

New models might also allow to better explain and add meaning to brain age estimates. Brain age seems to be influenced by multiple covariates which require careful mapping [145, 163]. Hence, brain age might also be considered as a dependent metric which becomes more meaningful when embedded in a structure of other markers. Thus, advanced statistical modelling, such as structural equation modelling might allow to better understand the multifactorial embedding of brain age and thereby its contributions to certain outcomes (see for an example study on the developmental trajectory in preterm neonates [183]). Brain age could also be used to improve clinical predictions, not only of diagnoses, but also prognoses for disease such as Alzheimer's, as brain age is associated with various neurodegenerative disorder biomarkers [54, 211].

Finally, more validation and replication work is necessary to establish generalisable models with the ability to robustly predict clinical outcome. A first step is to transparently report procedures in the respective studies and to share models, data, and code. The next step is to critically evaluate predictions both on a group level and on a single individual level, as done in Study C. Only after appropriate and stringent validation procedures, brain age models can be moved closer to make a true impact in clinical settings.

4.1.4 Imaging genetics

Imaging genetics comprises tools which allow the exploration and evaluation of genetic polymorphisms of the brain [40]. Such biological underpinnings of brain morphology (and its changes) are not apparent when observing associations of morphology and age or behaviour on their own, and can provide insights about potential mechanisms of the ageing brain associated with the development of disorders [53, 151, 311]. While there are multiple methods to assess brain-gene associations, we chose to examine associations between WM and polygenic risk scores (PGRS) of common psychiatric disorders as well as Alzheimer's Disease (Study E). PGRS describe a person's genetically determined

propensity to develop a certain brain disorder based on previously conducted genome-wide association studies (GWAS) [306]. Hence, PGRS leverage already conducted and well-informed studies to estimate risk scores in independent samples. Thereby, PGRS allow for direct inference without the necessity of genetic discovery or validation in the same study as would be necessary in GWAS [201], and without the loss of statistical power [58] as a result of sample splitting for validation.

We identified small effects, which is consistent with association strengths reported in the literature, particularly using WM-PGRS associations in the UK Biobank [8, 25, 136, 255, 283]. In the literature, such small effect were also observed when using T_1-weighted MRI [8, 101, 225], and both larger samples with different age distributions, examining, for example, attention deficit hyperactivity disorder (ADHD) PGRS-WM associations in adolescents [5], or young adults [293]. We do also not identify significant associations, which is a function of the observed small effect sizes paired with a conservative approach of adjusting the α-level (Bonferroni correction, Study E). The absence of significant effects was however also outlined in a previous review of schizophrenia PGRS with brain structure [312]. These findings show that mere disease PGRS associations with WM might be limited when examining largely healthy controls. Another argument might be that other, crucial third variables were not included in these analyses, such as behavioural measures, which will be subject of future investigations. However, we revealed an important effect of PGRS being stronger associated with the annual change in WM than cross-sectional measures. A previous study shows that longitudinal examinations associated higher PGRS with stronger WM degeneration, parameterised by DTI, in Parkinson's disease [116]. We made similar observations across brain regions. Yet, certain brain regions show stronger associations with PGRS of different disorders (Study E). Future work might, hence, focus on singular regions, other approaches to boost statistical power [310], use larger cohorts, and more embedded analysis strategies such as structural equation modelling.

4.2 Brain region-specific biomarkers

In each of the present studies, we presented age associations at different spatial scales. We benchmark age associations of global or whole-brain WM metrics by using the largest yet available cross-sectional and longitudinal sample, the UK Biobank. Our results provide a general idea of expectable trends when examining different samples, which might be smaller research or clinical samples. The observed trends include lower multi-compartment fractional anisotropy, intra-axonal water fraction, and kurtosis values, and higher axial, radial, mean diffusivity metrics, as well as free water at a higher age. Feasible treatment targets or biomarkers would optimally entail high spatial specificity. Hence, one possibility to establish biomarkers is to localise different markers specific to one brain region, to then establish predictive models based off these localisations [327]. However, several unsuccessful clinical explorations using electric stimulation of identified target areas [327] suggest that, in addition to appropriate models, a more nuanced understating of the biological process is necessary before implementing interventions. Here, we identified various associations of central and deep brain regions – corpus callosum, brain stem, cerebellum, the limbic system and the ventricles – with age, ageing and the polygenic risk of Alzheimer's Disease. In the

following, we will discuss the findings of the presented studies, outlining repeatedly central, deep brain areas, close to the ventricles as particularly related to ageing and polygenicity of psychiatric disorders and Alzheimer's disease. We will also discuss a brain region of which different aspects were repeatedly observed as strongly associated with age: the fornix, and how these regions might potentially become useful clinical targets.

4.2.1   Do central brain regions hold the key to ageing and disease?

Based on the observations of the presented studies and a large body of literature, one could expect that the presented deep brain regions close to the ventricles are both involved in ageing-associated processes as well as psychiatric and neurodegenerative disorders [26, 38, 65, 171, 250, 251, 332, 340].

For example, limbic structures, such as the hippocampus presents relatively high levels of adult neurogenesis potentially explaining repeated findings of the region's associations with psychiatric disorders and disorder states such as depression, anxiety, schizophrenia, addiction, and psychosis [243, 278], and neurdegenerative disorders, especially Alzheimer's disease (AD) [215], but also ageing in general [309]. Extending this view, AD has been frequently linked to limbic system degeneration [23, 79, 127, 156], which contains the hippocampus and other observed structures of importance in terms of age-associations, such as the amygdala and the fornix. Hence, ageing might follow a trajectory with more and less vulnerable brain regions, for example in accordance to the "last-in-first-out" hypothesis of retrogenesis, stating that brain areas with more complex structures and hence longer developmental periods in comparison to other brain regions are also more vulnerable to negative ageing effects, such as degeneration [100]. Translating the retrogenesis hypothesis to the presented findings could mean that central brain regions, which develop first should also be some of the last to be affected by age-related neurodegeneration. Neurodegenerating central brain areas might therefore be a sign of advanced ageing. If the last-in-first-out hypothesis is true, it still remains unclear whether last-in brain regions are also more vulnerable to lesions and pathology, and whether this vulnerability shows cumulative effects towards a higher age, and how this connects back to first-in brain areas [100].

Other possible explanations for limbic deterioration in healthy and diseased ageing can be found when examining ventricular function. Ventricle enlargements have repeatedly been shown as a general feature of ageing [38, 65, 171]. Importantly, ventricular function is strongly related to the development of neurodegenerative disorders and an important feature of this class of disorders [137, 252, 287, 332]. Potential links between ventricular enlargements and limbic structures, or generally structures near the ventricles, might be impaired glymphatic and vascular functioning influencing areas near the ventricles. This is also shown by relatively strong associations of cardiometabolic factors and white matter [30, 32] and with AD biomarkers [71], which is also reflected by cardiometabolic risk factor associations with WM brain age (Studies B and D).

The extend to which these degeneration phenomena can be connected with hemo-dynamics observed in younger samples requires further investigations. However, there several studies reporting transdiagnostic patterns of up-regulated limbic net-

works across psychiatric disorders [173] and more specifically hippocampal structure [15, 124, 180, 233] and limbic and corpus callosum WM [152, 164, 315] as transdiagnostic marker (which can already be observed in children [9]). Finally, also the cerebellum (in Study E as cerebellar peduncle) has been associated with psychiatric [218, 245] and brain disorders [178, 272]. This emerging evidence of the involvement of limbic, cerebellar and corpus callosum structure and function in neurodegeneration and psychopathology suggests the possibility of cortical transdiagnostic co-alterations being partly dependent on subcortical (including limbic) circuit alterations [124]. Moreover, there are observable differences between healthy controls and various psychiatric as well as brain disorder groups when comparing limbic and ventricular volumes [38]. The biological underpinnings and clinical implications of these differences and how they can be understood, for example, in the context of ageing and disorder-specificity, require further investigation.

Overall, our findings support previously reported results in outlining the importance of subcortical structures near the ventricles in ageing. AD fits well into a conceptualisation of accelerated ageing considering the large spatial overlap with the outlined structures. Psychiatric disorders show, however, larger variability and a more complex and dynamic pattern of associated brain regions, in terms of their structure and functions.

4.2.2   *The fornix – an ageing and disease biomarker?*

The fornix is a central, C-shaped structure close to the ventricles, and the inmost, or most central arch of the three arches of the limbic system [55]. Fornix is additionally a connecting region between the discussed limbic system and the frontal lobes [161]. The majority of the research questions of the presented cross-sectional and longitudinal UK Biobank Studies A and E in the context of this project was addressed by presenting age-association patterns of brain metrics and specifically white matter. These associations pointed out the fornix as one of the most age-associated region.

The fornix has also previously been found as highly age sensitive [333], to develop differently during Alzheimer's Disease compared to controls [14], and is already used as a therapeutic target when treating Alzheimer's Disease with deep brain (electric current) stimulation [176, 182, 257].

Previous studies have not only shown that white matter microstructural is sensitively associated with clinical staging in Alzheimer's disease outlining the importance of the fornix in these associations [333] but also that fornix degradation occurs earlier during diseased compared to healthy ageing [14]. Furthermore, white matter glial damage might precede and potentially cause hippocampal and other limbic GM atrophy [208]. Hence, for example, Alzheimer's disease related pathology, such as phospho-tau or microtubule-associated protein tau gene (MAPT) neuropathology in hippocampus is reflected in the fornix [246], rendering the fornix a potentially interesting treatment target. Indeed, there is a growing body of literature on deep brain stimulation of the fornix for Alzheimer's disease [176, 182, 257] as well as focused ultrasound [236] (see [206] for the general application of focused ultrasound). However, our findings require further replication to exclude that the fornix findings are due to misalignment or partial volume effects. If the presented findings hold, further advances are required in order

to identify earlier and less invasive treatments or even prognostics, for which the fornix might lend itself.

4.3 Ethical considerations

Several large scale data sets were used in the presented projects. The data collection procedures followed strict data protection policies and were approved by ethics committees, such as the Regional Committee for Medical and Health Research Ethics (REK, Norway), prior to their execution. Data management and privacy issues were additionally evaluated and approved by the University of Oslo leading to the use of the safe data environment / research server (TSD).

When doing research with human participants, there are several ethical considerations that should be made. These are especially important when using sensitive data such as brain images. Hence, we first discuss general ethical considerations concerning the utilised datasets. Second, we detail ethical considerations on the use of machine learning on medical data and clinical predictions.

As we used data sets collected in three different countries, also different procedures were applied to these datasets: Most datasets were only treated within TSD, whereas the Frequently Travelling Human Phantom (FTHP) dataset is openly available online [234]. FTHP data were processed locally, whereas all other datasets were processed within TSD. Only three utilised datasets, all used in Study C, were locally collected in Norway (Bergen and Oslo). All other data were collected in Germany (FTHP) and the United Kingdom (UK Biobank). The collections sides for these datasets were abroad and collections conducted by other researchers. Hence, our control over how ethical principles, such as data protection or participant information and rights, were followed during data collection is limited. Nevertheless, based on the ethics statements documented for each of the datasets [234, 292], it was and is still possible, in case of the UK Biobank study, for participants to withdraw at any time or to skip certain parts of the study. In case of the FTHP study, there was only one subject, a researcher himself, willing to share his data publicly, disallowing the opt-out option. All data handled with high confidentiality.

While this high level of confidentiality is important, there is also the question whether it can be evenly unethical to not make data openly available so that other researchers can use it [28]. The main motivation is to use data optimally in order to avoid wasting participant time and efforts as well as, usually public, resources. While MRI is safe, it can cause discomfort to participants: high levels of noise, a narrow space inside the scanner, and not being able to move for prolonged periods of times are standard features of the MRI scanning experience. Participants might also accept long scanning sessions which are actually outside their comfort zone just to complete the task they have started, also called completion bias [197, 339].

Allowing only a limited pool of researchers to access a dataset limits the extend to which the data are being used. TSD offers a solution, as researchers can here apply for access to datasets, following the specific ethical regulations of the datasets (some requiring multiple ethics admissions). This way, data can be shared with approved researchers while still being respectful of data protection issues.

Another highly relevant issue is the use of machine learning on health data, as

it was done here with brain age predictions. Machine learning models and their predictions depend on the characteristics of the training sample. Hence, imbalances, expressed in uneven sex, gender or ethnicity ratios can lead to poor representations of the population [209]. The following predictions might then contain biases, leading to misrepresentations and discrimination of the underrepresented group [204, 209]. Biased predictions can have multiple direct and indirect downstream consequences. Direct consequences could be poor or wrong predictions, and indirect consequences hidden decisions containing discriminatory bias, or systematically suggesting wrong treatments or disorders based on non-medical factors.

Questions which will become more and more pressing are how individual freedom will be influenced by commercial AI applications. This entails automatic profiling and decision making, and manipulation of users, for example by selective content presentations, as done on social media platforms. Considering that these issues are problematic in themselves, if they rely on biased assumptions, discrimination can be aggravated. Discrimination resulting from machine learning models' predictions can also be a consequence of the absence or poor model tuning and refining [204, 209]. Cross-validations (if possible across independent samples) and the monitoring of models and their performance are hence crucial steps in reducing undesirable consequences of machine learning applications (in the health sector but also other fields of application) [204].

Steps which aid monitoring and improving models is reducing black box components by making workflows (including models) as transparent and predictions as explainable as possible [17, 222, 318]. Hence, both components which are most weighted when making predictions, but also uncertainty around estimates help interpreting model output. The movement towards explainable AI is ongoing.

4.4 Conclusion and Outlook

In this project, we explored markers of ageing, including brain age as well as whole brain and region-averaged grey and white matter metrics. Particularly, the identification of multiple limbic and ventricle-near structures entails the potential for further exploration of biomarkers in clinical contexts. Each study can be extended and inform future studies in different ways. Moreover, there are multiple possibilities for following up the presented exploration with the goal of identifying biomarkers. In the following, we will give a non-comprehensive overview of potential future avenues.

Clinical advances as a direct outcome of this work are hidden by an attempt to firstly understand "healthy pathways" of ageing furthering attempts to identify clinical biomarkers. However, as outlined, our studies suggest that central and ventricle-near areas such as the fornix might be a useful prognostic and treatment targets for a range of disorders. In particular, in Alzheimer's disease different forms of brain stimulation including deep brain stimulation or focused ultrasound are already applied to stimulation sites such as the fornix. Moreover, the exploration of the vasculature of the identified regions close to the ventricles seem to be particularly affected by vascular disease. Hence, potentially, these regions are more sensitive to age, ageing, and ageing-related disease. The presented results on brain age offer multiple pathways to further develop brain age as a clinical marker. Yet, achieving the translation of the

brain age concept to the clinic and producing meaningful results still requires various methodological advances.

Methodological advances presented in this project include experiments and validation of brain age methods as well as suggestions for further study designs considering the different used spatial levels of analysis.

Brain age can be estimated using different algorithms, and different data. There are little constraints, and more individual and, for example, disorder-specific prediction strategies need to be explored, such as utilising regional or hemispheric brain age for unilateral stroke and tumour or general disorder detection. Furthermore, advanced models are required, embedding brain age and other markers to merge information for clinical meaningful predictions. Some suggestions would be to provide uncertainty estimates and more individual- or group-specific models. Additionally, there are many possibilities for feature and model architectural design. In the end, the most important part of the model is its explainability, and to, hence, maximise inference, following the simplest possible strategy is recommended (Occam's razor). Brain age models will also continuously need to be validated in different test and unseen (at best diverse) samples, not only as presented here with densely sampled data of few individuals, but also labelled clinical data providing an approximation of a ground truth. This will allow to better identify the model's predictive power and specificity across the heterogeneous phenotypes of patients in real-world clinical situations.

For better reproducibility, voxel-level analyses should be extended in order to not only create β-weighted maps or other representations of effect sizes, but also to include additional covariates such as body mass index, which have been repeatedly shown to influence brain structure. A more nuanced mapping of brain-body interactions and integrative biomarkers is an ongoing and large field of research. In terms of fundamental mechanisms, longitudinal changes in white matter asymmetries are largely unexplored in large samples and require further investigations.

Finally, the presented findings in this project could be strengthened by replication and validation in other independent samples, not necessarily matching age-range and health status, but potentially also deviating from such. In the following steps, the findings can then be examined in non-control or clinical groups and additional densely sampled individuals to move closer to clinical utility and the potential establishing of brain biomarkers.

BIBLIOGRAPHY

[1] E. B. Aamodt, D. Alnæs, A.-M. G. de Lange, S. Aam, T. Schellhorn, I. Saltvedt, M. K. Beyer, and L. T. Westlye. Longitudinal brain age prediction and cognitive function after stroke. *Neurobiology of Aging*, 122:55–64, 2023. 4.1.3

[2] A. Acheson, S. A. Wijtenburg, L. M. Rowland, A. Winkler, C. W. Mathias, L. E. Hong, N. Jahanshad, B. Patel, P. M. Thompson, S. A. McGuire, et al. Reproducibility of tract-based white matter microstructural measures using the enigma-dti protocol. *Brain and behavior*, 7(2):e00615, 2017. 4.1.2.1

[3] J. N. Acosta, G. J. Falcone, P. Rajpurkar, and E. J. Topol. Multimodal biomedical ai. *Nature Medicine*, 28(9):1773–1784, 2022. 4.1.3

[4] O. Agcaoglu, R. Miller, A. R. Mayer, K. Hugdahl, and V. D. Calhoun. Lateralization of resting state networks and relationship to age and gender. *Neuroimage*, 104:310–325, 2015. 1.3

[5] M. D. Albaugh, J. J. Hudziak, A. Ing, B. Chaarani, E. Barker, T. Jia, H. Lemaitre, R. Watts, C. Orr, P. A. Spechler, et al. White matter microstructure is associated with hyperactive/inattentive symptomatology and polygenic risk for attention-deficit/hyperactivity disorder in a population-based sample of adolescents. *Neuropsychopharmacology*, 44(9):1597–1603, 2019. 4.1.4

[6] S. Albawi, T. A. Mohammed, and S. Al-Zawi. Understanding of a convolutional neural network. In *2017 international conference on engineering and technology (ICET)*, pages 1–6. Ieee, 2017. 2.2.2

[7] F. Alfaro-Almagro, M. Jenkinson, N. K. Bangerter, J. L. Andersson, L. Griffanti, G. Douaud, S. N. Sotiropoulos, S. Jbabdi, M. Hernandez-Fernandez, E. Vallee, et al. Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. *Neuroimage*, 166:400–424, 2018. 2.3.1

[8] C. Alloza, M. Blesa-Cábez, M. Bastin, J. Madole, C. Buchanan, J. Janssen, J. Gibson, I. Deary, E. Tucker-Drob, H. Whalley, et al. Psychotic-like experiences, polygenic risk scores for schizophrenia, and structural properties of the salience, default mode, and central-executive networks in healthy participants from uk biobank. *Translational psychiatry*, 10(1):122, 2020. 4.1.4

[9] D. Alnæs, T. Kaufmann, N. T. Doan, A. Córdova-Palomera, Y. Wang, F. Bettella, T. Moberget, O. A. Andreassen, and L. T. Westlye. Association of heritable cognitive ability and psychopathology with white matter properties in children and adolescents. *JAMA psychiatry*, 75(3):287–295, 2018. 4.2.1

[10] S.-i. Amari. Backpropagation and stochastic gradient descent method. *Neurocomputing*, 5(4-5):185–196, 1993. 2.2.2, 2.2.2

[11] A. Amgalan, A. S. Maher, S. Ghosh, H. C. Chui, P. Bogdan, and A. Irimia. Brain age estimation reveals older adults' accelerated senescence after traumatic brain injury. *GeroScience*, 44(5):2509–2525, 2022. 1.2.1

[12] M. Anatürk, T. Kaufmann, J. H. Cole, S. Suri, L. Griffanti, E. Zsoldos, N. Filippini, A. Singh-Manoux, M. Kivimäki, L. T. Westlye, et al. Prediction of brain age and cognitive age: Quantifying brain and cognitive maintenance in aging. *Human brain mapping*, 42(6):1626–1640, 2021. 2.2.1

[13] M. Arain, M. Haque, L. Johal, P. Mathur, W. Nel, A. Rais, R. Sandhu, and S. Sharma. Maturation of the adolescent brain. *Neuropsychiatric disease and treatment*, pages 449–461, 2013. 1.1

[14] D. B. Archer, K. Schilling, N. Shashikumar, V. Jasodanand, E. E. Moore, K. R. Pechman, M. Bilgel, L. L. Beason-Held, Y. An, A. Shafer, et al. Leveraging longitudinal diffusion mri data to quantify differences in white matter microstructural decline in normal and abnormal aging. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 15(4):e12468, 2023. 1, 1.1, 4.2.2

[15] D. Arnone, A. McIntosh, K. Ebmeier, M. Munafò, and I. Anderson. Magnetic resonance imaging studies in unipolar depression: systematic review and meta-regression analyses. *European Neuropsychopharmacology*, 22(1):1–16, 2012. 4.2.1

[16] P. Arrondo, Ó. Elía-Zudaire, G. Martí-Andrés, M. A. Fernández-Seara, and M. Riverol. Grey matter changes on brain mri in subjective cognitive decline: a systematic review. *Alzheimer's Research & Therapy*, 14(1):1–16, 2022. 1.1

[17] A. Asatiani, P. Malo, P. R. Nagbøl, E. Penttinen, T. Rinta-Kahila, and A. Salovaara. Challenges of explaining the behavior of black-box ai systems. *MIS Quarterly Executive*, 19(4):259–278, 2020. 4.3

[18] E. Bacas, I. Kahhalé, P. R. Raamana, J. B. Pablo, A. S. Anand, and J. L. Hanson. Probing multiple algorithms to calculate brain age: Examining reliability, relations with demographics, and predictive power. *Human Brain Mapping*, 2023. 2.5, 2.2, 4.1.3

[19] M. Bach, F. B. Laun, A. Leemans, C. M. Tax, G. J. Biessels, B. Stieltjes, and K. H. Maier-Hein. Methodological considerations on tract-based spatial statistics (tbss). *Neuroimage*, 100:358–369, 2014. 4.1.2, 4.1.2.1

[20] L. Baecker, J. Dafflon, P. F. Da Costa, R. Garcia-Dias, S. Vieira, C. Scarpazza, V. D. Calhoun, J. R. Sato, A. Mechelli, and W. H. Pinaya. Brain age prediction: A comparison between machine learning models using region-and voxel-based morphometric data. *Human brain mapping*, 42(8):2332–2346, 2021. 2.2.1

[21] L. Baecker, R. Garcia-Dias, S. Vieira, C. Scarpazza, and A. Mechelli. Machine learning for brain age prediction: Introduction to methods and clinical applications. *EBioMedicine*, 72, 2021. 2.2.3

[22] V. Baliyan, C. J. Das, R. Sharma, and A. K. Gupta. Diffusion weighted imaging: technique and applications. *World journal of radiology*, 8(9):785, 2016. 2

[23] M. J. Ball. Limbic predilection in alzheimer dementia: is reactivated herpesvirus involved? *Canadian Journal of Neurological Sciences*, 9(3):303–306, 1982. 4.2.1

[24] P. L. Ballester, J. S. Suh, N. C. Ho, L. Liang, S. Hassel, S. C. Strother, S. R. Arnott, L. Minuzzi, R. B. Sassi, R. W. Lam, et al. Gray matter volume drives the brain age gap in schizophrenia: a shap study. *Schizophrenia*, 9(1):3, 2023. 1.2.1

[25] M. C. Barbu, Y. Zeng, X. Shen, S. R. Cox, T.-K. Clarke, J. Gibson, M. J. Adams, M. Johnstone, C. S. Haley, S. M. Lawrie, et al. Association of whole-genome and netrin1 signaling pathway–derived polygenic risk scores for major depressive disorder and white matter microstructure in the uk biobank. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 4(1):91–100, 2019. 4.1.4

[26] A. Bari, T. Niu, J.-P. Langevin, and I. Fried. Limbic neuromodulation: implications for addiction, posttraumatic stress disorder, and memory. *Neurosurgery Clinics*, 25(1):137–145, 2014. 4.2.1

[27] P. J. Basser, J. Mattiello, and D. LeBihan. Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994. 2.1.2.1, **??**, **??**

[28] H. Bauchner, R. M. Golub, and P. B. Fontanarosa. Data sharing: an ethical and scientific imperative. *Jama*, 315(12):1238–1240, 2016. 4.3

[29] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Biometrika*, 45(3-4):296–315, 1958. 2.1.2.3

[30] D. Beck, A.-M. G. De Lange, D. Alnæs, I. I. Maximov, M. L. Pedersen, O. D. Leinhard, J. Linge, R. Simon, G. Richard, K. M. Ulrichsen, et al. Adipose tissue distribution from body mri is associated with cross-sectional and longitudinal brain age in adults. *NeuroImage: Clinical*, 33:102949, 2022. 1.3, 2.2.1, 4.2.1

[31] D. Beck, A.-M. G. de Lange, I. I. Maximov, G. Richard, O. A. Andreassen, J. E. Nordvik, and L. T. Westlye. 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, 2021. 1.2, 2.1.2.2, 2.1.2.3, 2.2, 2.2.1

[32] D. Beck, A.-M. G. de Lange, M. L. Pedersen, D. Alnæs, I. I. Maximov, I. Voldsbekk, G. Richard, A.-M. Sanders, K. M. Ulrichsen, E. S. Dørum, et al. Cardiometabolic risk factors associated with brain age and accelerate brain ageing. *Human brain mapping*, 43(2):700–720, 2022. 1.2.2, 1.3, 4.2.1

[33] I. Beheshti, P. Gravel, O. Potvin, L. Dieumegarde, and S. Duchesne. A novel patch-based procedure for estimating brain age across adulthood. *Neuroimage*, 197:618–624, 2019. 4.1.2

[34] I. Beheshti, O. Potvin, and S. Duchesne. Patch-wise brain age longitudinal reliability. *Human Brain Mapping*, 42(3):690–698, 2021. 4.1.2

[35] C. M. Bennett, M. B. Miller, and G. L. Wolford. Neural correlates of interspecies perspective taking in the post-mortem atlantic salmon: An argument for multiple comparisons correction. *Neuroimage*, 47(Suppl 1):S125, 2009. 4.1

[36] A. Berger. Magnetic resonance imaging. *BMJ*, 324(7328):35, 2002. 2.1

[37] R. A. Bethlehem, C. Paquola, J. Seidlitz, L. Ronan, B. Bernhardt, K. A. Tsvetanov, C.-C. Consortium, et al. Dispersion of functional gradients across the adult lifespan. *Neuroimage*, 222:117299, 2020. 1.2

[38] R. A. Bethlehem, J. Seidlitz, S. R. White, J. W. Vogel, K. M. Anderson, C. Adamson, S. Adler, G. S. Alexopoulos, E. Anagnostou, A. Areces-Gonzalez, et al. Brain charts for the human lifespan. *Nature*, 604(7906):525–533, 2022. 1.1, 1.2, 1.2.1, 1.2.2, 4.2.1

[39] N. Bhagwat, A. Barry, E. W. Dickie, S. T. Brown, G. A. Devenyi, K. Hatano, E. DuPre, A. Dagher, M. Chakravarty, C. M. Greenwood, et al. Understanding the impact of preprocessing pipelines on neuroimaging cortical surface analyses. *GigaScience*, 10(1):giaa155, 2021. 2.1.1, 4.1.2.2

[40] K. L. Bigos and D. R. Weinberger. Imaging genetics—days of future past. *Neuroimage*, 53(3):804–809, 2010. 4.1.4

[41] N. A. Bishop, T. Lu, and B. A. Yankner. Neural mechanisms of ageing and cognitive decline. *Nature*, 464(7288):529–535, 2010. 1.1

[42] Y. Blinkouskaya, A. Caçoilo, T. Gollamudi, S. Jalalian, and J. Weickenmeier. Brain aging mechanisms with mechanical manifestations. *Mechanisms of ageing and development*, 200:111575, 2021. 1.2.1, 1.2.2, 1.2, 1.2.2

[43] R. Botvinik-Nezer, F. Holzmeister, C. F. Camerer, A. Dreber, J. Huber, M. Johannesson, M. Kirchler, R. Iwanir, J. A. Mumford, R. A. Adcock, et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, 582(7810):84–88, 2020. 2.1.1

[44] J. Bozek, L. Griffanti, S. Lau, and M. Jenkinson. Normative models for neuroimaging markers: Impact of model selection, sample size and evaluation criteria. *NeuroImage*, 268:119864, 2023. 1.1

[45] K. Brosch, F. Stein, S. Schmitt, J.-K. Pfarr, K. G. Ringwald, F. Thomas-Odenthal, T. Meller, O. Steinsträter, L. Waltemate, H. Lemke, et al. Reduced hippocampal gray matter volume is a common feature of patients with major depression, bipolar disorder, and schizophrenia spectrum disorders. *Molecular Psychiatry*, 27(10):4234–4243, 2022. 1.2.1

[46] S. N. Burke and C. A. Barnes. Neural plasticity in the ageing brain. *Nature reviews neuroscience*, 7(1):30–40, 2006. 1.2

[47] R. M. Califf. Biomarker definitions and their applications. *Experimental Biology and Medicine*, 243(3):213–221, 2018. 1

[48] B. J. Casey, T. Cannonier, M. I. Conley, A. O. Cohen, D. M. Barch, M. M. Heitzeg, M. E. Soules, T. Teslovich, D. V. Dellarco, H. Garavan, et al. The adolescent brain cognitive development (abcd) study: imaging acquisition across 21 sites. *Developmental cognitive neuroscience*, 32:43–54, 2018. 1.1

[49] L. Cecchetti and G. Handjaras. Reproducible brain-wide association studies do not necessarily require thousands of individuals. 2022. 4.1

[50] C. Chatfield. Model uncertainty, data mining and statistical inference. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 158(3):419–444, 1995. 4.1.2

[51] C.-M. Chen, P. Yang, M.-T. Wu, T.-C. Chuang, and T.-Y. Huang. Deriving and validating biomarkers associated with autism spectrum disorders from a large-scale resting-state database. *Scientific reports*, 9(1):9043, 2019. 1.3

[52] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016. 2.2.1

[53] W. Cheng, O. Frei, D. van der Meer, Y. Wang, K. S. O'Connell, Y. Chu, S. Bahrami, A. A. Shadrin, D. Alnæs, G. F. Hindley, et al. Genetic association between schizophrenia and cortical brain surface area and thickness. *JAMA psychiatry*, 78(9):1020–1030, 2021. 4.1.4

[54] U.-S. Choi, J. Y. Park, J. J. Lee, K. Y. Choi, S. Won, and K. H. Lee. Brain age prediction improves the early detection of alzheimer's disease in east asian elderly. *medRxiv*, pages 2023–02, 2023. 4.1.3

[55] Y. J. Choi, E. J. Lee, and J. E. Lee. The fornix: functional anatomy, normal neuroimaging, and various pathological conditions. *Investigative Magnetic Resonance Imaging*, 25(2):59–75, 2021. 4.2.2

[56] J. Chojdak-Łukasiewicz, E. Dziadkowiak, A. Zimny, and B. Paradowski. Cerebral small vessel disease: A review. *Advances in Clinical and Experimental Medicine*, 30(3):349–356, 2021. 1.2.2

[57] S. Coelho, S. H. Baete, G. Lemberskiy, B. Ades-Aron, G. Barrol, J. Veraart, D. S. Novikov, and E. Fieremans. Reproducibility of the standard model of diffusion in white matter on clinical mri systems. *NeuroImage*, 257:119290, 2022. 4.1.2, 4.1.2.1

[58] J. Cohen. Statistical power analysis. *Current directions in psychological science*, 1(3):98–101, 1992. 4.1.4

[59] S. Colangeli, M. Boccia, P. Verde, P. Guariglia, F. Bianchini, and L. Piccardi. Cognitive reserve in healthy aging and alzheimer's disease: a meta-analysis of fmri studies. *American Journal of Alzheimer's Disease & Other Dementias®*, 31(5):443–449, 2016. 1.1

[60] J. H. Cole. Multimodality neuroimaging brain-age in uk biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of aging*, 92:34–42, 2020. 1.3, 2.2, 4.1.3

[61] J. H. Cole, R. Leech, D. J. Sharp, and A. D. N. Initiative. Prediction of brain age suggests accelerated atrophy after traumatic brain injury. *Annals of neurology*, 77(4):571–581, 2015. 1.2.1

[62] J. H. Cole, R. E. Marioni, S. E. Harris, and I. J. Deary. Brain age and other bodily 'ages': implications for neuropsychiatry. *Molecular psychiatry*, 24(2):266–281, 2019. 1.1, 1.1, 1.3

[63] J. H. Cole, R. P. Poudel, D. Tsagkrasoulis, M. W. Caan, C. Steves, T. D. Spector, and G. Montana. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage*, 163:115–124, 2017. 1.3, 2.2, 2.2, 2.2.2, 4.1.2, 4.1.3

[64] J. H. Cole, J. Raffel, T. Friede, A. Eshaghi, W. J. Brownlee, D. Chard, N. De Stefano, C. Enzinger, L. Pirpamer, M. Filippi, et al. Longitudinal assessment of multiple sclerosis with the brain-age paradigm. *Annals of neurology*, 88(1):93–105, 2020. 1.2.1

[65] P. Coupé, J. V. Manjón, E. Lanuza, and G. Catheline. Lifespan changes of the human brain in alzheimer's disease. *Scientific reports*, 9(1):3998, 2019. 4.2.1

[66] S. R. Cox, M. Harris, S. J. Ritchie, C. Buchanan, M. d. C. Valdés Hernández, J. Corley, A. Taylor, J. W. Madole, S. Harris, H. Whalley, et al. Three major dimensions of human brain cortical ageing in relation to cognitive decline across the eighth decade of life. *Molecular psychiatry*, 26(6):2651–2662, 2021. 1.1

[67] F. Crivello, N. Tzourio-Mazoyer, C. Tzourio, and B. Mazoyer. Longitudinal assessment of global and regional rate of grey matter atrophy in 1,172 healthy older adults: modulation by sex and age. *PloS one*, 9(12):e114478, 2014. 1.2.1

[68] V. L. Cropley, P. Klauser, R. K. Lenroot, J. Bruggemann, S. Sundram, C. Bousman, A. Pereira, M. A. Di Biase, T. W. Weickert, C. S. Weickert, et al. Accelerated gray and white matter deterioration with age in schizophrenia. *American Journal of Psychiatry*, 174(3):286–295, 2017. 1.2.1, 1.2.2

[69] M. Dadar, A. L. Manera, S. Ducharme, and D. L. Collins. White matter hyperintensities are associated with grey matter atrophy and cognitive decline in alzheimer's disease and frontotemporal dementia. *Neurobiology of aging*, 111:54–63, 2022. 1.1

[70] A. Daducci, E. J. Canales-Rodríguez, H. Zhang, T. B. Dyrby, D. C. Alexander, and J.-P. Thiran. Accelerated microstructure imaging via convex optimization (amico) from diffusion mri data. *Neuroimage*, 105:32–44, 2015. 2.1.2.3

[71] H. E. Dark, C. Paterson, G. N. Daya, Z. Peng, M. R. Duggan, M. Bilgel, Y. An, A. Moghekar, C. Davatzikos, S. M. Resnick, et al. Proteomic indicators of health predict alzheimer's disease biomarker levels and dementia risk. *Annals of Neurology*, 2023. 4.2.1

[72] A.-M. G. de Lange, M. Anatürk, J. Rokicki, L. K. Han, K. Franke, D. Alnæs, K. P. Ebmeier, B. Draganski, T. Kaufmann, L. T. Westlye, et al. Mind the gap:

Performance metric evaluation in brain-age prediction. *Human Brain Mapping*, 43(10):3113–3129, 2022. 4.1.3

[73] A.-M. G. De Lange, M. Anatürk, S. Suri, T. Kaufmann, J. H. Cole, L. Griffanti, E. Zsoldos, D. E. Jensen, N. Filippini, A. Singh-Manoux, et al. Multimodal brain-age prediction and cardiovascular risk: The whitehall ii mri sub-study. *NeuroImage*, 222:117292, 2020. 2.2, 4.1.3

[74] A.-M. G. de Lange and J. H. Cole. Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clinical*, 26, 2020. 2.2.3, 4.1.3

[75] A.-M. G. de Lange, T. Kaufmann, D. van der Meer, L. A. Maglanoc, D. Alnæs, T. Moberget, G. Douaud, O. A. Andreassen, and L. T. Westlye. Population-based neuroimaging reveals traces of childbirth in the maternal brain. *Proceedings of the National Academy of Sciences*, 116(44):22341–22346, 2019. 2.2.1, 2.2.3

[76] C. Depp, T. Sun, A. O. Sasmita, L. Spieth, S. A. Berghoff, T. Nazarenko, K. Overhoff, A. A. Steixner-Kumar, S. Subramanian, S. Arinrad, et al. Myelin dysfunction drives amyloid-β deposition in models of alzheimer's disease. *Nature*, pages 1–9, 2023. 1.2, 1.2.2

[77] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006. 2.1.1

[78] D. D. DeSouza, M. Hodaie, and K. D. Davis. Structural magnetic resonance imaging can identify trigeminal system abnormalities in classical trigeminal neuralgia. *Frontiers in neuroanatomy*, 10:95, 2016. 2.4

[79] M. A. DeTure and D. W. Dickson. The neuropathological diagnosis of alzheimer's disease. *Molecular neurodegeneration*, 14(1):1–18, 2019. 4.2.1

[80] C. G. DeYoung, T. A. Sassenberg, R. Abend, T. Allen, R. Beaty, M. Bellgrove, S. D. Blain, D. Bzdok, R. S. Chavez, S. A. Engel, et al. Reproducible between-person brain-behavior associations do not always require thousands of individuals. 2022. 4.1

[81] B. Dhital, M. Reisert, E. Kellner, and V. G. Kiselev. Intra-axonal diffusivity in brain white matter. *NeuroImage*, 189:543–550, 2019. 2.1.2.3

[82] T. Dhollander, A. Clemente, M. Singh, F. Boonstra, O. Civier, J. D. Duque, N. Egorova, P. Enticott, I. Fuelscher, S. Gajamange, et al. Fixel-based analysis of diffusion mri: methods, applications, challenges and opportunities. *Neuroimage*, 241:118417, 2021. 4.1.2.1

[83] M. A. Di Biase, Y. E. Tian, R. A. Bethlehem, J. Seidlitz, A. F. Alexander-Bloch, B. T. Yeo, and A. Zalesky. Mapping human brain charts cross-sectionally and longitudinally. *Proceedings of the National Academy of Sciences*, 120(20):e2216798120, 2023. 1.3

[84] R. P. Dörfel, J. M. Arenas-Gomez, P. M. Fisher, M. Ganz, G. M. Knudsen, J. Svensson, and P. Plavén-Sigray. Prediction of brain age using structural magnetic resonance imaging: A comparison of accuracy and test-retest reliability of publicly available software packages. *BioRxiv*, pages 2023–01, 2023. 2.2.2

[85] J. Duan, Y. Wei, F. Y. Womer, X. Zhang, M. Chang, Y. Zhu, Z. Liu, C. Li, Z. Yin, R. Zhang, et al. Neurobiological substrates of major psychiatry disorders: transdiagnostic associations between white matter abnormalities, neuregulin 1 and clinical manifestation. *Journal of Psychiatry and Neuroscience*, 46(5):E506–E515, 2021. 1.2.2

[86] B. N. Dugger and D. W. Dickson. Pathology of neurodegenerative diseases. *Cold Spring Harbor perspectives in biology*, 9(7):a028035, 2017. 1.2.2

[87] J. Dukart, S. Holiga, M. Rullmann, R. Lanzenberger, P. C. Hawkins, M. A. Mehta, S. Hesse, H. Barthel, O. Sabri, R. Jech, et al. Juspace: A tool for spatial correlation analyses of magnetic resonance imaging data with nuclear imaging derived neurotransmitter maps. Technical report, Wiley Online Library, 2021. 4.1.2.2

[88] K. Eckle and J. Schmidt-Hieber. A comparison of deep networks with relu activation function and linear spline-type methods. *Neural Networks*, 110:232–242, 2019. 2.2.2

[89] S. B. Eickhoff, B. T. Yeo, and S. Genon. Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience*, 19(11):672–686, 2018. 2.1.1

[90] A. Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4, 1905. 2.1.2

[91] A. Eklund, T. E. Nichols, and H. Knutsson. Cluster failure: Why fmri inferences for spatial extent have inflated false-positive rates. *Proceedings of the national academy of sciences*, 113(28):7900–7905, 2016. 4.1, 4.1.2

[92] F. N. Emamzadeh and A. Surguchov. Parkinson's disease: biomarkers, treatment, and risk factors. *Frontiers in neuroscience*, 12:612, 2018. 1

[93] T. Espeseth, A. Christoforou, A. J. Lundervold, V. M. Steen, S. Le Hellard, and I. Reinvang. Imaging and cognitive genetics: the norwegian cognitive neurogenetics sample. *Twin Research and Human Genetics*, 15(3):442–452, 2012. 2.3.4

[94] O. Esteban, D. Birman, M. Schaer, O. O. Koyejo, R. A. Poldrack, and K. J. Gorgolewski. Mriqc: Advancing the automatic prediction of image quality in mri from unseen sites. *PloS one*, 12(9):e0184661, 2017. 2.1.1

[95] O. Esteban, C. J. Markiewicz, R. W. Blair, C. A. Moodie, A. I. Isik, A. Erramuzpe, J. D. Kent, M. Goncalves, E. DuPre, M. Snyder, et al. fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, 16(1):111–116, 2019. 2.1.1

[96] A. M. Ferreira, F. Costa, A. Tralhão, H. Marques, N. Cardim, and P. Adragão. Mri-conditional pacemakers: current perspectives. *Medical Devices: Evidence and Research*, pages 115–124, 2014. 2

[97] E. Fieremans, J. H. Jensen, and J. A. Helpern. White matter characterization with diffusional kurtosis imaging. *Neuroimage*, 58(1):177–188, 2011. 2.1.2.1, **??**, 2.1.2.3, **??**

[98] M. Filippi, W. Brück, D. Chard, F. Fazekas, J. J. Geurts, C. Enzinger, S. Hametner, T. Kuhlmann, P. Preziosa, À. Rovira, et al. Association between pathological and mri findings in multiple sclerosis. *The Lancet Neurology*, 18(2):198–210, 2019. 1

[99] B. Fischl. Freesurfer. *Neuroimage*, 62(2):774–781, 2012. 2.1.1, 2.1.1, 4.1.2.1, 4.1.2.2

[100] A. M. Fjell, L. McEvoy, D. Holland, A. M. Dale, K. B. Walhovd, A. D. N. Initiative, et al. What is normal in normal aging? effects of aging, amyloid and alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in neurobiology*, 117:20–40, 2014. 4.2.1

[101] H. Foo, A. Thalamuthu, J. Jiang, F. Koch, K. A. Mather, W. Wen, and P. S. Sachdev. Associations between alzheimer's disease polygenic risk scores and hippocampal subfield volumes in 17,161 uk biobank participants. *Neurobiology of Aging*, 98:108–115, 2021. 4.1.4

[102] A. Fornito, A. Zalesky, and M. Breakspear. The connectomics of brain disorders. *Nature Reviews Neuroscience*, 16(3):159–172, 2015. 1.1

[103] K. Förster, R. H. Horstmann, U. Dannlowski, J. Houenou, and P. Kanske. Progressive grey matter alterations in bipolar disorder across the life span–a systematic review. *Bipolar Disorders*, 2023. 1.2.1

[104] K. Franke and C. Gaser. Longitudinal changes in individual brainage in healthy aging, mild cognitive impairment, and alzheimer's disease. *GeroPsych*, 2012. 1.2.1

[105] K. Franke and C. Gaser. Ten years of brainage as a neuroimaging biomarker of brain aging: what insights have we gained? *Frontiers in neurology*, page 789, 2019. 1.1, 1.3, 2.2, 4.1.3

[106] A. Fry, T. J. Littlejohns, C. Sudlow, N. Doherty, L. Adamska, T. Sprosen, R. Collins, and N. E. Allen. Comparison of sociodemographic and health-related characteristics of uk biobank participants with those of the general population. *American journal of epidemiology*, 186(9):1026–1034, 2017. 4.1.1

[107] R. K. Gebre, M. L. Senjem, S. Raghavan, C. G. Schwarz, J. L. Gunter, E. I. Hofrenning, R. I. Reid, K. Kantarci, J. Graff-Radford, D. S. Knopman, et al. Cross–scanner harmonization methods for structural mri may need further work: A comparison study. *Neuroimage*, 269:119912, 2023. 4.1.2.2

[108] M. Ghadimi and A. Sapra. Magnetic resonance imaging contraindications. 2019. 2

[109] A. Giorgio, L. Santelli, V. Tomassini, R. Bosnell, S. Smith, N. De Stefano, and H. Johansen-Berg. Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage*, 51(3):943–951, 2010. 1.2.2

[110] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016. 4.1.2.2

[111] E. Gokcal, M. J. Horn, S. J. van Veluw, A. Frau-Pascual, A. S. Das, M. Pasi, P. Fotiadis, A. D. Warren, K. Schwab, J. Rosand, et al. Lacunes, microinfarcts, and vascular dysfunction in cerebral amyloid angiopathy. *Neurology*, 96(12):e1646–e1654, 2021. 1.2.2

[112] M. R. Gold, T. Sommer, J. Schwitter, A. Al Fagih, T. Albert, B. Merkely, M. Peterson, A. Ciuffo, S. Lee, L. Landborg, et al. Full-body mri in patients with an implantable cardioverter-defibrillator: primary results of a randomized study. *Journal of the American College of Cardiology*, 65(24):2581–2588, 2015. 2

[113] R. G. González. Clinical mri of acute ischemic stroke. *Journal of Magnetic Resonance Imaging*, 36(2):259–271, 2012. 1

[114] E. M. Gordon, T. O. Laumann, A. W. Gilmore, D. J. Newbold, D. J. Greene, J. J. Berg, M. Ortega, C. Hoyt-Drazen, C. Gratton, H. Sun, et al. Precision functional mapping of individual human brains. *Neuron*, 95(4):791–807, 2017. 1.2.3, 1.3, 4.1, 4.1.1

[115] E. C. Grossner, R. A. Bernier, E. K. Brenner, K. S. Chiou, and F. G. Hillary. Prefrontal gray matter volume predicts metacognitive accuracy following traumatic brain injury. *Neuropsychology*, 32(4):484, 2018. 2.1

[116] L. Gu, X. Guan, T. Gao, C. Zhou, W. Yang, D. Lv, J. Wu, Y. Fang, T. Guo, Z. Song, et al. The effect of polygenic risk on white matter microstructural degeneration in parkinson's disease: A longitudinal diffusion tensor imaging study. *European Journal of Neurology*, 29(4):1000–1010, 2022. 4.1.4

[117] L. Gu and Z. Zhang. Exploring structural and functional brain changes in mild cognitive impairment: a whole brain ale meta-analysis for multimodal mri. *ACS chemical neuroscience*, 10(6):2823–2829, 2019. 1.1

[118] L. V. Haar, T. Elvira, and O. Ochoa. An analysis of explainability methods for convolutional neural networks. *Engineering Applications of Artificial Intelligence*, 117:105606, 2023. 4.1.2

[119] T. Hahn, J. Ernsting, N. R. Winter, V. Holstein, R. Leenings, M. Beisemann, L. Fisch, K. Sarink, D. Emden, N. Opel, et al. An uncertainty-aware, shareable, and transparent neural network architecture for brain-age modeling. *Science advances*, 8(1):eabg9471, 2022. 4.1.3

[120] J. Han, S. Y. Kim, J. Lee, and W. H. Lee. Brain age prediction: A comparison between machine learning models using brain morphometric data. *Sensors*, 22(20):8077, 2022. 2.2, 2.2.1, 4.1.2.2

[121] M. J. Hawrylycz, E. S. Lein, A. L. Guillozet-Bongaarts, E. H. Shen, L. Ng, J. A. Miller, L. N. Van De Lagemaat, K. A. Smith, A. Ebbert, Z. L. Riley, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489(7416):391–399, 2012. 4.1.2.2

[122] E. P. Hedges, M. Dimitrov, U. Zahid, B. B. Vega, S. Si, H. Dickson, P. McGuire, S. Williams, G. J. Barker, and M. J. Kempton. Reliability of structural mri measurements: The effects of scan session, head tilt, inter-scan interval, acquisition sequence, freesurfer version and processing stream. *Neuroimage*, 246:118751, 2022. 4.1.2.2

[123] J. Henrich, S. J. Heine, and A. Norenzayan. Most people are not weird. *Nature*, 466(7302):29–29, 2010. 4.1.1

[124] M. Hettwer, S. Larivière, B. Park, O. A. van den Heuvel, L. Schmaal, O. Andreassen, C. Ching, M. Hoogman, J. Buitelaar, D. van Rooij, et al. Coordinated cortical thickness alterations across six neurodevelopmental and psychiatric disorders. *Nature communications*, 13(1):6851, 2022. 4.2.1

[125] R. A. Hill, A. M. Li, and J. Grutzendler. Lifelong cortical myelin plasticity and age-related degeneration in the live mammalian brain. *Nature neuroscience*, 21(5):683–695, 2018. 1.2.2

[126] E. A. Høgestøl, T. Kaufmann, G. O. Nygaard, M. K. Beyer, P. Sowa, J. E. Nordvik, K. Kolskår, G. Richard, O. A. Andreassen, H. F. Harbo, et al. Cross-sectional and longitudinal mri brain scans reveal accelerated brain aging in multiple sclerosis. *Frontiers in neurology*, 10:450, 2019. 1.2.1

[127] M. Hopper and F. Vogel. The limbic system in alzheimer's disease. a neuropathologic investigation. *The American journal of pathology*, 85(1):1, 1976. 4.2.1

[128] C. Horien, S. Noble, A. S. Greene, K. Lee, D. S. Barron, S. Gao, D. O'Connor, M. Salehi, J. Dadashkarimi, X. Shen, et al. A hitchhiker's guide to working with large, open-source neuroimaging datasets. *Nature human behaviour*, 5(2):185–193, 2021. 1.1, 1.2.1

[129] K. Hua, J. Zhang, S. Wakana, H. Jiang, X. Li, D. S. Reich, P. A. Calabresi, J. J. Pekar, P. C. van Zijl, and S. Mori. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage*, 39(1):336–347, 2008. 2.1.2.4

[130] W. Huang, X. Li, H. Li, W. Wang, K. Chen, K. Xu, J. Zhang, Y. Chen, D. Wei, N. Shu, et al. Accelerated brain aging in amnestic mild cognitive impairment: relationships with individual cognitive decline, risk factors for alzheimer disease, and clinical progression. *Radiology: Artificial Intelligence*, 3(5):e200171, 2021. 1.2.1

[131] D. E. Huber, K. W. Potter, and L. D. Huszar. Less "story" and more "reliability" in cognitive neuroscience. *Cortex; a journal devoted to the study of the nervous system and behavior*, 113:347, 2019. 4.1

[132] M. Huber, J. A. Knottnerus, L. Green, H. Van Der Horst, A. R. Jadad, D. Kromhout, B. Leonard, K. Lorig, M. I. Loureiro, J. W. Van der Meer, et al. How should we define health? *Bmj*, 343, 2011. 1.1

[133] M. A. Ibrahim, B. Hazhirkarzar, and A. B. Dublin. Gadolinium magnetic resonance imaging. 2018. 2

[134] C. R. Jack, D. S. Knopman, W. J. Jagust, L. M. Shaw, P. S. Aisen, M. W. Weiner, R. C. Petersen, and J. Q. Trojanowski. Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade. *The Lancet Neurology*, 9(1):119–128, 2010. 1

[135] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, et al. The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(4):685–691, 2008. 2.3.3

[136] B. M. Jacobs, C. Watson, C. Marshall, A. Noyce, and R. Dobson. No evidence for association between polygenic risk of multiple sclerosis and mri phenotypes in˜ 30,000 healthy adult uk biobank participants. *Multiple Sclerosis Journal*, 28(10):1656–1657, 2022. 4.1.4

[137] D. Jaraj, S. Agerskov, K. Rabiei, T. Marlow, C. Jensen, X. Guo, S. Kern, C. Wikkelsø, and I. Skoog. Vascular factors in suspected normal pressure hydrocephalus: a population-based study. *Neurology*, 86(7):592–599, 2016. 4.2.1

[138] I. O. Jelescu and M. D. Budde. Design and validation of diffusion mri models of white matter. *Frontiers in physics*, 5:61, 2017. 2.1.2.1, 2.1.2.2, 2.1.2.2, 2.1.2.3

[139] I. O. Jelescu, J. Veraart, V. Adisetiyo, S. S. Milla, D. S. Novikov, and E. Fieremans. One diffusion acquisition and different white matter models: how does microstructure change in human early development based on wmti and noddi? *Neuroimage*, 107:242–256, 2015. 2.1.2.3

[140] I. O. Jelescu, J. Veraart, E. Fieremans, and D. S. Novikov. Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR in Biomedicine*, 29(1):33–47, 2016. 2.1.2.3

[141] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith. Fsl. *Neuroimage*, 62(2):782–790, 2012. 2.1.1

[142] J. H. Jensen, J. A. Helpern, A. Ramani, H. Lu, and K. Kaczynski. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005. ??, 2.1.2.2, ??

[143] S. N. Jespersen, J. L. Olesen, B. Hansen, and N. Shemesh. Diffusion time dependence of microstructural parameters in fixed spinal cord. *Neuroimage*, 182:329–342, 2018. 2.1.2.3

[144] X. Ji, F. Yuan, and Y. Zhao. What we know and what we don't know about the proton spin after 30 years. *Nature Reviews Physics*, 3(1):27–38, 2021. 2.1

[145] R. J. Jirsaraie, T. Kaufmann, V. Bashyam, G. Erus, J. L. Luby, L. T. Westlye, C. Davatzikos, D. M. Barch, and A. Sotiras. Benchmarking the generalizability of brain age models: challenges posed by scanner variance and prediction bias. *Human Brain Mapping*, 44(3):1118–1128, 2023. 2.2, 2.2, 2.2.1, 2.2.2, 2.2.3, 4.1.2, 4.1.3

[146] J. Jovicich, M. Marizzoni, B. Bosch, D. Bartrés-Faz, J. Arnold, J. Benninghoff, J. Wiltfang, L. Roccatagliata, A. Picco, F. Nobili, et al. Multisite longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging of healthy elderly subjects. *Neuroimage*, 101:390–403, 2014. 4.1.2.1

[147] P. M. Jungmann, C. A. Agten, C. W. Pfirrmann, and R. Sutter. Advances in mri around metal. *Journal of magnetic resonance imaging*, 46(4):972–991, 2017. 2

[148] E. Kaden, N. D. Kelm, R. P. Carson, M. D. Does, and D. C. Alexander. Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139:346–359, 2016. 2.1.2.1, **??**, 2.1.2.3, **??**

[149] E. Kaden, F. Kruggel, and D. C. Alexander. Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic resonance in medicine*, 75(4):1752–1763, 2016. 2.1.2.1, **??**, 2.1.2.3, **??**

[150] V. R. Karolis, M. Corbetta, and M. Thiebaut de Schotten. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nature communications*, 10(1):1417, 2019. 1.3

[151] T. Kaufmann, D. van der Meer, N. T. Doan, E. Schwarz, M. J. Lund, I. Agartz, D. Alnæs, D. M. Barch, R. Baur-Streubel, A. Bertolino, et al. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature neuroscience*, 22(10):1617–1623, 2019. 1, 1.2.1, 2.6, 2.2.3, 4.1.3, 4.1.4

[152] S. Kelly, N. Jahanshad, A. Zalesky, P. Kochunov, I. Agartz, C. Alloza, O. Andreassen, C. Arango, N. Banaj, S. Bouix, et al. Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the enigma schizophrenia dti working group. *Molecular psychiatry*, 23(5):1261–1269, 2018. 4.1.2.1, 4.2.1

[153] F. Khozeimeh, D. Sharifrazi, N. H. Izadi, J. H. Joloudari, A. Shoeibi, R. Alizadehsani, J. M. Gorriz, S. Hussain, Z. A. Sani, H. Moosaei, et al. Combining a convolutional neural network with autoencoders to predict the survival chance of covid-19 patients. *Scientific Reports*, 11(1):15343, 2021. 4.1.3

[154] M.-J. Kim, J.-H. Lee, F. Juarez Anaya, J. Hong, W. Miller, S. Telu, P. Singh, M. Y. Cortes, K. Henry, G. L. Tye, et al. First-in-human evaluation of [11 c] ps13, a novel pet radioligand, to quantify cyclooxygenase-1 in the brain. *European journal of nuclear medicine and molecular imaging*, 47:3143–3151, 2020. 4.1.2.2

[155] K. M. King, A. K. Littlefield, C. J. McCabe, K. L. Mills, J. Flournoy, and L. Chassin. Longitudinal modeling in developmental neuroimaging research: Common challenges, and solutions from developmental psychology. *Developmental cognitive neuroscience*, 33:54–72, 2018. 1.2.3

[156] S. Knafo. Amygdala in alzheimer's disease. *The Amygdala–A Discrete Multitasking Manager. IntechOpen*, pages 375–383, 2012. 4.2.1

[157] F. Knolle, S. S. Arumugham, R. A. Barker, M. W. Chee, A. Justicia, N. Kamble, J. Lee, S. Liu, A. Lenka, S. J. Lewis, et al. A multicentre study on grey matter morphometric biomarkers for classifying early schizophrenia and parkinson's disease psychosis. *npj Parkinson's Disease*, 9(1):87, 2023. 1.2.1

[158] X.-Z. Kong, S. R. Mathias, T. Guadalupe, E. L. W. Group, D. C. Glahn, B. Franke, F. Crivello, N. Tzourio-Mazoyer, S. E. Fisher, P. M. Thompson, et al. Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *Proceedings of the National Academy of Sciences*, 115(22):E5154–E5163, 2018. 1.3

[159] X.-Z. Kong, M. C. Postema, T. Guadalupe, C. de Kovel, P. S. Boedhoe, M. Hoogman, S. R. Mathias, D. Van Rooij, D. Schijven, D. C. Glahn, et al. Mapping brain asymmetry in health and disease through the enigma consortium. *Human brain mapping*, 43(1):167–181, 2022. 1.3

[160] M. Korbmacher, F. Azevedo, C. R. Pennington, H. Hartmann, M. Pownall, K. Schmidt, M. Elsherif, N. Breznau, O. Robertson, T. Kalandadze, et al. The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology*, 1(1):3, 2023. 4.1

[161] M. Korbmacher, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, and I. I. Maximov. Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing. *Human brain mapping*, 2023. 4.2.2

[162] M. Korbmacher, D. van der Meer, D. Beck, A.-M. G. de Lange, E. Eikefjord, A. Lundervold, O. Andreassen, L. T. Westlye, and I. I. Maximov. Brain asymmetries from midlife to old adulthood and hemispheric brain age. *bioRxiv*, pages 2023–08, 2023. 1.2.2

[163] M. Korbmacher, M.-Y. Wang, R. Eikeland, R. Buchert, O. A. Andreassen, T. Espeseth, E. Leonardsen, L. T. Westlye, I. I. Maximov, and K. Specht. Considerations on brain age predictions from repeatedly sampled data across time. *Brain and Behavior*, page e3219, 2023. 4.1.3

[164] D. Koshiyama, M. Fukunaga, N. Okada, K. Morita, K. Nemoto, K. Usui, H. Yamamori, Y. Yasuda, M. Fujimoto, N. Kudo, et al. White matter microstructural alterations across four major psychiatric disorders: mega-analysis study in 2937 individuals. *Molecular psychiatry*, 25(4):883–895, 2020. 1.2.2, 4.2.1

[165] S. Krantic, N. Mechawar, S. Reix, and R. Quirion. Molecular basis of programmed cell death involved in neurodegeneration. *Trends in neurosciences*, 28(12):670–676, 2005. 1.2.1

[166] D. L. Kurtin, V. Giunchiglia, J. Vohryzek, J. Cabral, A. C. Skeldon, and I. R. Violante. Moving from phenomenological to predictive modelling: Progress and pitfalls of modelling brain stimulation in-silico. *Neuroimage*, 272:120042, 2023. 1.1

[167] T. O. Laumann, E. M. Gordon, B. Adeyemo, A. Z. Snyder, S. J. Joo, M.-Y. Chen, A. W. Gilmore, K. B. McDermott, S. M. Nelson, N. U. Dosenbach, et al. Functional system and areal organization of a highly sampled individual human brain. *Neuron*, 87(3):657–670, 2015. 1.3

[168] D.-K. Lee, H. Lee, K. Park, E. Joh, C.-E. Kim, and S. Ryu. Common gray and white matter abnormalities in schizophrenia and bipolar disorder. *PLoS One*, 15(5):e0232826, 2020. 1.2.1

[169] J. Lee and H.-J. Kim. Normal aging induces changes in the brain and neurodegeneration progress: review of the structural, biochemical, metabolic, cellular, and molecular changes. *Frontiers in Aging Neuroscience*, 14:931536, 2022. 1.2

[170] P.-L. Lee, C.-Y. Kuo, P.-N. Wang, L.-K. Chen, C.-P. Lin, K.-H. Chou, and C.-P. Chung. Regional rather than global brain age mediates cognitive function in cerebral small vessel disease. *Brain Communications*, 4(5):fcac233, 2022. 1.3

[171] E. H. Leonardsen, H. Peng, T. Kaufmann, I. Agartz, O. A. Andreassen, E. G. Celius, T. Espeseth, H. F. Harbo, E. A. Høgestøl, A.-M. De Lange, et al. Deep neural networks learn general and clinically relevant representations of the ageing brain. *NeuroImage*, 256:119210, 2022. 1.3, 2.2, 2.2, 2.2.2, 2.8, 2.2.2, 4.1.2, 4.2.1

[172] A. D. Leow, I. Yanovsky, N. Parikshak, X. Hua, S. Lee, A. W. Toga, C. R. Jack Jr, M. A. Bernstein, P. J. Britson, J. L. Gunter, et al. Alzheimer's disease neuroimaging initiative: a one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. *Neuroimage*, 45(3):645–655, 2009. 1

[173] C. Li, M. Dong, F. Y. Womer, S. Han, Y. Yin, X. Jiang, Y. Wei, J. Duan, R. Feng, L. Zhang, et al. Transdiagnostic time-varying dysconnectivity across major psychiatric disorders. *Human brain mapping*, 42(4):1182–1196, 2021. 4.2.1

[174] H.-J. Li, X.-H. Hou, H.-H. Liu, C.-L. Yue, G.-M. Lu, and X.-N. Zuo. Putting age-related task activation into large-scale brain networks: a meta-analysis of 114 fmri studies on healthy aging. *Neuroscience & Biobehavioral Reviews*, 57:156–174, 2015. 1.1

[175] P. Li, E. Ensink, S. Lang, L. Marshall, M. Schilthuis, J. Lamp, I. Vega, and V. Labrie. Hemispheric asymmetry in the human brain and in parkinson's disease is linked to divergent epigenetic patterns in neurons. *Genome biology*, 21(1):1–23, 2020. 1.3

[176] R. Li, C. Zhang, Y. Rao, and T.-F. Yuan. Deep brain stimulation of fornix for memory improvement in alzheimer's disease: A critical review. *Ageing research reviews*, 79:101668, 2022. 4.2.2

[177] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 2021. 2.2.2

[178] K. J. Liang and E. S. Carlson. Resistance, vulnerability and resilience: A review of the cognitive cerebellum in aging and neurodegenerative diseases. *Neurobiology of Learning and Memory*, 170:106981, 2020. 4.2.1

[179] Z.-P. Liang and P. C. Lauterbur. *Principles of magnetic resonance imaging*. SPIE Optical Engineering Press Belllingham, WA, 2000. 2.1

[180] J. Lieberman, R. Girgis, G. Brucato, H. Moore, F. Provenzano, L. Kegeles, D. Javitt, J. Kantrowitz, M. Wall, C. Corcoran, et al. Hippocampal dysfunction in the pathophysiology of schizophrenia: a selective review and hypothesis for early detection and intervention. *Molecular psychiatry*, 23(8):1764–1772, 2018. 4.2.1

[181] S.-L. Liew, N. Schweighofer, J. H. Cole, A. Zavaliangos-Petropulu, B. P. Lo, L. K. Han, T. Hahn, L. Schmaal, M. R. Donnelly, J. N. Jeong, et al. Association of brain age, lesion volume, and functional outcome in patients with stroke. *Neurology*, 100(20):e2103–e2113, 2023. 1

[182] H. Liu, Y. Temel, J. Boonstra, and S. Hescham. The effect of fornix deep brain stimulation in brain diseases. *Cellular and Molecular Life Sciences*, 77:3279–3291, 2020. 4.2.2

[183] M. Liu, S. Kim, B. Duffy, S. Yuan, J. H. Cole, A. W. Toga, N. Jahanshad, A. J. Barkovich, D. Xu, and H. Kim. Brain age predicted using graph convolutional neural network explains developmental trajectory in preterm neonates. *bioRxiv*, pages 2021–05, 2021. 4.1.3

[184] S. Liu, A. Abdellaoui, K. J. Verweij, and G. A. van Wingen. Replicable brain–phenotype associations require large-scale neuroimaging data. *Nature Human Behaviour*, pages 1–13, 2023. 1.1, 4.1

[185] C. Lois, I. González, D. Izquierdo-García, N. R. Zurcher, P. Wilkens, M. L. Loggia, J. M. Hooker, and H. D. Rosas. Neuroinflammation in huntington's disease: New insights with 11c-pbr28 pet/mri. *ACS chemical neuroscience*, 9(11):2563–2571, 2018. 4.1.2.2

[186] A. Lombardi, A. Monaco, G. Donvito, N. Amoroso, R. Bellotti, and S. Tangaro. Brain age prediction with morphological features using deep neural networks: results from predictive analytic competition 2019. *Frontiers in Psychiatry*, 11:619629, 2021. 2.2.1

[187] C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer. The hallmarks of aging. *Cell*, 153(6):1194–1217, 2013. 1.1

[188] L. D. Lotter, A. Saberi, J. Y. Hansen, B. Misic, G. J. Barker, A. L. Bokde, S. Desrivieres, H. Flor, A. Grigis, H. Garavan, et al. Human cortex development is shaped by molecular and cellular brain systems. *bioRxiv*, pages 2023–05, 2023. 4.1.2.2

[189] N. Lubben, E. Ensink, G. A. Coetzee, and V. Labrie. The enigma and implications of brain hemispheric asymmetry in neurodegenerative diseases. *Brain Communications*, 3(3):fcab211, 2021. 1.3

[190] D. M. Lyall, T. Quinn, L. M. Lyall, J. Ward, J. J. Anderson, D. J. Smith, W. Stewart, R. J. Strawbridge, M. E. Bailey, and B. Cullen. Quantifying bias in psychological and physical health in the uk biobank imaging sub-sample. *Brain communications*, 4(3):fcac119, 2022. 4.1.1

[191] T. Madhyastha, S. Merillat, S. Hirsiger, L. Bezzola, F. Liem, T. Grabowski, and L. Jäncke. Longitudinal reliability of tract-based spatial statistics in diffusion tensor imaging. *Human brain mapping*, 35(9):4544–4555, 2014. 4.1.2.1

[192] C. Makowski, T. T. Brown, W. Zhao, D. J. Hagler, P. Parekh, H. Garavan, T. E. Nichols, T. L. Jernigan, and A. M. Dale. Reports of the death of brain-behavior associations have been greatly exaggerated. *bioRxiv*, pages 2023–06, 2023. 4.1

[193] M. Marek, M. Horyniecki, M. Frączek, and E. Kluczewska. Leukoaraiosis–new concepts and modern imaging. *Polish journal of radiology*, 83:76–81, 2018. 1.2.2

[194] S. Marek, B. Tervo-Clemmens, F. J. Calabro, D. F. Montez, B. P. Kay, A. S. Hatoum, M. R. Donohue, W. Foran, R. L. Miller, T. J. Hendrickson, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603(7902):654–660, 2022. 1.1, 1.2.1, 1.2.2, 4.1, 4.1.2

[195] A. F. Marquand, S. M. Kia, M. Zabihi, T. Wolfers, J. K. Buitelaar, and C. F. Beckmann. Conceptualizing mental disorders as deviations from normative functioning. *Molecular psychiatry*, 24(10):1415–1424, 2019. 1.1, 1.2.1

[196] A. F. Marquand, I. Rezek, J. Buitelaar, and C. F. Beckmann. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biological psychiatry*, 80(7):552–561, 2016. 1.1, 1.2.1

[197] E. Masicampo and R. F. Baumeister. Consider it done! plan making can eliminate the cognitive effects of unfulfilled goals. *Journal of personality and social psychology*, 101(4):667, 2011. 4.3

[198] P. M. Matthews and C. Sudlow. The uk biobank, 2015. 1.2.1, 1.3, 2.3.1

[199] I. I. Maximov, D. van Der Meer, A.-M. G. de Lange, T. Kaufmann, A. Shadrin, O. Frei, T. Wolfers, and L. T. Westlye. Fast quality control method for derived diffusion metrics (yttrium) in big data analysis: Uk biobank 18,608 example. *Human brain mapping*, 42(10):3141–3155, 2021. 2.1.2.4

[200] I. I. Maximov and L. T. Westlye. Comparison of different neurite density metrics with brain asymmetry evaluation. *Zeitschrift für Medizinische Physik*, 2023. 2.1.2.3, 4.1.2.1, 4.1.2.1

BIBLIOGRAPHY

[201] D. Mayer and D. Butler. Statistical validation. *Ecological modelling*, 68(1-2):21–32, 1993. 4.1.4

[202] N. B. McNamara, D. A. Munro, N. Bestard-Cuche, A. Uyeda, J. F. Bogie, A. Hoffmann, R. K. Holloway, I. Molina-Gonzalez, K. E. Askew, S. Mitchell, et al. Microglia regulate central nervous system myelin growth and integrity. *Nature*, 613(7942):120–129, 2023. 1.2.2

[203] D. W. McRobbie, E. A. Moore, M. J. Graves, and M. R. Prince. *MRI from Picture to Proton*. Cambridge university press, 2017. 2.1, 2.1.2, 1

[204] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan. A survey on bias and fairness in machine learning. *ACM computing surveys (CSUR)*, 54(6):1–35, 2021. 4.3

[205] T. R. Melzer, R. J. Keenan, G. J. Leeper, S. Kingston-Smith, S. A. Felton, S. K. Green, K. J. Henderson, N. J. Palmer, R. Shoorangiz, M. M. Almuqbel, et al. Test-retest reliability and sample size estimates after mri scanner relocation. *Neuroimage*, 211:116608, 2020. 4.1.2.1

[206] Y. Meng, K. Hynynen, and N. Lipsman. Applications of focused ultrasound in the brain: from thermoablation to drug delivery. *Nature Reviews Neurology*, 17(1):7–22, 2021. 4.2.2

[207] H. Merisaari, J. J. Tuulari, L. Karlsson, N. M. Scheinin, R. Parkkola, J. Saunavaara, T. Lähdesmäki, S. J. Lehtola, M. Keskinen, J. D. Lewis, et al. Test-retest reliability of diffusion tensor imaging metrics in neonates. *NeuroImage*, 197:598–607, 2019. 4.1.2.1

[208] C. Metzler-Baddeley, J. P. Mole, R. Sims, F. Fasano, J. Evans, D. K. Jones, J. P. Aggleton, and R. J. Baddeley. Fornix white matter glia damage causes hippocampal gray matter damage during age-dependent limbic decline. *Scientific reports*, 9(1):1060, 2019. 4.2.2

[209] V. Mhasawade, Y. Zhao, and R. Chunara. Machine learning and algorithmic fairness in public and population health. *Nature Machine Intelligence*, 3(8):659–666, 2021. 4.3

[210] M. Miłkowski, W. M. Hensel, and M. Hohol. Replicability or reproducibility? on the replication crisis in computational neuroscience and sharing only relevant detail. *Journal of computational neuroscience*, 45(3):163–172, 2018. 4.1

[211] P. R. Millar, B. A. Gordon, P. H. Luckett, T. L. Benzinger, C. Cruchaga, A. M. Fagan, J. J. Hassenstab, R. J. Perrin, S. E. Schindler, R. F. Allegri, et al. Multimodal brain age estimates relate to alzheimer disease biomarkers and cognition in early stages: a cross-sectional observational study. *Elife*, 12:e81869, 2023. 4.1.3

[212] D. B. Miller and J. P. O'Callaghan. Biomarkers of parkinson's disease: present and future. *Metabolism*, 64(3):S40–S46, 2015. 1

[213] K. L. Miller, F. Alfaro-Almagro, N. K. Bangerter, D. L. Thomas, E. Yacoub, J. Xu, A. J. Bartsch, S. Jbabdi, S. N. Sotiropoulos, J. L. Andersson, et al. Multimodal population brain imaging in the uk biobank prospective epidemiological study. *Nature neuroscience*, 19(11):1523–1536, 2016. 2.3.1

[214] A. Modabbernia, H. C. Whalley, D. C. Glahn, P. M. Thompson, R. S. Kahn, and S. Frangou. Systematic evaluation of machine learning algorithms for neuroanatomically-based age prediction in youth. Technical report, Wiley Online Library, 2022. 2.2.1

[215] K. Moodley and D. Chan. The hippocampus in neurodegenerative disease. *The hippocampus in clinical neuroscience*, 34:95–108, 2014. 4.2.1

[216] S. More, G. Antonopoulos, F. Hoffstaedter, J. Caspers, S. B. Eickhoff, K. R. Patil, A. D. N. Initiative, et al. Brain-age prediction: A systematic comparison of machine learning workflows. *NeuroImage*, 270:119947, 2023. 4.1.3

[217] S. Mori, S. Wakana, P. C. Van Zijl, and L. Nagae-Poetscher. *MRI atlas of human white matter*. Elsevier, 2005. 2.1.2.4

[218] E. Mormina, M. Petracca, G. Bommarito, N. Piaggio, S. Cocozza, and M. Inglese. Cerebellum and neurodegenerative diseases: beyond conventional magnetic resonance imaging. *World journal of radiology*, 9(10):371, 2017. 4.2.1

[219] P. Mörters and Y. Peres. *Brownian motion*, volume 30. Cambridge University Press, 2010. 2.1.2

[220] D. Moujalled, A. Strasser, and J. R. Liddell. Molecular mechanisms of cell death in neurological diseases. *Cell Death & Differentiation*, 28(7):2029–2044, 2021. 1.2.1

[221] M. H. Murdock and L.-H. Tsai. Insights into alzheimer's disease from single-cell genomic approaches. *Nature Neuroscience*, 26(2):181–195, 2023. 1.2

[222] P. I. Nakagawa, L. F. Pires, J. L. R. Moreira, and L. O. Bonino. Towards semantic description of explainable machine learning workflows. In *2021 IEEE 25th International Enterprise Distributed Object Computing Workshop (EDOCW)*, pages 236–244. IEEE, 2021. 4.3

[223] S. E. Nasrabady, B. Rizvi, J. E. Goldman, and A. M. Brickman. White matter changes in alzheimer's disease: a focus on myelin and oligodendrocytes. *Acta neuropathologica communications*, 6(1):1–10, 2018. 1.2.2

[224] N. Nazlee, G. D. Waiter, and A.-L. Sandu. Age-associated sex and asymmetry differentiation in hemispheric and lobar cortical ribbon complexity across adulthood: A uk biobank imaging study. *Human Brain Mapping*, 44(1):49–65, 2023. 1.3

[225] E. Neilson, X. Shen, S. R. Cox, T.-K. Clarke, E. M. Wigmore, J. Gibson, D. M. Howard, M. J. Adams, M. A. Harris, G. Davies, et al. Impact of polygenic risk for schizophrenia on cortical structure in uk biobank. *Biological psychiatry*, 86(7):536–544, 2019. 4.1.4

[226] D. Nielsen. Tree boosting with xgboost-why does xgboost win" every" machine learning competition? Master's thesis, NTNU, 2016. 2.2.1

[227] S. Noble, D. Scheinost, and R. T. Constable. Cluster failure or power failure? evaluating sensitivity in cluster-level inference. *Neuroimage*, 209:116468, 2020. 4.1.2

[228] L. B. Norbom, L. Ferschmann, N. Parker, I. Agartz, O. A. Andreassen, T. Paus, L. T. Westlye, and C. K. Tamnes. New insights into the dynamic development of the cerebral cortex in childhood and adolescence: Integrating macro-and microstructural mri findings. *Progress in Neurobiology*, 204:102109, 2021. 4.1.2.2

[229] D. S. Novikov, E. Fieremans, S. N. Jespersen, and V. G. Kiselev. Quantifying brain microstructure with diffusion mri: Theory and parameter estimation. *NMR in Biomedicine*, 32(4):e3998, 2019. 2.1.2, 2.1.2.1, 2.1.2.2, 4.1.2.1

[230] D. S. Novikov, V. G. Kiselev, and S. N. Jespersen. On modeling. *Magnetic resonance in medicine*, 79(6):3172–3193, 2018. 1.3, 2.1.2.3

[231] D. S. Novikov, J. Veraart, I. O. Jelescu, and E. Fieremans. Rotationally-invariant mapping of scalar and orientational metrics of neuronal microstructure with diffusion mri. *NeuroImage*, 174:518–538, 2018. 2.1.2.3

[232] S. Ocklenburg, P. Friedrich, O. Güntürkün, and E. Genç. Intrahemispheric white matter asymmetries: the missing link between brain structure and functional lateralization? *Reviews in the Neurosciences*, 27(5):465–480, 2016. 1.3

[233] N. Opel, J. Goltermann, M. Hermesdorf, K. Berger, B. T. Baune, and U. Dannlowski. Cross-disorder analysis of brain structural abnormalities in six major psychiatric disorders: a secondary analysis of mega-and meta-analytical findings from the enigma consortium. *Biological Psychiatry*, 88(9):678–686, 2020. 4.2.1

[234] R. Opfer, J. Krüger, L. Spies, A.-C. Ostwaldt, H. H. Kitzler, S. Schippling, and R. Buchert. Automatic segmentation of the thalamus using a massively trained 3d convolutional neural network: higher sensitivity for the detection of reduced thalamus volume by improved inter-scanner stability. *European Radiology*, 33(3):1852–1861, 2023. 1.3, 2.3.3, 4.3

[235] P. Parekh, G. V. Bhalerao, R. Rao, V. S. Sreeraj, B. Holla, J. Saini, G. Venkatasubramanian, J. P. John, S. Jain, and A. consortium. Protocol for magnetic resonance imaging acquisition, quality assurance, and quality check for the accelerator program for discovery in brain disorders using stem cells. *International Journal of Methods in Psychiatric Research*, 30(3):e1871, 2021. 4.1.2.1

[236] W. E. Parker, E. K. Weidman, J. L. Chazen, S. N. Niogi, R. Uribe-Cardenas, M. G. Kaplitt, and C. E. Hoffman. Magnetic resonance–guided focused ultrasound for ablation of mesial temporal epilepsy circuits: modeling and theoretical feasibility of a novel noninvasive approach. *Journal of neurosurgery*, 133(1):63–70, 2019. 4.2.2

[237] T. Paus. Mapping brain maturation and cognitive development during adolescence. *Trends in cognitive sciences*, 9(2):60–68, 2005. 1.1

[238] T. Paus. Imaging microstructure in the living human brain: A viewpoint. *NeuroImage*, 182:3–7, 2018. 4.1.2.2

[239] T. Paus, M. Keshavan, and J. N. Giedd. Why do many psychiatric disorders emerge during adolescence? *Nature reviews neuroscience*, 9(12):947–957, 2008. 1.1

[240] N. Peel, H. Bartlett, and R. McClure. Healthy ageing: how is it defined and measured? *Australasian Journal on Ageing*, 23(3):115–119, 2004. 1.1

[241] W. Penfield. *The second career: with other essays and addresses*. Boston: Little, Brown, 1963. I

[242] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols. *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011. 2.1.1

[243] L. Peyton, A. Oliveros, D.-S. Choi, and M.-H. Jang. Hippocampal regenerative medicine: neurogenic implications for addiction and mental disorders. *Experimental & molecular medicine*, 53(3):358–368, 2021. 4.2.1

[244] A. Pfefferbaum, T. Rohlfing, M. J. Rosenbloom, W. Chu, I. M. Colrain, and E. V. Sullivan. Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85 years) measured with atlas-based parcellation of mri. *Neuroimage*, 65:176–193, 2013. 1.2.1

[245] J. R. Phillips, D. H. Hewedi, A. M. Eissa, and A. A. Moustafa. The cerebellum and psychiatric disorders. *Frontiers in public health*, page 66, 2015. 4.2.1

[246] E. D. Plowey and J. L. Ziskin. Hippocampal phospho-tau/mapt neuropathology in the fornix in alzheimer disease: an immunohistochemical autopsy study. *Acta Neuropathologica Communications*, 4(1):1–13, 2016. 4.2.2

[247] R. A. Poldrack, T. O. Laumann, O. Koyejo, B. Gregory, A. Hover, M.-Y. Chen, K. J. Gorgolewski, J. Luci, S. J. Joo, R. L. Boyd, et al. Long-term neural and physiological phenotyping of a single human. *Nature communications*, 6(1):8885, 2015. 4.1, 4.1.1

[248] S. G. Popescu, B. Glocker, D. J. Sharp, and J. H. Cole. Local brain-age: a u-net model. *Frontiers in Aging Neuroscience*, 13:761954, 2021. 1.3, 2.2, 2.2, 2.2.2, 4.1.2

[249] L. Qin, Z. Guo, M. A. McClure, and Q. Mu. White matter changes from mild cognitive impairment to alzheimer's disease: A meta-analysis. *Acta Neurologica Belgica*, 121:1435–1447, 2021. 1.1

[250] V. Rajmohan and E. Mohandas. The limbic system. *Indian journal of psychiatry*, 49(2):132, 2007. 4.2.1

[251] R. Redlich, N. Opel, C. Bürger, K. Dohm, D. Grotegerd, K. Förster, D. Zaremba, S. Meinert, J. Repple, V. Enneking, et al. The limbic system in youth depression: brain structural and functional alterations in adolescent in-patients with severe depression. *Neuropsychopharmacology*, 43(3):546–554, 2018. 4.2.1

[252] B. C. Reeves, J. K. Karimy, A. J. Kundishora, H. Mestre, H. M. Cerci, C. Matouk, S. L. Alper, I. Lundgaard, M. Nedergaard, and K. T. Kahle. Glymphatic system impairment in alzheimer's disease and idiopathic normal pressure hydrocephalus. *Trends in molecular medicine*, 26(3):285–295, 2020. 4.2.1

[253] M. Reisert, E. Kellner, B. Dhital, J. Hennig, and V. G. Kiselev. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. *Neuroimage*, 147:964–975, 2017. 2.1.2.1, **??**, 2.1.2.3, **??**

[254] M. Reisert, V. G. Kiselev, and B. Dhital. A unique analytical solution of the white matter standard model using linear and planar encodings. *Magnetic resonance in medicine*, 81(6):3819–3825, 2019. 2.1.2.3

[255] L. M. Reus, X. Shen, J. Gibson, E. Wigmore, L. Ligthart, M. J. Adams, G. Davies, S. R. Cox, S. P. Hagenaars, M. E. Bastin, et al. Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in uk biobank. *Scientific reports*, 7(1):42140, 2017. 4.1.4

[256] G. Richard, K. Kolskår, K. M. Ulrichsen, T. Kaufmann, D. Alnæs, A.-M. Sanders, E. S. Dørum, J. M. Sánchez, A. Petersen, H. Ihle-Hansen, et al. Brain age prediction in stroke patients: Highly reliable but limited sensitivity to cognitive performance and response to cognitive training. *NeuroImage: Clinical*, 25:102159, 2020. 4.1.3

[257] A. S. Ríos, S. Oxenford, C. Neudorfer, K. Butenko, N. Li, N. Rajamani, A. Boutet, G. J. Elias, J. Germann, A. Loh, et al. Optimal deep brain stimulation sites and networks for stimulation of the fornix in alzheimer's disease. *Nature communications*, 13(1):7707, 2022. 4.2.2

[258] C. P. Robert, G. Casella, C. P. Robert, and G. Casella. Monte carlo integration. *Monte Carlo statistical methods*, pages 71–138, 1999. 2.1.2.3

[259] T. D. Robinson, Y. L. Sun, P. T. Chang, and J. J. Chen. In search of a unifying theory of white matter aging: tract morphometry-microstructure relationships. *bioRxiv*, pages 2023–03, 2023. 2, 4.1.2.1

[260] J. M. Roe, D. Vidal-Piñeiro, Ø. Sørensen, A. M. Brandmaier, S. Düzel, H. A. Gonzalez, R. A. Kievit, E. Knights, S. Kühn, U. Lindenberger, et al. Asymmetric thinning of the cerebral cortex across the adult lifespan is accelerated in alzheimer's disease. *Nature communications*, 12(1):721, 2021. 1, 1.1, 1.2.1, 1.3

[261] J. Rokicki, T. Wolfers, W. Nordhøy, N. Tesli, D. S. Quintana, D. Alnæs, G. Richard, A.-M. G. de Lange, M. J. Lund, L. Norbom, et al. Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Human brain mapping*, 42(6):1714–1726, 2021. 2.2, 4.1.3

[262] V. Romeo, T. H. Helbich, and K. Pinker. Breast pet/mri hybrid imaging and targeted tracers. *Journal of Magnetic Resonance Imaging*, 57(2):370–386, 2023. 2

[263] A. Rosberg, J. J. Tuulari, V. Kumpulainen, M. Lukkarinen, E. P. Pulli, E. Silver, A. Copeland, E. Saukko, J. Saunavaara, J. D. Lewis, et al. Test–retest reliability of diffusion tensor imaging scalars in 5-year-olds. *Human Brain Mapping*, 43(16):4984–4994, 2022. 4.1.2.1

[264] A. F. Rosen, D. R. Roalf, K. Ruparel, J. Blake, K. Seelaus, L. P. Villa, R. Ciric, P. A. Cook, C. Davatzikos, M. A. Elliott, et al. Quantitative assessment of structural image quality. *Neuroimage*, 169:407–418, 2018. 2.1.1

[265] S. Rutherford, C. Fraza, R. Dinga, S. M. Kia, T. Wolfers, M. Zabihi, P. Berthet, A. Worker, S. Verdi, D. Andrews, et al. Charting brain growth and aging at high spatial precision. *elife*, 11:e72904, 2022. 1.1, 1.2.1

[266] S. Rutherford, S. M. Kia, T. Wolfers, C. Fraza, M. Zabihi, R. Dinga, P. Berthet, A. Worker, S. Verdi, H. G. Ruhe, et al. The normative modeling framework for computational psychiatry. *Nature protocols*, 17(7):1711–1734, 2022. 1.2.1

[267] M. R. Sabuncu, E. Konukoglu, and A. D. N. Initiative. Clinical prediction from structural brain mri scans: a large-scale empirical study. *Neuroinformatics*, 13:31–46, 2015. 4.1.3

[268] R. Sala-Llonch, D. Bartrés-Faz, and C. Junqué. Reorganization of brain networks in aging: a review of functional connectivity studies. *Frontiers in psychology*, 6:663, 2015. 1.2

[269] A. Salih, I. B. Galazzo, A. Jaggi, Z. R. Estabragh, S. E. Petersen, K. Lekadir, P. Radeva, and G. Menegaz. Multi-modal brain age estimation: a comparative study confirms the importance of microstructure. In *Computational Diffusion MRI: International MICCAI Workshop, Lima, Peru, October 2020*, pages 239–250. Springer, 2021. 2.2

[270] K. Saltoun, R. Adolphs, L. K. Paul, V. Sharma, J. Diedrichsen, B. T. Yeo, and D. Bzdok. Dissociable brain structural asymmetry patterns reveal unique phenome-wide profiles. *Nature Human Behaviour*, pages 1–18, 2022. 1.3

[271] S. G. Samko. Fractional integrals and derivatives. *Theory and applications*, 1993. 2.2.2

[272] M. Samson and D. O. Claassen. Neurodegeneration and the cerebellum. *Neurodegenerative Diseases*, 17(4-5):155–165, 2017. 4.2.1

[273] M. A. Sayed, W. Eldahshan, M. Abdelbary, B. Pillai, W. Althomali, M. H. Johnson, A. S. Arbab, A. Ergul, and S. C. Fagan. Stroke promotes the development of brain atrophy and delayed cell death in hypertensive rats. *Scientific reports*, 10(1):20233, 2020. 1.2.1

[274] E. Schäffner, M. Bosch-Queralt, J. M. Edgar, M. Lehning, J. Strauß, N. Fleischer, T. Kungl, P. Wieghofer, S. A. Berghoff, T. Reinert, et al. Myelin insulation as a risk factor for axonal degeneration in autoimmune demyelinating disease. *Nature Neuroscience*, 26(7):1218–1228, 2023. 1.2.2

[275] D. Schijven, M. C. Postema, M. Fukunaga, J. Matsumoto, K. Miura, S. M. de Zwarte, N. E. Van Haren, W. Cahn, H. E. Hulshoff Pol, R. S. Kahn, et al. Large-scale analysis of structural brain asymmetries in schizophrenia via the enigma consortium. *Proceedings of the National Academy of Sciences*, 120(14):e2213880120, 2023. 1.3

[276] K. G. Schilling, V. Nath, C. Hansen, P. Parvathaneni, J. Blaber, Y. Gao, P. Neher, D. B. Aydogan, Y. Shi, M. Ocampo-Pineda, et al. Limits to anatomical accuracy of diffusion tractography using modern approaches. *Neuroimage*, 185:1–11, 2019. 4.1.2.1

[277] T. Schoeler, D. Speed, E. Porcu, N. Pirastu, J.-B. Pingault, and Z. Kutalik. Participation bias in the uk biobank distorts genetic associations and downstream analyses. *Nature Human Behaviour*, pages 1–12, 2023. 4.1.1

[278] T. J. Schoenfeld and H. A. Cameron. Adult neurogenesis and mental illness. *Neuropsychopharmacology*, 40(1):113–128, 2015. 4.2.1

[279] C. G. Schwarz, R. I. Reid, J. L. Gunter, M. L. Senjem, S. A. Przybelski, S. M. Zuk, J. L. Whitwell, P. Vemuri, K. A. Josephs, K. Kantarci, et al. Improved dti registration allows voxel-based analysis that outperforms tract-based spatial statistics. *Neuroimage*, 94:65–78, 2014. 4.1.2.1

[280] J. Seidlitz, A. Nadig, S. Liu, R. A. Bethlehem, P. E. Vértes, S. E. Morgan, F. Váša, R. Romero-Garcia, F. M. Lalonde, L. S. Clasen, et al. Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. *Nature communications*, 11(1):3358, 2020. 4.1.2.2

[281] C. E. Sexton, U. G. Kalu, N. Filippini, C. E. Mackay, and K. P. Ebmeier. A meta-analysis of diffusion tensor imaging in mild cognitive impairment and alzheimer's disease. *Neurobiology of aging*, 32(12):2322–e5, 2011. 1.1

[282] S. Shahab, B. H. Mulsant, M. L. Levesque, N. Calarco, A. Nazeri, A. L. Wheeler, G. Foussias, T. K. Rajji, and A. N. Voineskos. Brain structure, cognition, and brain age in schizophrenia, bipolar disorder, and healthy controls. *Neuropsychopharmacology*, 44(5):898–906, 2019. 1.2.1

[283] X. Shen, D. M. Howard, M. J. Adams, W. D. Hill, T.-K. Clarke, I. J. Deary, H. C. Whalley, and A. M. McIntosh. A phenome-wide association and mendelian randomisation study of polygenic risk for depression in uk biobank. *Nature communications*, 11(1):2301, 2020. 4.1.4

[284] A. Shukla-Dave, N. A. Obuchowski, T. L. Chenevert, S. Jambawalikar, L. H. Schwartz, D. Malyarenko, W. Huang, S. M. Noworolski, R. J. Young, M. S. Shiroishi, et al. Quantitative imaging biomarkers alliance (qiba) recommendations

for improved precision of dwi and dce-mri derived biomarkers in multicenter oncology trials. *Journal of Magnetic Resonance Imaging*, 49(7):e101–e121, 2019. 2

[285] R. Shwartz-Ziv and A. Armon. Tabular data: Deep learning is not all you need. *Information Fusion*, 81:84–90, 2022. 2.2, 2.2.1, 4.1.2

[286] J. P. Simmons, L. D. Nelson, and U. Simonsohn. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 22(11):1359–1366, 2011. 4.1.3

[287] P. Skalický, A. Mládek, A. Vlasák, P. De Lacy, V. Beneš, and O. Bradáč. Normal pressure hydrocephalus—an overview of pathophysiological mechanisms and diagnostic procedures. *Neurosurgical Review*, 43(6):1451–1464, 2020. 4.2.1

[288] S. M. Smith, M. Jenkinson, H. Johansen-Berg, D. Rueckert, T. E. Nichols, C. E. Mackay, K. E. Watkins, O. Ciccarelli, M. Z. Cader, P. M. Matthews, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*, 31(4):1487–1505, 2006. 2.1.2.4

[289] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al. Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, 23:S208–S219, 2004. 2.1.2.4, 2.4

[290] S. M. Smith and T. E. Nichols. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, 44(1):83–98, 2009. 2.4, 4.1.2

[291] J. Stumme, C. Jockwitz, F. Hoffstaedter, K. Amunts, and S. Caspers. Functional network reorganization in older adults: Graph-theoretical analyses of age, cognition and sex. *NeuroImage*, 214:116756, 2020. 1.2

[292] C. Sudlow, J. Gallacher, N. Allen, V. Beral, P. Burton, J. Danesh, P. Downey, P. Elliott, J. Green, M. Landray, et al. Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, 2015. 1.1, 1.2.1, 1.3, 2.3.1, 4.1, 4.3

[293] G. Sudre, J. Frederick, W. Sharp, A. Ishii-Takahashi, A. Mangalmurti, S. Choudhury, and P. Shaw. Mapping associations between polygenic risks for childhood neuropsychiatric disorders, symptoms of attention deficit hyperactivity disorder, cognition, and the brain. *Molecular psychiatry*, 25(10):2482–2492, 2020. 4.1.4

[294] A. Szafer, J. Zhong, and J. C. Gore. Theoretical model for water diffusion in tissues. *Magnetic resonance in medicine*, 33(5):697–712, 1995. 2.1.2.3

[295] D. Szucs and J. P. Ioannidis. Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *NeuroImage*, 221:117164, 2020. 4.1.1

[296] R. H. Takahashi, M. Yokotsuka, M. Tobiume, Y. Sato, H. Hasegawa, T. Nagao, and G. K. Gouras. Accumulation of cellular prion protein within β-amyloid oligomer plaques in aged human brains. *Brain Pathology*, 31(5):e12941, 2021. 1.2

[297] M. Tanveer, M. Ganaie, I. Beheshti, T. Goel, N. Ahmad, K.-T. Lai, K. Huang, Y.-D. Zhang, J. Del Ser, and C.-T. Lin. Deep learning for brain age estimation: A systematic review. *Information Fusion*, 2023. 2.2, 2.2, 2.2.2, 2.2.2, 4.1.2

[298] P. A. Taylor, R. C. Reynolds, V. Calhoun, J. Gonzalez-Castillo, D. A. Handwerker, P. A. Bandettini, A. F. Mejia, and G. Chen. Highlight results, don't hide them: Enhance interpretation, reduce biases and improve reproducibility. *NeuroImage*, 274:120138, 2023. 4.1.2

[299] P. M. Thompson, N. Jahanshad, C. R. Ching, L. E. Salminen, S. I. Thomopoulos, J. Bright, B. T. Baune, S. Bertolín, J. Bralten, W. B. Bruin, et al. Enigma and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Translational psychiatry*, 10(1):100, 2020. 1.2.1

[300] Y. E. Tian, V. Cropley, A. B. Maier, N. T. Lautenschlager, M. Breakspear, and A. Zalesky. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nature Medicine*, 29(5):1221–1231, 2023. 1.1

[301] S. Tønnesen, T. Kaufmann, N. T. Doan, D. Alnæs, A. Córdova-Palomera, D. v. d. Meer, J. Rokicki, T. Moberget, T. P. Gurholt, U. K. Haukvik, et al. White matter aberrations and age-related trajectories in patients with schizophrenia and bipolar disorder revealed by diffusion tensor imaging. *Scientific reports*, 8(1):14129, 2018. 2.3.4

[302] L. Torrey and J. Shavlik. Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pages 242–264. IGI global, 2010. 4.1.3

[303] O. Trofimova, A. Latypova, G. DiDomenicantonio, A. Lutti, A.-M. G. de Lange, M. Kliegel, S. Stringhini, P. Marques-Vidal, J. Vaucher, P. Vollenweider, et al. Topography of associations between cardiovascular risk factors and myelin loss in the ageing human brain. *Communications Biology*, 6(1):392, 2023. 1.2.2

[304] A. Tsapanou, C. Habeck, Y. Gazes, Q. Razlighi, J. Sakhardande, Y. Stern, and T. A. Salthouse. Brain biomarkers and cognition across adulthood. *Human brain mapping*, 40(13):3832–3842, 2019. 1

[305] S. Turrini, B. Wong, M. Eldaief, D. Press, D. A. Sinclair, G. Koch, A. Avenanti, and E. Santarnecchi. The multifactorial nature of healthy brain ageing: brain changes, functional decline and protective factors. *Ageing Research Reviews*, page 101939, 2023. 1.2, 1.2.2

[306] E. Uffelmann, Q. Q. Huang, N. S. Munung, J. De Vries, Y. Okada, A. R. Martin, H. C. Martin, T. Lappalainen, and D. Posthuma. Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1):59, 2021. 4.1.4

[307] S. van Alten, B. W. Domingue, T. Galama, and A. T. Marees. Reweighting the uk biobank to reflect its underlying sampling population substantially reduces pervasive selection bias due to volunteering. *medRxiv*, pages 2022–05, 2022. 4.1.1

[308] L. A. van de Mortel, R. M. Thomas, G. A. van Wingen, A. D. N. Initiative, et al. Grey matter loss at different stages of cognitive decline: A role for the thalamus in developing alzheimer's disease. *Journal of Alzheimer's Disease*, 83(2):705–720, 2021. 1.1

[309] L. Van de Pol, A. Hensel, F. Barkhof, H. Gertz, P. Scheltens, and W. Van Der Flier. Hippocampal atrophy in alzheimer disease: age matters. *Neurology*, 66(2):236–238, 2006. 4.2.1

[310] D. van der Meer, O. Frei, T. Kaufmann, A. A. Shadrin, A. Devor, O. B. Smeland, W. K. Thompson, C. C. Fan, D. Holland, L. T. Westlye, et al. Understanding the genetic determinants of the brain with mostest. *Nature communications*, 11(1):3512, 2020. 4.1.4

[311] D. Van der Meer, A. A. Shadrin, K. O'Connell, F. Bettella, S. Djurovic, T. Wolfers, D. Alnæs, I. Agartz, O. B. Smeland, I. Melle, et al. Boosting schizophrenia genetics by utilizing genetic overlap with brain morphology. *Biological psychiatry*, 92(4):291–298, 2022. 4.1.4

[312] C. Van der Merwe, R. Passchier, M. Mufford, R. Ramesar, S. Dalvie, and D. Stein. Polygenic risk for schizophrenia and associated brain structural changes: A systematic review. *Comprehensive Psychiatry*, 88:77–82, 2019. 4.1.4

[313] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013. 1.1

[314] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, et al. The human connectome project: a data acquisition perspective. *Neuroimage*, 62(4):2222–2231, 2012. 1.2.1, 4.1.2.2

[315] L. D. Vanes and R. J. Dolan. Transdiagnostic neuroimaging markers of psychiatric risk: A narrative review. *NeuroImage: Clinical*, 30:102634, 2021. 4.2.1

[316] D. Vidal-Pineiro, Y. Wang, S. K. Krogsrud, I. K. Amlien, W. F. Baaré, D. Bartres-Faz, L. Bertram, A. M. Brandmaier, C. A. Drevon, S. Düzel, et al. Individual variations in 'brain age'relate to early-life factors more than to longitudinal brain change. *elife*, 10:e69995, 2021. 4.1.3

[317] R. P. Viviano and J. S. Damoiseaux. Functional neuroimaging in subjective cognitive decline: current status and a research path forward. *Alzheimer's Research & Therapy*, 12(1):1–18, 2020. 1.1

[318] W. J. von Eschenbach. Transparency and the black box problem: Why we do not trust ai. *Philosophy & Technology*, 34(4):1607–1622, 2021. 4.3

BIBLIOGRAPHY

[319] L. L. Wald, P. C. McDaniel, T. Witzel, J. P. Stockmann, and C. Z. Cooley. Low-cost and portable mri. *Journal of Magnetic Resonance Imaging*, 52(3):686–696, 2020. 2

[320] K. Wang, C. Xu, G. Li, Y. Zhang, Y. Zheng, and C. Sun. Combining convolutional neural networks and self-attention for fundus diseases identification. *Scientific Reports*, 13(1):76, 2023. 4.1.3

[321] M.-Y. Wang, M. Korbmacher, R. Eikeland, and K. Specht. Deep brain imaging of three participants across 1 year: The bergen breakfast scanning club project. *Frontiers in Human Neuroscience*, 16:1021503, 2022. 1.2.3, 1.3, 2.2, 2.3.2, 4.1

[322] M.-Y. Wang, M. Korbmacher, R. Eikeland, and K. Specht. The bergen breakfast scanning club dataset: a deep brain imaging dataset. *bioRxiv*, pages 2023–05, 2023. 1.3, 2.2, 2.3.2, 4.1.1

[323] J. M. Wardlaw, C. Smith, and M. Dichgans. Small vessel disease: mechanisms and clinical implications. *The Lancet Neurology*, 18(7):684–696, 2019. 1.2.2

[324] H.-Y. Wey, T. M. Gilbert, N. R. Zürcher, A. She, A. Bhanot, B. D. Taillon, F. A. Schroeder, C. Wang, S. J. Haggarty, and J. M. Hooker. Insights into neuroepigenetics through human histone deacetylase pet imaging. *Science translational medicine*, 8(351):351ra106–351ra106, 2016. 4.1.2.2

[325] B. L. Wilkoff, D. Bello, M. Taborsky, J. Vymazal, E. Kanal, H. Heuer, K. Hecking, W. B. Johnson, W. Young, B. Ramza, et al. Magnetic resonance imaging in patients with a pacemaker system designed for the magnetic resonance environment. *Heart rhythm*, 8(1):65–73, 2011. 2

[326] T. Wise, J. Radua, E. Via, N. Cardoner, O. Abe, T. Adams, F. Amico, Y. Cheng, J. Cole, C. de Azevedo Marques Périco, et al. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis. *Molecular psychiatry*, 22(10):1455–1463, 2017. 1.2.1

[327] C.-W. Woo, L. J. Chang, M. A. Lindquist, and T. D. Wager. Building better biomarkers: brain models in translational neuroimaging. *Nature neuroscience*, 20(3):365–377, 2017. 1, 1.1, 4.2

[328] J. Wrigglesworth, P. Ward, I. H. Harding, D. Nilaweera, Z. Wu, R. L. Woods, and J. Ryan. Factors associated with brain ageing-a systematic review. *BMC neurology*, 21(1):312, 2021. 1.2.1

[329] T. Wyss-Coray. Ageing, neurodegeneration and brain rejuvenation. *Nature*, 539(7628):180–186, 2016. 1.1

[330] Y. Xie, L. Xie, F. Kang, J. Jiang, T. Yao, G. Mao, R. Fang, J. Fan, and D. Wu. Association between white matter alterations and domain-specific cognitive impairment in cerebral small vessel disease: a meta-analysis of diffusion tensor imaging. *Frontiers in Aging Neuroscience*, 14:1019088, 2022. 1.1

[331] M. Xiong, L. Lin, Y. Jin, W. Kang, S. Wu, and S. Sun. Comparison of machine learning models for brain age prediction using six imaging modalities on middle-aged and older adults. *Sensors*, 23(7):3622, 2023. 4.1.3

[332] S. Yamada, M. Ishikawa, and K. Nozaki. Exploring mechanisms of ventricular enlargement in idiopathic normal pressure hydrocephalus: a role of cerebrospinal fluid dynamics and motile cilia. *Fluids and Barriers of the CNS*, 18(1):1–11, 2021. 4.2.1

[333] Y. Yang, K. Schilling, N. Shashikumar, V. Jasodanand, E. E. Moore, K. R. Pechman, M. Bilgel, L. L. Beason-Held, Y. An, A. Shafer, et al. White matter microstructural metrics are sensitively associated with clinical staging in alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 15(2):e12425, 2023. 4.2.2

[334] Z. A. Yaple, W. D. Stevens, and M. Arsalidou. Meta-analyses of the n-back working memory task: fmri evidence of age-related changes in prefrontal cortex involvement across the adult lifespan. *NeuroImage*, 196:16–31, 2019. 1.1

[335] T. Yarkoni. The generalizability crisis. *Behavioral and Brain Sciences*, 45:e1, 2022. 4.1.2

[336] H. Yu, Y.-j. Meng, X.-j. Li, C. Zhang, S. Liang, M.-l. Li, Z. Li, W. Guo, Q. Wang, W. Deng, et al. Common and distinct patterns of grey matter alterations in borderline personality disorder and bipolar disorder: voxel-based meta-analysis. *The British Journal of Psychiatry*, 215(1):395–403, 2019. 1.2.1

[337] A. Zalesky. Moderating registration misalignment in voxelwise comparisons of dti data: a performance evaluation of skeleton projection. *Magnetic resonance imaging*, 29(1):111–125, 2011. 4.1.2.1

[338] A. Zavaliangos-Petropulu, B. Lo, M. R. Donnelly, N. Schweighofer, K. Lohse, N. Jahanshad, G. Barisano, N. Banaj, M. R. Borich, L. A. Boyd, et al. Chronic stroke sensorimotor impairment is related to smaller hippocampal volumes: An enigma analysis. *Journal of the American Heart Association*, 11(10):e025109, 2022. 1

[339] B. Zeigarnik. On finished and unfinished tasks. 1938. 4.3

[340] M. Zelikowsky, S. Hersman, M. K. Chawla, C. A. Barnes, and M. S. Fanselow. Neuronal ensembles in amygdala, hippocampus, and prefrontal cortex track differential components of contextual fear. *Journal of neuroscience*, 34(25):8462–8466, 2014. 4.2.1

[341] H. Zhang, T. Schneider, C. A. Wheeler-Kingshott, and D. C. Alexander. Noddi: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage*, 61(4):1000–1016, 2012. 2.1.2.3

[342] H. Zheng, Z. Yang, W. Liu, J. Liang, and Y. Li. Improving deep neural networks using softplus units. In *2015 International joint conference on neural networks (IJCNN)*, pages 1–4. IEEE, 2015. 2.2.2

[343] X. Zhou. Understanding the convolutional neural networks with gradient descent and backpropagation. In *Journal of Physics: Conference Series*, volume 1004, page 012028. IOP Publishing, 2018. 2.2.2

[344] X. Zhou, R. Wu, Y. Zeng, Z. Qi, S. Ferraro, L. Xu, X. Zheng, J. Li, M. Fu, S. Yao, et al. Choice of voxel-based morphometry processing pipeline drives variability in the location of neuroanatomical brain markers. *Communications Biology*, 5(1):913, 2022. 2.1.1, 4.1.2.2

Part II

ARTICLES

BRAIN-WIDE ASSOCIATIONS BETWEEN WHITE MATTER AND AGE HIGHLIGHT THE ROLE OF FORNIX MICROSTRUCTURE IN BRAIN AGEING

M. Korbmacher, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov

RESEARCH ARTICLE

WILEY

Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing

Max Korbmacher[1,2,3] | Ann Marie de Lange[2,4,5] | Dennis van der Meer[2,6] | Dani Beck[2,7,8] | Eli Eikefjord[1,3] | Arvid Lundervold[1,3,9,10] | Ole A. Andreassen[2,11] | Lars T. Westlye[2,8,11] | Ivan I. Maximov[1,2]

[1]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway

[2]NORMENT Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway

[3]Mohn Medical Imaging and Visualisation Center (MMIV), Bergen, Norway

[4]Department of Psychiatry, University of Oxford, Oxford, UK

[5]LREN, Centre for Research in Neurosciences–Department of Clinical Neurosciences, CHUV and University of Lausanne, Lausanne, Switzerland

[6]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

[7]Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, Norway

[8]Department of Psychology, University of Oslo, Oslo, Norway

[9]Department of Radiology, Haukeland University Hospital, Bergen, Norway

[10]Department of Biomedicine, University of Bergen, Bergen, Norway

[11]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

Correspondence
Max Korbmacher and Ivan I. Maximov, Department of Health and Functioning, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway.
Email: max.korbmacher@hvl.no and ivan.maximov@hvl.no

Abstract

Unveiling the details of white matter (WM) maturation throughout ageing is a fundamental question for understanding the ageing brain. In an extensive comparison of brain age predictions and age-associations of WM features from different diffusion approaches, we analyzed UK Biobank diffusion magnetic resonance imaging (dMRI) data across midlife and older age ($N = 35{,}749$, 44.6–82.8 years of age). Conventional and advanced dMRI approaches were consistent in predicting brain age. WM-age associations indicate a steady microstructure degeneration with increasing age from midlife to older ages. Brain age was estimated best when combining diffusion approaches, showing different aspects of WM contributing to brain age. Fornix was found as the central region for brain age predictions across diffusion approaches in complement to forceps minor as another important region. These regions exhibited a general pattern of positive associations with age for intra axonal water fractions, axial, radial diffusivities, and negative relationships with age for mean diffusivities, fractional anisotropy, kurtosis. We encourage the application of multiple dMRI approaches for detailed insights into WM, and the further investigation of fornix and forceps as potential biomarkers of brain age and ageing.

1 | INTRODUCTION

Along the past decades, neuroscientific research, and particularly magnetic resonance imaging (MRI) have increased our understanding of the biological mechanisms associated with brain tissue maturation and ageing effects (Grady, 2012; Symms et al., 2004; Wrigglesworth et al., 2021). A fundamental basis for that are large-scale MRI databases, such as UK Biobank (UKB; Sudlow et al., 2015) or the Human Connectome Project (Van Essen et al., 2012), allowing one to provide larger generalizability for revealed effects (Marek et al., 2022). Simultaneously, large-scale data provide sufficient power for the application of advanced multivariate statistical models, and machine learning (ML) techniques. Brain age prediction is an example of such technique, translating large amounts of complex multidimensional data into practically interpretable outputs. Brain age prediction involves training a ML model to determine trajectories of brain ageing from a series of brain MRI features. Once the model is trained, it can predict the age of brains not included in the training data. The disparity between chronological age and predicted age, the so-called brain age gap (BAG), can be used as an indicator of various disorders and potentially general health status (Beck et al., 2022; Cole et al., 2017; Franke & Ten Gaser, 2019; Kaufmann et al., 2019; Leonardsen et al., 2021). For example, BAG has been associated with stroke history, diabetes, smoking, alcohol intake, several cognitive measures (Cole, 2020; Leonardsen et al., 2021), cardiovascular risk factors (Beck et al., 2022), stroke risk (de Lange et al., 2020), and loneliness (de Lange et al., 2021), mortality risk, different brain and psychiatric disorders, particularly Alzheimer's disease and schizophrenia (Cole & Franke, 2017; Franke & Ten Gaser, 2019; Kaufmann et al., 2019; Rokicki et al., 2021). Yet, the effect of brain age on brain maturation remains unclear (Vidal-Pineiro et al., 2021), indicating the need for further investigation.

BAG and age trajectories offer paths toward a better understanding of the ageing brain. There are various detectable age-related brain changes, such as GM and white matter (WM) atrophy (Lawrence et al., 2021), WM de-differentiation (Cox et al., 2016a), and functional connectivity changes (Wrigglesworth et al., 2021) which have hence informed the choice of brain-age modeling-parameters (Beck et al., 2021; Beck et al., 2022; Cole, 2020; de Lange et al., 2020; Le Chen et al., 2020; Richard et al., 2018; Salih et al., 2021). In that context, many ML approaches have been used to make robust and clinically relevant brain age predictions from different MRI modalities (Baecker et al., 2021; Dosenbach et al., 2010; Franke et al., 2010; Kaufmann et al., 2019); yet, particularly the eXtreme Gradient Boosting (Chen & Guestrin, 2016) regressor model, using a decision tree approach, is increasingly used for brain age predictions from large-scale data due to its precision and speed (Beck et al., 2021; de Lange et al., 2019; Kaufmann et al., 2019). Especially diffusion magnetic resonance imaging (dMRI) and structural MRI have been shown useful for brain age predictions (Beck et al., 2021; Beck et al., 2022;

Cole, 2020; de Lange et al., 2020; Le Chen et al., 2020; Richard et al., 2018; Salih et al., 2021). However, further systematic, sufficiently powered assessments of dMRI-derived brain age and how diffusion metrics map onto age are needed. To this end, there are only a few publications about the influence of diffusion derived metrics on brain age predictions. Moreover, studies on the relationships between age and diffusion metrics usually focus on diffusion tensor imaging (DTI; Basser et al., 1994). In turn, advanced dMRI approaches (Fieremans et al., 2011; Jensen et al., 2005; Kaden et al., 2016a; Kaden et al., 2016b; Novikov et al., 2019; Reisert et al., 2017; Westlye et al., 2010) which offer additional details on WM microstructure and, hence, brain maturation processes require further research. In order to address this shortcoming, this study focusses in dMRI-derived measures from a large midlife-to-older adult sample and the measures' associations with age.

DMRI-derived measures consist of unique parameters allowing both to reveal WM changes at micrometer scale and to provide the basis for a prediction of macroscopic outcomes, such as age. Conventionally, WM brain architecture is described using DTI (Basser et al., 1994). However, recent advances offer more biophysically meaningful approaches (Novikov et al., 2019), and sensible foundation for cross-validation and better comparability (Beck et al., 2021). DTI-derived measures, namely fractional anisotropy (FA), and axial (AD), mean (MD), and radial (RD) diffusivity have all been shown to be highly age sensitive (Beck et al., 2021; Cox et al., 2016a; Westlye et al., 2010). Nevertheless, the DTI approach is limited by the Gaussian diffusion assumption and is unable to take into account entangled WM microstructure features (Beck et al., 2021). In the present work, we consider (1) the Bayesian rotationally invariant approach (BRIA; Reisert et al., 2017), (2) diffusion kurtosis imaging (DKI; Jensen et al., 2005), (3) kurtosis derived supplement, known as *white matter tract integrity* (WMTI; Fieremans et al., 2011) (4) spherical mean technique (SMT; Kaden et al., 2016a), and (5) *multi-compartment spherical mean technique* (mcSMT; Kaden et al., 2016b) in addition to DTI. Only a few studies have compared dMRI models directly as original brain age predictors (Beck et al., 2021; Maximov et al., 2021; Raghavan et al., 2021). Yet, brain age and age curve assessments of DTI, BRIA, DKI, WMTI, SMT, mcSMT (Table S10) in a representative sample present a great interest, as well as most influential WM regions for brain ageing. Our assessments focus on the process of ageing (from midlife to late adulthood), starting by associating BAG across diffusion approaches and compare–predicted versus chronological-age correlations in order to assess predictors' consistency. As fornix was identified as most contributing feature in these predictions, and forceps minor as another influential region, post-hoc analyses focused on both fornix, forceps minor, and whole-brain relationships with age. Fornix was the strongest correlate of age, and fornix and forceps minor features were highly correlated across approaches. Finally, we created fornix, forceps minor, and whole-brain-age curves expecting

curvilinear relationships reflecting brain-tissue-composition at different ageing stages (Beck et al., 2021; Davis et al., 2009; Westlye et al., 2010).

2 | METHODS

2.1 | Sample characteristics

The original UKB (Sudlow et al., 2015) diffusion MRI data consisted of $N = 42,208$ participants. After exclusions, based on later withdrawn consent and an ICD-10 diagnosis from categories F, G, I, and stroke (excluded: $N = 3521$), and data sets not meeting quality control standards ($N = 2938$) using the YTTRIUM method (Maximov et al., 2021), we obtained a final sample consisting of 35,749 healthy adults (age range 44.57–82.75, Mage = 64.46, SDage = 7.62, Mdage = 64.97; 52.96% females, 47.04% males). In brief, YTTRIUM converts diffusion scalar metric into 2D format using a structural similarity extension (Wang et al., 2004) of each scalar map to their mean image in order to create a 2D distribution of image and diffusion parameters. The quality check is based on a two-step clustering algorithm applied to identify subjects located out of the main distribution. We define healthy here as the absence of mental and behavioral disorder (ICD-10 category F), disease of the nervous system (ICD-10 category G), and disease of the circulatory system (ICD-10 category I). Included participants showed generally higher cognitive test performance and took less medication than excluded subjects (Table 1). Participants were recruited and scanned at four different sites: 57.62% in Cheadle, 26.30% in Newcastle, 15.96% in Reading, and 0.12% in Bristol (Figure 1). Imbalances in age distributions in the Bristol sample can be attributed to the small number of participants sampled ($N = 43$).

2.2 | MRI acquisition, diffusion pipeline, and tract-based spatial statistic analysis

UKB MRI data acquisition procedures are described elsewhere (Miller et al., 2016; Sudlow et al., 2015). The brain scan protocol (https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=2367) was applied at each scanner site (see also documentation: https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=1977). Shortly, the diffusion protocol consists of two b-values (1000 and 2000 s/mm^2) with 50 noncoplanar diffusion weighting gradients per each shell. For a susceptibility artefact correction, nondiffusion weighted images with an opposite gradient encoding direction were acquitted as well.

Diffusion data preprocessing was conducted as described in Maximov et al. (2019), using an optimized pipeline which includes corrections for noise (Veraart et al., 2016), Gibbs ringing (Kellner et al., 2016), susceptibility-induced and motion distortions, and eddy current artefacts (Andersson & Sotiropoulos, 2016). Isotropic Gaussian smoothing was carried out with the FSL (Jenkinson et al., 2012) function *fslmaths* with a Gaussian kernel of 1 mm^3. After that DTI, DKI, and WMTI metrics were estimated using Matlab 2017b (Mathworks, 2017). Employing the multishell data, DKI and WMTI metrics were estimated using Matlab code (https://github.com/NYU-DiffusionMRI/DESIGNER; Fieremans et al., 2011). SMT, and mcSMT metrics were estimated using original code (https://github.com/ekaden/smt; Kaden et al., 2016a), as well as Bayesian estimates/BRIA were estimated by the original Matlab code (https://bitbucket.org/reisert/baydiff/src/master/; Reisert et al., 2017).

In total, we obtained 28 metrics from 6 diffusion approaches (DTI, DKI, WMTI, SMT, mcSMT, BRIA; Beck et al., 2021; Kaden et al., 2016b; Maximov et al., 2019; Benitez et al., 2018; Hope et al., 2019; Pines et al., 2020). In order to normalize all metrics, we used TBSS (Smith et al., 2006), as part of FSL (Smith et al., 2004). In brief, initially all BET-extracted (Smith, 2002) FA images were aligned to MNI space using nonlinear transformation (FNIRT; Jenkinson et al., 2012). Afterward, the mean FA image and related mean FA skeleton were derived. Each diffusion scalar map was projected onto the mean FA skeleton using the TBSS procedure. In order to provide a quantitative description of diffusion metrics we evaluated averaged values over the skeleton and two white matter atlases, namely the JHU atlas (Mori & Wakana, 2005) and the JHU tractographic atlas (Hua et al., 2008). Finally, we obtained 20 WM tracts and 48 regions of interest (ROIs) based on a probabilistic white matter

TABLE 1 Included and excluded sample characteristics.

Variable	Excluded ($N = 6459$)	Included ($N = 35,749$)	P-value	Cohens d
Number of medications	2.812 (2.782)	1.784 (2.034)	<.001	0.474
Self-rated health	2.204 (0.764)	1.965 (0.644)	<.001	0.360
Number of correctly solved matrix puzzles	7.671 (2.191)	8.012 (2.126)	<.001	−0.159
Number of correctly solved tower puzzles	9.650 (3.318)	9.917 (3.224)	<.001	−0.083
Number of correct symbol digit matches	17.808 (5.414)	18.998 (5.246)	<.001	−0.226
Number of incorrectly matched pairs	2.239 (1.282)	2.215 (1.274)	0.250	0.019
Matrix puzzle response time in seconds	81.116 (16.605)	83.011 (15.873)	<.001	−0.119
Maximum number of remembered digits	6.497 (1.642)	6.678 (1.538)	<.001	−0.117
Fluid intelligence	6.429 (2.096)	6.634 (2.054)	<.001	−0.099
Prospective memory score	1.069 (0.433)	1.068 (0.397)	0.783	0.004

Note: Mean (SD) for each sample's variables. *p*-values are indicated for Welch two sample t-tests.
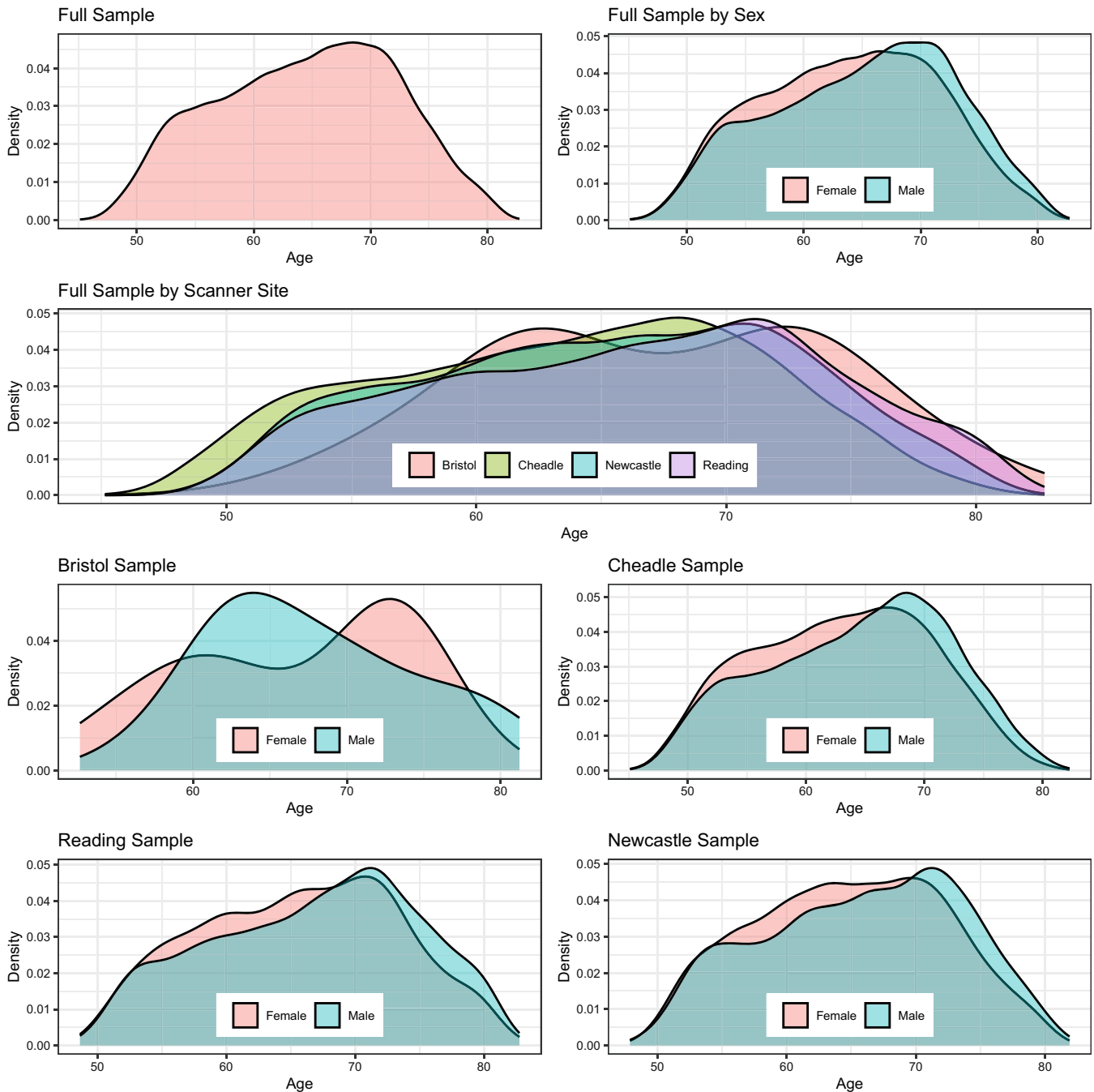
FIGURE 1 Density plots for the sample's age by sex and scanner site. The *y*-axis indicates the probability of age scaled to 1.

atlas (JHU; Hua et al., 2008) for each of the 28 metrics, including the mean skeleton values. Altogether, 1932 features per individual were derived (28 metrics * [48 ROIs +1 skeleton mean + 20 tracts]; see number of dMRI features in Table 2)). We included both whole-brain average metrics in addition to tracts and regional averages, as these provide spatially differential information (Figure S16), also expressed the metrics' relationships with age (Barrick et al., 2010; Beck et al., 2021; Eikenes et al., 2023; Kochunov et al., 2007; Westlye et al., 2010).

2.3 | Brain age predictions

First, brain age predictions were performed using XGBoost (Chen & Guestrin, 2016) in Python (v3.7.1). To evaluate how much data was needed for hyper-parameter tuning while accurately predicting brain age from all 1932 brain features, we divided the full dataset (*N* = 35,749) into two equal parts: one validation set and one hyper-parameter tuning set for independent parameter-tuning. From the hyper-parameter tuning set, data was randomly sampled into

TABLE 2 Performance of brain age prediction models.

Approach[a]	Number of MRI features	R^2 (SD)	RMSE (SD)	MAE (SD)	Prediction-age correlation* [95% CI]	Corrected prediction-age correlation* [95% CI]
BRIA	690	0.550 (0.012)	5.007 (0.057)	4.002 (0.042)	0.742 [0.737, 0.747]	0.892[b] [0.889, 0.894]
DKI	207	0.576 (0.015)	4.958 (0.077)	3.975 (0.068)	0.754 [0.755, 0.764]	0.903 [0.901, 0.905]
DTI	276	0.571 (0.014)	4.983 (0.072)	3.984 (0.062)	0.756 [0.751, 0.761]	0.900 [0.897, 0.902]
SMT	276	0.531 (0.010)	5.214 (0.053)	4.183 (0.036)	0.729 [0.724, 0.734]	0.899 [0.897, 0.901]
mcSMT	276	0.519 (0.011)	5.175 (0.045)	4.153 (0.036)	0.721 [0.716, 0.726]	0.892[b] [0.889, 0.894]
WMTI	207	0.585 (0.012)	4.903 (0.065)	3.928 (0.050)	0.765 [0.761, 0.770]	0.902 [0.900, 0.904]
Mean multimodal	28	0.393 (0.012)	5.932 (0.051)	4.812 (0.046)	0.627 [0.621, 0.634]	0.905 [0.903, 0.907]
Full multimodal	1932	0.645 (0.011)	4.534 (0.041)	3.624 (0.037)	0.804 [0.800, 0.808]	0.907 [0.905, 0.909]

Note: R^2, RMSE, MAE are displayed in the format Mean (Standard Deviation), Pearson's correlations are displayed in the format Correlation Score 95% Confidence Interval (Lower Bound, Upper Bound). Mean multimodal refers to diffusion metrics averaged over the skeleton for all six diffusion approaches. Full multimodal refers to all diffusion data from the six diffusion approaches, that is, mean multimodal data in addition to metrics averaged over the JHU atlas regions.

Abbreviations: BRIA, Bayesian rotationally invariant approach; MAE = mean absolute error; R^2 = variance explained; RMSE = root mean squared error.

[a]For an overview of the metrics contained in each of the diffusion approaches see Table S10.

[b]Details on the smallest correlation: BRIA Corrected Prediction-Age Correlation $r = 0.89173$, mcSMT Corrected Prediction-Age Correlation $r = 0.89176$.

*All correlation were significant at $p < .001$.

subsamples consisting of 358, 715, 1073, 1430, 1788, 2145, 2503, 2860, 3218, 3575, 7150, 10,725, 14,300, or 17,875 participants, corresponding to 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 20%, 30%, 40%, and 50% of the total subjects, respectively (Figure 2). Hyperparameters were tuned on these sub-samples and then tested on the remaining half, that is, the validation sample, using 10-fold cross validation showing model performance to not further improve past the 10% (tuning) data mark, informing our tuning-validation-split (Figure 2, Table S1, trained models in S2).

Second, in order to compare the different diffusion approaches, based on the previous steps, the training-test split was fixed at previously used 10% training data ($N = 3575$) and 90% test data ($N = 32,174$) which indicated a best fit at a learning rate = 0.05, max layers/depth = 3, and number of trees = 750. These tuned parameters were used for 10-fold cross-validations brain age predictions on the test data of all six individual models, one multimodal model combining all metrics from all diffusion models, and one multimodal model using only mean values from all diffusion models (Table 2).

Third, uncorrected BAG was calculated as the difference between chronological age Ω and predicted age P:

$$BAG_u = P - \Omega \tag{1}$$

We calculated BAG as it is the commonly used metric indicative of general health when using brain age predictions (Beck et al., 2022; Cole, 2020; Cole et al., 2017; Cole & Franke, 2017; de Lange et al., 2020; de Lange et al., 2021; Franke & Ten Gaser, 2019; Kaufmann et al., 2019; Leonardsen et al., 2021; Rokicki et al., 2021; Vidal-Pineiro et al., 2021). BAG is, however, sensitive to the age distribution of the sample (de Lange et al., 2019; de Lange & Cole, 2020). Hence as a supplement, age-bias-corrected predicted age was calculated from the intercept and slope of age predictions as previously described (de Lange et al., 2019; de Lange & Cole, 2020):

$$P = \alpha \times \Omega + \beta \tag{2}$$
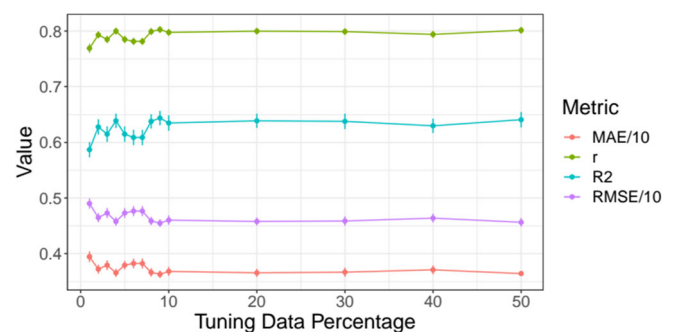
FIGURE 2 Model performance for different train-test splits. Model metrics R^2, root mean squared error (RMSE), mean absolute error (MAE) and their standard deviations, as well as the Pearson's correlations between predicted and chronological age and its 95% confidence interval are displayed for different training data percentages of the total data (x-axis). For visualization purposes, RMSE and MAE were divided by 10. For exact values see Table S1.

$$BAGc = (P + [\Omega - (\alpha \times \Omega + \beta)]) - \Omega \qquad (3)$$

P represents predicted age modelled from chronological age Ω, with intercept β and slope α. This age-bias correction allowed for a bias-corrected BAG estimate (BAGc). See Figure 3 for both uncorrected and age-bias-corrected brain ages over age.

2.4 | Statistical analyses

All statistical analyses were carried out using R (v3.6.0; www.r-project.org/). *p*-values were adjusted for multiple comparison using Holm correction (Holm, 1979). Model performance for brain ages estimations across different diffusion approaches are presented in addition to top five features for each brain age model ranked based on their model contributions (variance explained, as determined by permutation feature importance testing). Then, the correlation structure of age, brain age, BAG, and brain features (identified as main contributors in the model and whole-brain-average scores) were examined across diffusion approaches. In detail: first, brain ages were correlated across diffusion-approach-specific brain ages. Then, the correlations between true and estimated age across diffusion approaches were compared. Second, BAGs were correlated across diffusion approaches. Third, we present the correlation structure of fornix and age, and present brain-age crude and adjusted age-relationships for all included metrics (M).

$$M = \beta 0 + \beta 1 \text{Age} + \beta 2 \text{Age}^2 + \beta 3 \times \text{Site}^* \text{Sex} + \beta 4 \text{Sex}^* \text{Age} + \beta 5 \text{Sex} + \beta 6 \text{Site} \qquad (4)$$

Fourth, we plot absolute/crude whole-brain and fornix diffusion metrics by age, and contrast these with diffusion metrics (M) adjusted for age, sex, and site. To test the age-sensitivity of the metrics, we removed age from the model and compared the models using Likelihood Ratio tests.

$$M = \beta 0 + \beta 1 \text{Site}^* \text{Sex} + \beta 2 \text{Sex} + \beta 3 \text{Site} \qquad (5)$$

We also assess to which extent the regression lines can be called linear by comparing model fit of generalized additive models with simple linear regression models for fornix and whole brain features. Finally, we associate the first two principal components of all WM features with the different brain ages to assess the relationship between BAG and WM. For an overview of the analyses see Figure 4.

3 | RESULTS

3.1 | Brain age predictions

Table 2 presents a comparison between different diffusion approaches in predicting brain age for each diffusion approach. The strongest correlation between uncorrected age predictions and
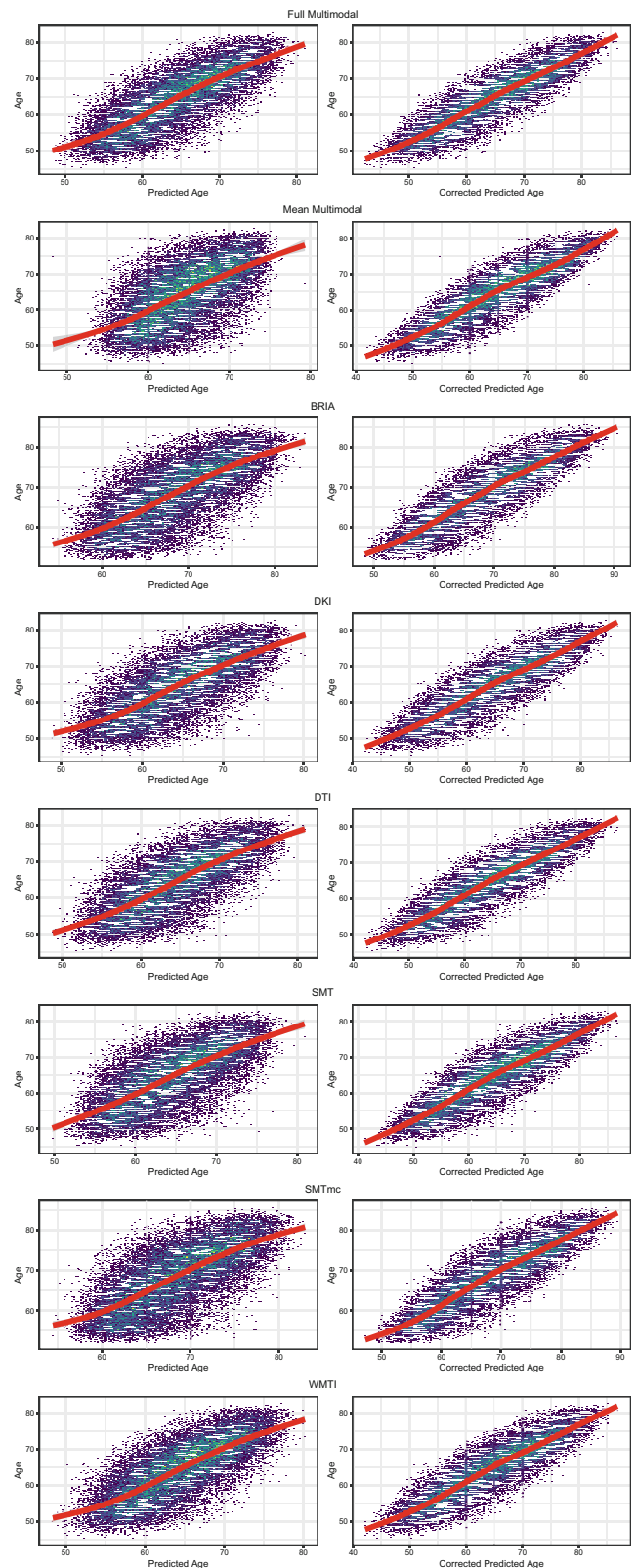
FIGURE 3 Corrected and uncorrected brain age by age for each of the utilized brain age models.

chronological age was observed for WMTI Pearson's $r = 0.765$, 95% CI [0.761, 0.770], $p < .001$, and the smallest for mcSMT Pearson's $r = 0.721$, 95% CI [0.716, 0.726], $p < .001$.
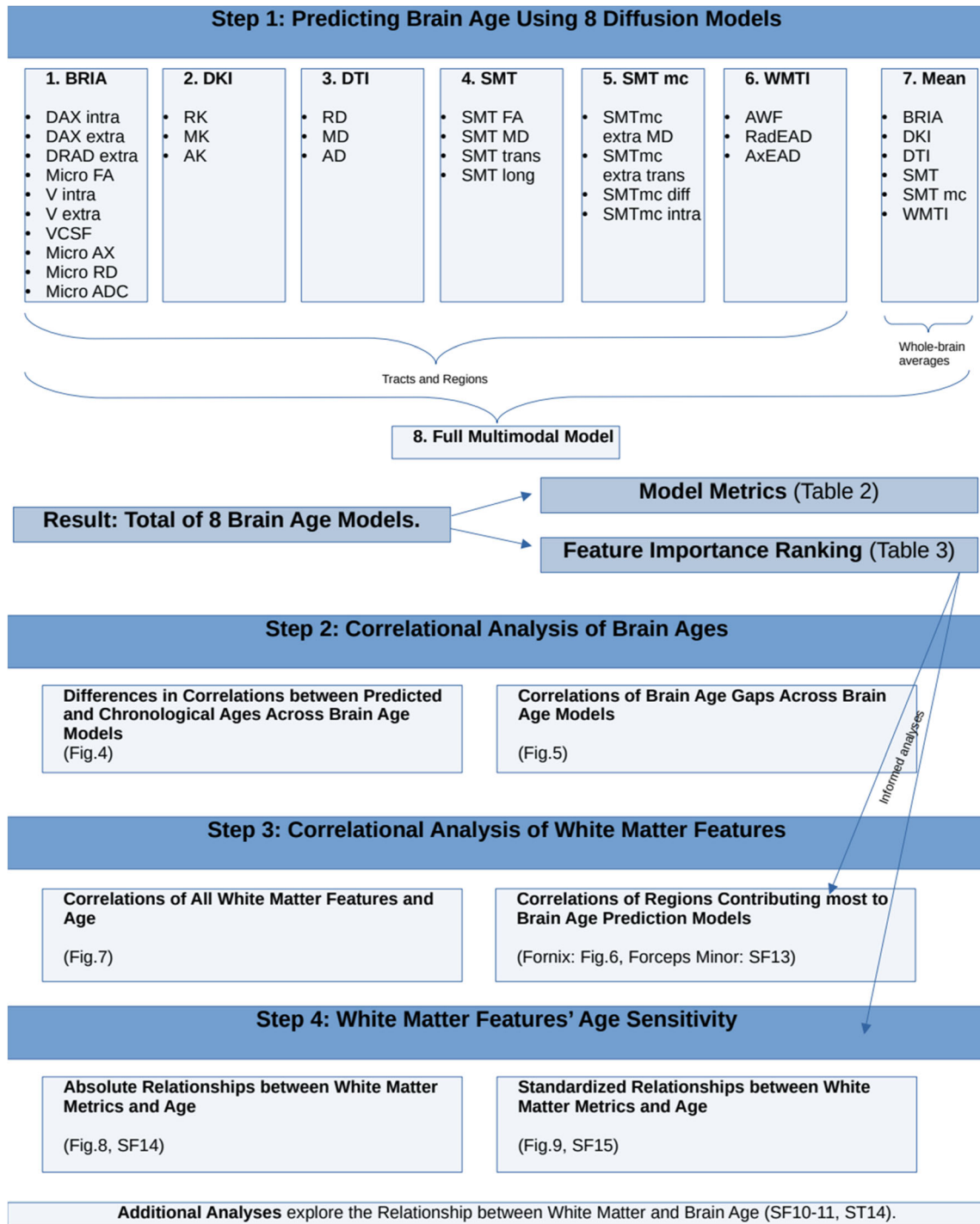
FIGURE 4 Overview of the analysis steps.

Hotelling's (Hotelling, 1936) *t*-tests were used to compare correlations between uncorrected predicted age and chronological age across diffusion models. Zou's (Zou, 2007) method was used to estimate the confidence intervals around the correlation differences (Figure 5 and Table S3; Figure S8 and Table S2 for corrected prediction correlation comparisons). These differences were not significantly different from each other for model pairs DKI and DTI (*p* 1).

All other correlations were different from each other, Pearson's *rs*diff ≤0.15, *p* < .001, with the biggest difference observed between mean and full multimodal scores' correlations (Table S2 for exact values).

Permutation feature importance estimates across diffusion models showed that fornix contributed strongest to variance explained (Table 3), which was in correspondence with feature
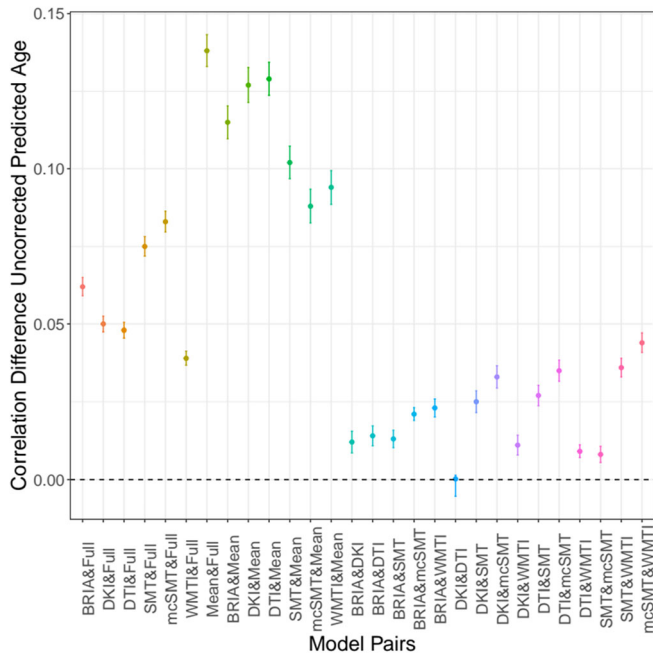
FIGURE 5 Differences between Pearson's correlations of chronological and uncorrected predicted ages across diffusion approaches with 95% confidence interval. Differences between Pearson's correlation coefficients of chronological and uncorrected predicted age by diffusion approach. See Figure S8 for correlational differences between approaches for corrected brain age predictions.

rankings by gain score (XBGoost Developers, 2021; Table S15). Follow-up models which had fornix features removed had lower model fit, explained less variance in age, and predicted-chronological-age correlations were smaller than for models containing fornix ($rs_{diff} < -0.003$, $ps < .001$; Table S16). Another potentially important region was the forceps minor, also contributing significantly to age predictions (Table 3).

3.2 | BAG across diffusion approaches and age

In order to compare uncorrected BAG (BAGu) calculations across the used diffusion approaches, BAGu was correlated from different diffusion approaches and with age. Correlations between the six diffusion approaches ranged between $r = 0.857$ and $r = 0.966$ (Figures 6 and S1 for corrected BAG correlations). Overall, BAGu scores from the different approaches were strongest related to WMTI BAGc (range: $r = 0.873$–0.952), and weakest to mean multimodal BAGu (range: $r = 0.779$–0.828), and could be observed in one cluster containing DKI, DTI, WMTI, and multimodal BAGu and a second cluster containing BRIA, SMT, and SMTmc. However, DKI, BAGu was more strongly correlated with full multimodal BAGc than with other well-performing approaches DTI (Pearson's $rdiff = 0.03$, $p < .001$) and WMTI ($rdiff = 0.03$, $p < .001$). Vice versa, DTI BAGc correlated strongest with WMTI BAGc ($r = 0.905$, $p < .001$).

TABLE 3 Top five diffusion metrics ranked by their contribution to variance explained (R^2) in age.

BRIA	DKI	DTI	SMT	mcSMT	WMTI	Multimodal
Micro FA fornix 0.1954 ± 0.0027	AK right anterior limb of internal capsule 0.0984 ± 0.0014	MD fornix 0.0712 ± 0.0013	MD fornix 0.0795 ± 0.0018	Extratrans fornix 0.0498 ± 0.0013	AWF fornix 0.1699 ± 0.0023	Micro FA fornix 0.0914 ± 0.0011
Vextra forceps minor 0.0278 ± 0.0007	RK fornix 0.0884 ± 0.0016	FA forceps minor 0.0533 ± 0.0011	FA right superior longitudinal fasciculus 0.0267 ± 0.0007	Intra forceps minor 0.0444 ± 0.0009	radEAD fornix to right striaterminalis 0.0283 ± 0.0007	AK anterior limb of internal capsule 0.0055 ± 0.0011
Vextra body of the corpus callosum 0.0261 ± 0.0007	MK left external capsule 0.0259 ± 0.0006	RD fornix to right Striaterminalis 0.0462 ± 0.0009	Longitudinal fornix 0.0251 ± 0.0006	Intra fornix 0.0289 ± 0.0009	AWF forceps minor 0.0194 ± 0.0005	FA forceps minor 0.0219 ± 0.0006
Micro FA fornix to right Striaterminalis 0.0203 ± 0.0006	MK right superior longitudinal fasciculus 0.0214 ± 0.0006	FA right superior cerebellar peduncle 0.0221 ± 0.0006	Trans fornix to right striaterminalis 0.0204 ± 0.0006	Extratrans fornix to right Striaterminalis 0.0201 ± 0.0006	axEAD forceps minor 0.0193 ± 0.0007	RD right fornix stria terminalis 0.0214 ± 0.0006
Vintra right superior cerebellar peduncle 0.0194 ± 0.0006	RK forceps minor 0.0208 ± 0.0005	FA body of the corpus callosum 0.0218 ± 0.0006	FA fornix 0.0192 ± 0.0006	Extratrans right external capsule 0.0163 ± 0.0007	axEAD left posterior limb of internal capsule 0.0173 ± 0.0006	AK Genu corpus callosum 0.0095 ± 0.0003

Note: Variance explained (R^2) by a single feature refers here to the part of the total variance explained by the respective feature in each of the brain age models presented in Table 2. Multimodal refers to an approach using the diffusion metrics from all diffusion approaches. Cells containing fornix are marked in green. Cells containing forceps minor are marked in blue. See Table S19 for an overview of all the features and their variance explained.
Abbreviations: BRIA, Bayesian rotationally invariant approach; DKI, diffusion kurtosis imaging; DTI, diffusion tensor imaging; mcSMT, multicompartment spherical mean technique; SMT, spherical mean technique; WMTI, white matter tract integrity.

3.3 | Associations between diffusion metrics and age

A correlational analysis was used to demonstrate associations among fornix diffusion metrics and age (Figure 7, including QC outliers: Figure S4). Association strengths ranged from to $r = -0.997$ (smtTrans and smtMCintra) to $r = 0.999$ (smtTrans and smtMD). Correlations between fornix metrics and age ranged from $r = -0.558$ (smtMCintra) to $r = 0.570$ (microRD), and between forceps minor metrics and age from $r = -0.519$ (FA) to $r = 0.493$ (RD, see Figure S13).

Correlations across all diffusion metrics and age (1933 × 1933 correlations), age-fornix associations were the strongest (Figure 8, Figure S12). Overall, the significant $N = 1823$ correlations (at $p_{Holm} < .001$) ranged from $|r| = 0.024$ to $|r| = 0.578$ with $|r|_{Mean} = 0.245$, $|r|_{SD} = 0.122$.

3.4 | Age trajectories of diffusion features

In Figure 9 we present absolute diffusion metrics for the whole brain (Figure 9a) and fornix (Figure 9b) across ages for the examined six diffusion approaches (for forceps see Figure S14; overview of metrics: Table S10). Age-metric relationships for fornix were approximating linearity closer than more curvilinear global age-curves.

Several fornix-age relationships for BRIA extra-axonal and intra-axonal radial and axonal diffusivity opposed age relationships of whole-brain-averages, whereas forceps-age relationships closely resembled these whole-brain-average metrics' age relationships.

Whole-brain (Figure 10), fornix (Figure S9), and forceps (Figure S15) diffusion metrics M were predicted from age, sex and scanner site to create age curves (Figure 10a,b) which can be compared to crude curves (Figure 10c,d). Highest SE, R^2adj, and variability across metrics was observed when predicting BRIA metrics (R^2adj = 0.21), as well as lowest R^2adj ≈ 0 in BRIA Vextra, respectively. While DTI metrics could also be predicted well from the model, lowest variability in R^2adj was found in WMTI and DKI. For fornix metrics, SE and R^2adj was generally higher across diffusion approaches (Figure S9).

Likelihood Ratio tests indicated age dependence across global metrics ($p_{Holm} < .001$), with the exception of WMTI axEAD ($\chi^2 = 6.66$, $p_{Holm} = .084$; Table S11), whereas all fornix (Table S4) and forceps (Table S17) features were age sensitive. While the regression lines show a slight curvature, model fit did not differ between linear and nonlinear models for whole-brain (Table S12), fornix metrics (Table S9), and forceps minor metrics (Table S18), indicating steady WM degeneration in mid-life to older ages.
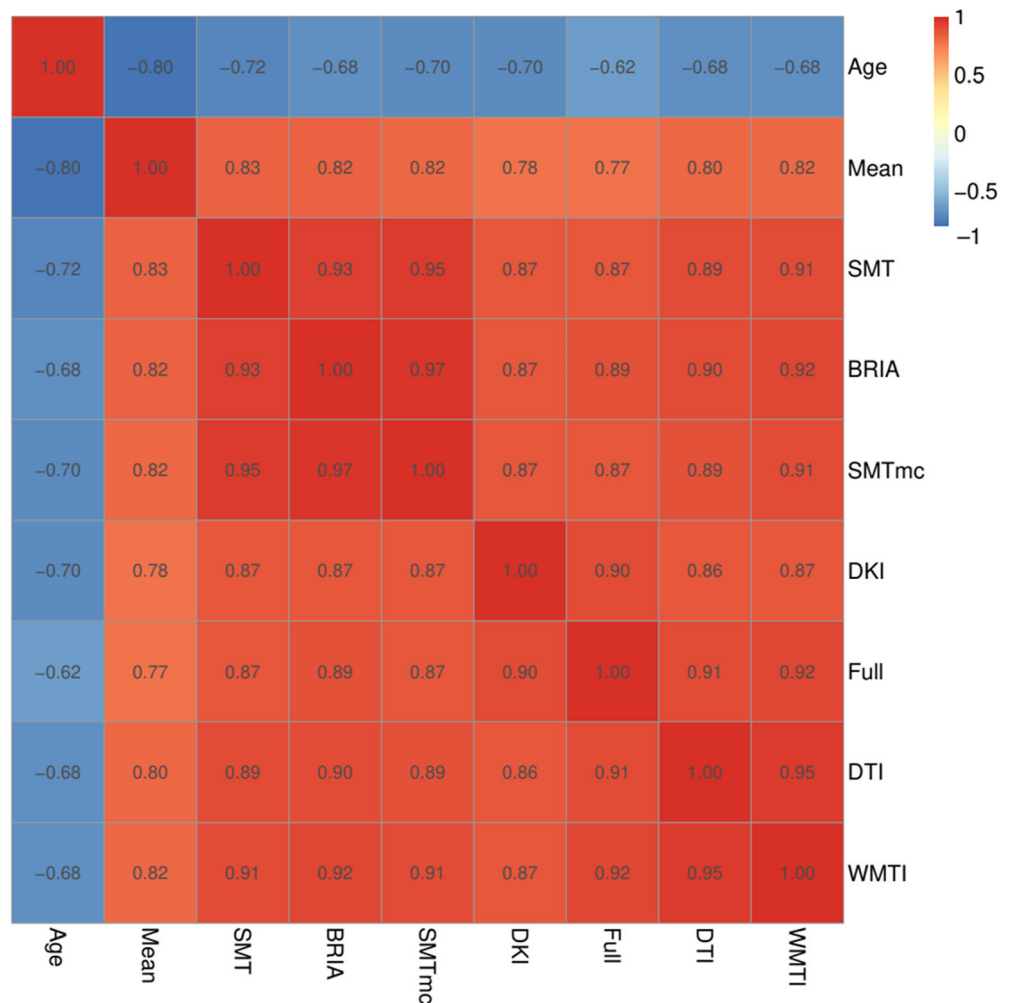
FIGURE 6 Correlations of uncorrected BAG and age across used diffusion approaches. Age-BAG correlations were significant at $p_{Holm} < .001$. For the corrected BAG correlations across models see Figure S1.
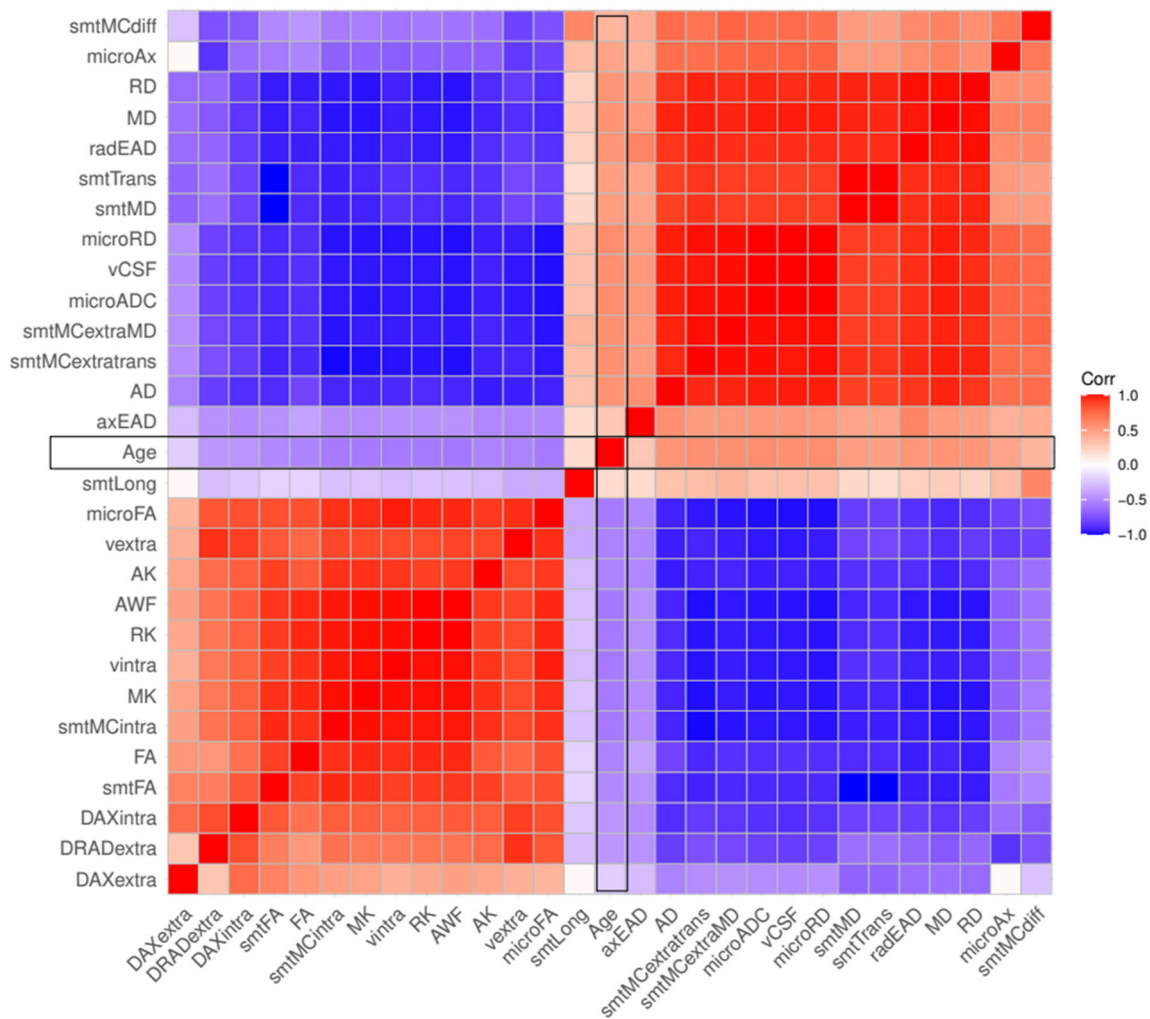
FIGURE 7 Correlation matrix for fornix diffusion metrics and chronological age. All correlations were significant at Holm-corrected *pHolm* < .05.

3.5 | Associations between BAG and WM

Finally, principal components of regional and whole-brain WM metrics for each of the eight models (Table 2) were only weakly correlated with uncorrected BAGu, and similarly related to corrected BAGc, chronological and predicted ages (Figure S10). Furthermore, when predicting either WM components which explain most variability (Figure S10, Table S14) or single regional or whole-brain metrics (Figure S11) from BAGc and BAGu and covariates, models predicted relatively small proportions of variance, with small contributions of BAG to the model (Figures S10, S11).

4 | DISCUSSION

We revealed that both conventional DTI and advanced diffusion approaches (WMTI, DKI, BRIA, SMT, mcSMT) perform consistently on brain age predictions, as indicated previously (Beck et al., 2021). As a novel finding, our results show strong contributions of fornix and

forceps minor microstructures to brain age prediction models. Additionally, among WM features, fornix shows strongest correlations with age. This suggest that the fornix and forceps minor are key WM region of cross-sectional brain age, with fornix and whole-brain dMRI metrics' age trajectories following similar patterns such as steepening slopes at later ages. Furthermore, WM microstructure is expected to steadily degenerate in midlife to older ages, in particular, in extra axonal space.

4.1 | Limitations

There are multiple challenges related to fornix and forceps minor as drivers of brain age estimates, particularly multicollinearity, which might bias estimates of the importance of fornix and forceps minor (gain and permutation feature importance) for brain age predictions, and second, data processing artefacts. UKB offers diffusion data acquired with the most typical two-shell-diffusion protocol. Nevertheless, the standard diffusion model (Novikov et al., 2018) based on

FIGURE 8 Correlations between diffusion metrics and age. Each point indicates one correlation between a diffusion metric and chronological age. Names of diffusion metrics are displayed when correlations between the metric and age reached a Pearson correlation of |*r*| > 0.5. Holm correction (Holm, 1979) was used for Holm-correction, and all displayed values were significant at *p* < .001. For the distribution of the correlations see Figure S12.

differentiation of intra- and extra-axonal water pools could not be solved using this measurement strategy (Novikov et al., 2018). As a result, the derived diffusion metrics have both numerical uncertainties and the variability introduced from nonbiological parameters (Novikov et al., 2018). Quantitative metrics derived from the different diffusion approaches allow to investigate such nonbiological variability and to grade the subject variability in terms of used covariances. Yet, the aforementioned technical limitation might play a decisive role in a clinical context (Novikov et al., 2018; Thomas et al., 2011).

Besides obstacles resulting from modelling assumptions, our sample is cross-sectional in design and limited to adults older than 40, which, in turn, influences predictions (de Lange et al., 2022). Additionally, the UKB imaging subsample shows better health than the non-imaging UKB subjects (Lyall et al., 2022). Another open question is the exact interpretation of BAG and its relationship with WM

metrics. This BAG-WM relationship was found to be small for principal WM components (Figure S10) and single diffusion metrics (Figure S11). Previous research indicates no relationship between the rate of change in longitudinal regional and global T1-weighted-feature-retrieved BAG (Vidal-Pineiro et al., 2021). Yet, further investigation of longitudinal, in particular voxel-wise WM-derived BAG provides additional avenues to increase the interpretability of BAG.

Diffusion metrics were highly correlated within fornix (Figure 7) and forceps (Figure S13) across diffusion approaches, and show similar age trajectories (fornix: Figure S9, forceps: Figure S15). This provokes the question of redundancy of some of the metrics. The identification of redundant metrics and the combination of metrics across diffusion approaches is a matter of future research comparing diffusion approaches by probing them in practical settings such as in clinical samples (Kantarci, 2014).
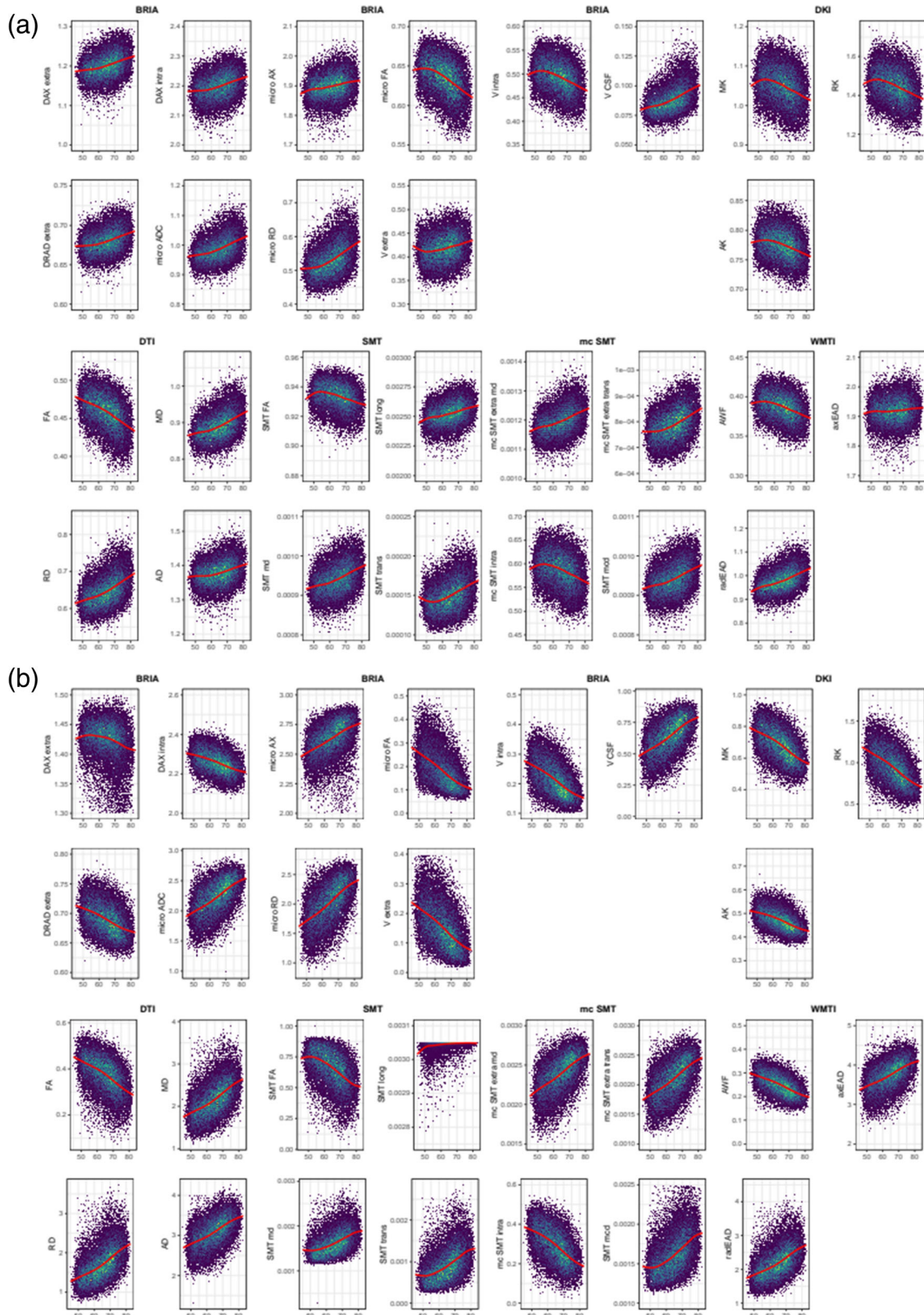
FIGURE 9 Whole-brain and fornix diffusion metrics across age. The presented plots represent diffusion metrics for each of the six diffusion models from the full sample N = 35,749 for (a) whole-brain diffusion metrics, (b) fornix diffusion metrics. Brighter colors indicate higher density and red lines are fitted lines to the relationship between age and diffusion metric. Plots for forceps can be found in Figure S14.

Only few studies (Chen et al., 2015; Christiansen et al., 2016) address the fornix across ages. A possible reason is fornix' artefact-susceptibility induced from its proximity to the cerebrospinal-fluid, while being a small tubular region. Recent processing pipelines such as TBSS minimize such artefacts (Smith et al., 2006). Yet, the influence of cerebrospinal-fluid artefacts in

100   **Paper A**
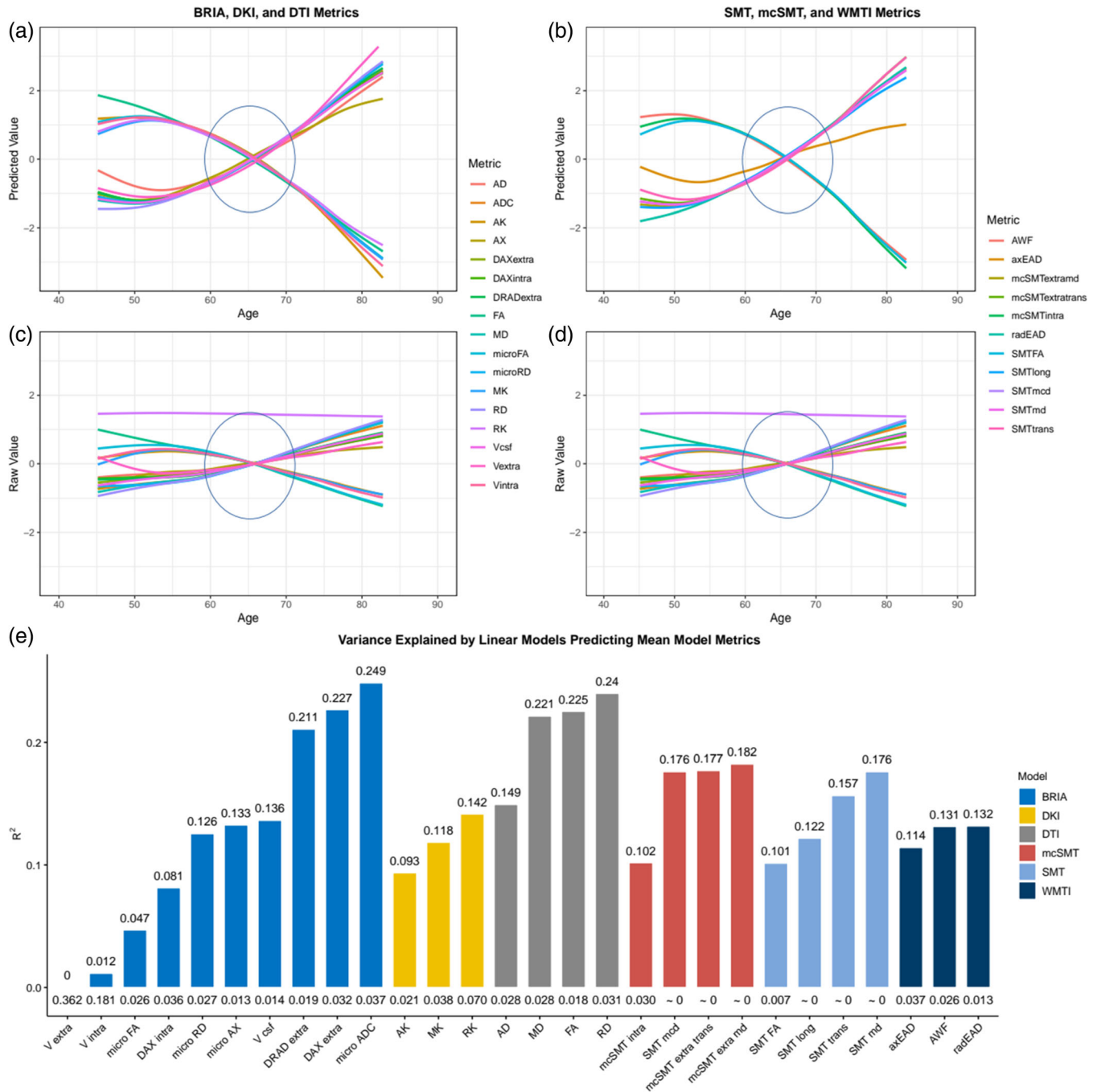
FIGURE 10 Raw and predicted whole-brain WM diffusion metrics by chronological age. Figure 10a–d shows age curves for each standardized (z-score) diffusion metric's mean skeleton value (y-axis) plotted as a function of age (x-axis). Shaded areas represent 95% CI. Curves fitted to raw values (Figure 10c,d) serve as a comparison to the lm-derived predicted values from Equation (4) (Figure 10a,b). Figure 10e indicates the model fit for the linear models from Figure 10a,b, showing R^2adj values on top and standard error (SE) on the bottom of the bars which each represent a Fornix skeleton value for one of the seven models. Lines crossing at age 65 are marked with ovals. Model summaries of all 28 mean models can be found in Table S5. The same visualization of fornix diffusion values can be found in Figure S9, and for the forceps minor in Figure S15.

small tubular structures like the fornix remains unclear (Bach et al., 2014). Fornix is a relatively small anatomical structure, and, for example, fornix BRIA cerebrospinal-fluid fraction is higher (vCSF > 0.5) than global measures (vCSF > 0.075), suggesting a presence of strong partial volume effect. In order to overcome

such distorting effects, voxel-wise techniques are recommended, demanding the development of novel approaches incorporating techniques such as deep learning showing better performance than traditional ML, especially on large population samples (Popescu et al., 2021).

4.2 | Consistency across diffusion approaches

Overall, the results of brain age predictions are similar across diffusion approaches, with WMTI, DTI, and DKI predicting age better than SMT, mcSMT, and BRIA considering model fit and prediction-outcome correlations (Table 2). This finding could be explained in terms of diffusion approaches, that is, the attempt to introduce more biophysically accurate parameters into the model might simultaneously reduce the general sensitivity of the used approaches to tissue changes. Integrative approaches such as DTI or DKI are able to localize brain changes, however, without providing information about the underlying mechanisms. Our study supports a previous study with a smaller but more age-differentiated sample ($n = 702$) of DTI and WMTI being superior to mcSMT at brain age predictions in terms of model performance (Beck et al., 2021). When examining additional diffusion models on a larger sample, and also including JHU ROIs in addition to tract and whole-brain average scores, we find DKI metrics to have higher predictive power than in Beck and colleagues (Beck et al., 2021). This effect might be partly due to added spatial detail from the added RIOs and their relationships to the tracts. Simultaneously, differences between diffusion approaches, and both variance explained and prediction error (root mean squared error, mean absolute error) were smaller in this study. These differences are likely due to the narrower age range in our study (de Lange et al., 2022), whereas our significantly larger sample emphasizes the reliability of our findings.

While brain age predictions from single diffusion approaches were grossly similar, predictions from combined approaches were most accurate (Table 2). Correlations between predicted and chronological age were consistent across diffusion approaches, as differences between correlations were small (Figure 5, Figure S8). This shows that addressing a wider range of WM characteristics improves predictive models compared to models with single diffusion approach metrics (e.g., only DTI), which would be intuitive when considering BAG as a general indicator of health (Beck et al., 2022; Cole et al., 2017; Kaufmann et al., 2019; Leonardsen et al., 2021).Vice versa, reducing spatial specificity by averaging diffusion metrics across all WM reduced prediction accuracy. Conventionally used DTI on its own is limited in its ability to present biophysically meaningful measures of the underlying microstructure. As a result, the advanced modelling is recalled including intra- and extra-axonal spaces and tissue peculiarities being influenced by individual differences in myelin and fiber architecture (crossing/bending fibers, and axonal characteristics; Beck et al., 2021). Hence, adding additional information to DTI better allow to infer the underlying neurobiology of tissue, for example, expressed in differential WM-age-dependences (Figures 9, 10, Figures S14, 15) or brain age predictions (Table 2; Beck et al., 2021).

We observed that BAG exhibits strong correlations across all diffusion approaches (Figure 6, Figure S1). Congruently with the correlational differences (Figure 5, Figure S8), BAG based on averaged skeleton values was least correlated to all other diffusion approaches (Figure 6), indicating inferiority of global compared to region-wide approaches. BAG obtained from WMTI, DTI, and DKI were closest

related to BAG from the multimodal approach (which predicted age best), both for age-bias corrected and uncorrected BAG (Figure 6, Figure S1). This is in agreement with the observed age-prediction model performance (Table 2). BAG correlations were observed in three clusters: (1) WMTI and DTI, (2) mcSMT, SMT, BRIA, and (3) DKI, indicative of similar measurements within these clusters (Figure 6, Figure S1). To a certain extent, these clusters reflect similarities in the underlying mathematics of the clustering diffusion approaches. For example, mcSMT and SMT are closely related models (Kaden et al., 2016a), whereas DKI's non-Gaussianity might reveal another quality of age-sensitive WM microstructures not captured by the other approaches (De Santis et al., 2011). Additionally, the cluster differences indicate that the observed diffusion approaches measure different age(ing)-sensitive characteristics, supporting the argument for a combination of diffusion approaches when assessing the ageing brain.

4.3 | Age trajectories and fornix and forceps minor as a brain age feature

Based on the presented findings on fornix, we further investigate details of fornix, keeping discussed limitations to the generalizability of the findings in mind. Diffusion metrics describing fornix microstructure were consistently related to each other and age across all diffusion approaches in two clusters. Values were positively correlated within each cluster and negatively between clusters (see Figure 7). In the first cluster, different approaches' FA, kurtosis metrics (MK, RK, AK), water fractions (vintra and vextra from BRIA and AWF from WMTI), and BRIA intra-axonal and extra-axonal radial and AD were positively correlated. The second cluster, which was negatively related to the first cluster but positive to age, contained metrics of MD, AD, and RD, and cerebrospinal-fluid fraction of the different diffusion approaches, which were positively related to each other. Interestingly, both clusters consisted of unit-less values, for example, water fractions, and diffusivities, which might have the same meaning as extra-axonal ADs from different diffusion approaches, for example, BRIA versus SMTmc. Such consistencies of similar metrics across diffusion approaches were more apparent for the fornix when QC-identified outliers were removed (compare Figure 7 and Figure S4), which supports the reliability of our findings of fornix-age-dependencies. Furthermore, fornix metrics were most strongly related to age across diffusion approaches (Figure 8, Figure S11), supporting the importance of fornix in reducing error of brain age predictions (Table 3). Correlations of diffusion metrics within the forceps minor were not as strong and consistent as in the fornix, and partly in the opposite direction as for the fornix (Figure S13). Not surprisingly, all fornix and forceps minor features were age-sensitive (Tables S4, S17), and more age sensitive than whole-brain metrics (compare: Table S11). Whole-brain trajectories are in agreement with previous results, showing-age sensitivity of various mean diffusion metrics (Beck et al., 2021), and the same directionality of age trajectories of metrics for DTI (Cox et al., 2016a; Westlye et al., 2010), mcSMT, DKI, WMTI (Beck et al., 2021).

We displayed differential behaviors of fornix microstructure measures across diffusion approaches (Figures 9, 10). Focusing on absolute diffusion values (Figure 9), diffusion measures which are correlated (Figures 6, 7, Figure S13) exhibit similar age dependences. Additionally, slopes of fornix compared to whole-brain diffusion metrics were generally steeper and closer approximating linearity, indicating stronger changes, such as quicker WM degeneration in the fornix compared to the whole-brain average (see Figure 9). Particularly BRIA metrics show visually detectable differences between the fornix and the whole brain (Figure 9, DAXextra, DAXintra, DRADextra, Vextra); as opposed to global age trends which are also strongly resembled by forceps minor (Figure S14), fornix intra and extra-axonal diffusion decreased, indicating fornix shrinkage with increasing age. Periventricular shrinkage is linked to enlarging ventricles (Kwon et al., 2014), which has been related to ageing and neurodegenerative disorder progression (Pinaya et al., 2021). This effect was observed by a positive relationship between age and cerebrospinal fluid (CSF) fraction in BRIA. Another metric which revealed larger differences in the fornix than for the whole-brain average was intra-axonal water fractions, which can be treated as a proxy for the axonal density, decreased with increasing age (see Figure 9, BRIA:Vintra; SMTmc:intra; WMTI:AWF) while the CSF fraction (BRIA) increases. Such WM microstructure changes are not only directly linked to different neurobiological features but can be markers of clinical outcomes, such as dementia (Meeter et al., 2017; Thomas et al., 2011).

A selection of metrics is comparable across diffusion approaches when taking DTI as reference point and focusing on similar age trends. DTI metrics AD, RD, and MD tend to increase over the lifespan and FA tends to decrease across brain regions (Figures 9, 10; Beck et al., 2021; Cox et al., 2016b; Davis et al., 2009; Westlye et al., 2010) as well as in fornix (Figure 9b, Figure S9), implying processes such as de-myelination, changes in axonal and general WM integrity. Such DTI-age-dependencies are reflected by according BRIA, SMT, and WMTI metrics, whereas DKI shows opposite age-relationships, as presented previously (Beck et al., 2021). Deterioration effects, measured by the age-dependency of axonal water fractions, were generally stronger in fornix compared to whole-brain metrics (Figure 9). Interestingly, opposed to global metrics, radial diffusivity measures from DKI and BRIA (DRADextra) decreased in fornix (Figure 9), suggesting higher fornix than global plasticity, potentially being an antecedent of age-related hippocampal changes (Metzler-Baddeley et al., 2019).

Additional, unique information about age dynamics was presented by standardized scores corrected for age, sex, and scanner site and crude standardized scores across ages (Figure 10, Figure S9). After corrections, most fornix metrics follow a tightly resembling near-linear trend either increasing or decreasing by age (Figure S9a,b), as opposed to forceps minor (Figure S15) and whole-brain metrics which follow a rather curvilinear line, as previously shown (Beck et al., 2021; Davis et al., 2009; Westlye et al., 2010). Diffusion metrics' variance explained across models indicates fornix metrics to be more sensitive to a combination of covariates age, sex, and scanner site than whole-brain metrics (Figure 10, Figure S9). In the fornix, only BRIA extra-axonal AD (DAX extra) and the SMT longitudinal diffusion

coefficient (SMT long) showed non-linear trajectories, however, both measures are weakly correlated to other diffusion parameters (Figure 10). Yet, when comparing model metrics such as variance explained of linear and nonlinear models predicting fornix, forceps minor, and whole-brain diffusion metrics from age, sex, and scanner site and their interactions, there were no apparent differences between models (Tables S9, S12, S15). This implies that contrary to previous research observing the entire lifespan presenting curvilinear DTI age trajectories (Beck et al., 2021; Westlye et al., 2010), or trends toward curvilinearity (with yet better linear fit for selected regions; Davis et al., 2009), we found that fornix and whole-brain age trajectories from age 40 can be described as linear when accounting for covariates sex, age, and scanner site. While the crossing of the x-axis at age 65 (Figure 10, Figures S9, S15) is a reflection of the sample's age distribution (Figure 1), in addition to the shapes of the different age-trajectories, it reveals that the different diffusion approaches are similarly age-sensitive or measure similar underlying ageing-related changes. For whole-brain metrics, changes become exacerbated from 65 onward (Figure 1), with reasons potentially laying in an accelerated neurodegeneration also reflected in the exponentially increasing risk to develop neurodegenerative disorders from age 65 onward (Nichols et al., 2022). For example, in the USA, 3% of 65–74 year olds, 17% of the 75–84 year olds, and 32% of those aged 85+ developed Alzheimer's dementia (Alzheimer's Association, 2020). Subclinical or preclinical states are, however, not captured by these approximations, and WM changes usually precede clinical detections. This makes WM monitoring a promising tool for early neurodegenerative disease detection.

Beyond WM, fornix changes seem to play an important role for GM changes, particularly in the hippocampus: for example, fornix glia damages lead to hippocampal GM atrophy (Metzler-Baddeley et al., 2019). This might be reflected by dis-connectivity of fornix with other brain regions as described by decreasing extra axonal space coefficients (Figure 8b), and following changes in fornix function. Potentially, the consequences of age-related fornix changes thereby affect functionality of a selection of brain regions, such as the hippocampus. While several studies have presented ageing-related fornix microstructure changes in humans (Chen et al., 2015; Christiansen et al., 2016) and monkeys (Peters et al., 2010) in small samples, only one large-scale study revealed findings connected to the fornix, namely strongest default mode network GM volume covariation with fornix WM microstructure (Kernbach et al., 2018). This suggests that fornix, a key connector of the limbic system with the cortex, might also be critical for default mode network functioning. Moreover, memory and episodic recall have been related to fornix (Senova et al., 2020). Hence, fornix changes might play an important role in known ageing-dependent temporal lobe changes, and specifically hippocampal changes for ageing-related pathological developments (Cabeza et al., 2018; Burke & Barnes, 2006; Hedden & Gabrieli, 2004; Pluvinage & Wyss-Coray, 2020). Previous studies presented age-related fornix DTI metric changes (Chen et al., 2015; Christiansen et al., 2016; Metzler-Baddeley et al., 2019) which potentially appear prior to hippocampal volume changes (Chen et al., 2015; Metzler-

Baddeley et al., 2019), and are related to declining episodic memory performance (Metzler-Baddeley et al., 2019). Hence, fornix changes potentially serve to predict future pathological development, suggesting fornix WM microstructure and changes in such as ageing biomarkers. This supports previous findings showing network re-activations, metabolic, and GM changes after fornix deep-brain-stimulation, antagonizing the progression of neurodegenerative disorders (Jakobs et al., 2020).

Different studies showed age-related deterioration effects in the forceps minor (Bastin et al., 2008; Fan et al., 2019), a subregion of the corpus callosum. Loss in WM integrity have also been associated with various phenotypes, for example, behavioral impacts, such as mental slowing (Jokinen et al., 2007), and various disorders, such as major depressive disorder (Won et al., 2016), schizophrenia (Kelly et al., 2018), dependencies on cocaine (Moeller et al., 2005) and alcohol (Pfefferbaum & Sullivan, 2005), with WM degeneracy explaining higher impulsivity in cocaine addiction (Moeller et al., 2005). Overall, the forceps are assumed to have an important role of connecting both hemispheres, which might be crucial for interhemispheric signal propagation (Voineskos et al., 2010). Previous research shows also that WM changes in FA and MD relate to GM thinning with the forceps being particularly vulnerable to such changes (Storsve et al., 2016). Moreover, cognitive test scores were related to forceps minor AD and MD scores in Alzheimer's Disease patients (Tu et al., 2017), and already at mild cognitive impaired forceps minor FA and MD scores were different from age-matched participants with subjective cognitive decline (Luo et al., 2020). FA was also shown in this study as important brain age feature for both multimodal and DTI models (Table 3). This suggests forceps as an important region for brain age and ageing.

The current study gives for the first time a detailed account on region-wise-to-global WM-age relationships for multiple diffusion approaches in a representative sample, and highlights fornix and forceps minor as an important structures for age predictions across diffusion approaches. Brain age was estimated best when combining diffusion approaches, showing different aspects of WM to contribute to brain age with fornix and forceps minor being the central regions for these predictions. Trained models are made available for further research to extend the reported brain age predictions to other samples (e.g., to clinical samples with a similar age structure), in addition to examining the discussed metrics in fornix and forceps.

AUTHOR CONTRIBUTIONS

Max Korbmacher: Study design, Software, Formal analysis, Visualizations, Project administration, Writing—original draft, Writing—review & editing. Ann Marie de Lange: Software, Writing—review & editing. Dennis van der Meer: Software, Writing – review & editing. Arvid Lundervold: Writing—review & editing, Funding acquisition. Eli Eikefjord: Writing—review & editing. Dani Beck: Writing—review & editing. Ole A. Andreassen: Writing—review & editing, Funding acquisition. Lars T. Westlye: Writing—review & editing, Funding acquisition. Ivan I. Maximov: Supervision, Study design, Data preprocessing and quality control, Writing—review & editing, Funding acquisition.

CONFLICT OF INTEREST STATEMENT

OOA has received a speaker's honorarium from Lundbeck and is a cosultant to Coretechs.ai.

DATA AVAILABILITY STATEMENT

All raw data are available from the UKB5 (www.ukbiobank.ac.uk). Synthetic datasets with the synthpop 2016 R package based on the original data for all six diffusion approaches (resulting in six datasets) to run the code and code needed to run brain age predictions in Python are openly available at the Open Science Framework: (https://osf.io/nv8ea/). Synthetic datasets are simulated datasets closely mimicking the statistical characteristics of the original data while protecting data privacy and anonymity. Finally, also the trained XG Boost models are made available in the same depository.

ORCID

Max Korbmacher https://orcid.org/0000-0002-8113-2560
Lars T. Westlye https://orcid.org/0000-0001-8644-956X

REFERENCES

Alzheimer's Association. (2020). Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, *16*, 391–460.

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078.

Bach, M., Laun, F. B., Leemans, A., Tax, C. M. W., Biessels, G. J., Stieltjes, B., & Maier-Hein, K. H. (2014). Methodological considerations on tract-based spatial statistics (TBSS). *NeuroImage*, *100*, 358–369.

Baecker, L., Garcia-Dias, R., Vieira, S., Scarpazza, C., & Mechelli, A. (2021). Machine learning for brain age prediction: Introduction to methods and clinical applications. *eBioMedicine*, *72*, 103600.

Barrick, T. R., Charlton, R. A., Clark, C. A., & Markus, H. S. (2010). White matter structural decline in normal ageing: A prospective longitudinal study using tract-based spatial statistics. *NeuroImage*, *51*, 565–577.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*, 259–267.

Bastin, M. E., Piatkowski, J. P., Storkey, A. J., Brown, L. J., MacLullich, A. M. J., & Clayden, J. D. (2008). Tract shape modelling

provides evidence of topological change in corpus callosum genu during normal ageing. *NeuroImage*, *43*, 20–28.

Beck, D., de Lange, A. M. G., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., & Westlye, L. T. (2021). 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.

Beck, D., Lange, A. M. G., Pedersen, M. L., Alnæs, D., Maximov, I. I., Voldsbekk, I., Richard, G., Sanders, A. M., Ulrichsen, K. M., Dørum, E. S., Kolskår, K. K., Høgestøl, E. A., Steen, N. E., Djurovic, S., Andreassen, O. A., Nordvik, J. E., Kaufmann, T., & Westlye, L. T. (2022). Cardiometabolic risk factors associated with brain age and accelerate brain ageing. *Human Brain Mapping*, *43*, 700–720.

Benitez, A., Jensen, J. H., Falangola, M. F., Nietert, P. J., & Helpern, J. A. (2018). Modeling white matter tract integrity in aging with diffusional kurtosis imaging. *Neurobiology of Aging*, *70*, 265–275.

Burke, S. N., & Barnes, C. A. (2006). Neural plasticity in the ageing brain. *Nature Reviews. Neuroscience*, *71*(7), 30–40.

Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018). Maintenance, reserve and compensation: The cognitive neuroscience of healthy ageing. *Nature Reviews. Neuroscience*, *19*, 701–710. https://doi.org/10.1038/s41583-018-0068-2

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

Chen, D. Q., Strauss, I., Hayes, D. J., Davis, K. D., & Hodaie, M. (2015). Age-related changes in diffusion tensor imaging metrics of fornix subregions in healthy humans. *Stereotactic and Functional Neurosurgery*, *93*, 151–159.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, August 14, 2017, 14, 785–794.

Christiansen, K., Aggleton, J. P., Parker, G. D., O'Sullivan, M. J., Vann, S. D., & Metzler-Baddeley, C. (2016). The status of the precommissural and postcommissural fornix in normal ageing and mild cognitive impairment: An MRI tractography study. *NeuroImage*, *130*, 35–47.

Cole, J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: Relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of Aging*, *92*, 34–42.

Cole, J. H., & Franke, K. (2017). Predicting age using neuroimaging: Innovative brain ageing biomarkers. *Trends in Neurosciences*, *40*, 681–690. https://doi.org/10.1016/j.tins.2017.10.001

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

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., Wardlaw, J. M., Gale, C. R., Bastin, M. E., & Deary, I. J. (2016b). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nature Communications*, *7*, 1-13.

Davis, S. W., Dennis, N. A., Buchler, N. G., White, L. E., Madden, D. J., & Cabeza, R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *NeuroImage*, *46*, 530–541.

de Lange, A. M. G., Anatürk, M., Rokicki, J., Han, L. K. M., Franke, K., Alnæs, D., Ebmeier, K. P., Draganski, B., Kaufmann, T., Westlye, L. T., Hahn, T., & Cole, J. H. (2022). Mind the gap: Performance metric evaluation in brain-age prediction. *Human Brain Mapping*, *43*, 3113–3129.

de Lange, A. M. G., Anatürk, M., Suri, S., Kaufmann, T., Cole, J. H., Griffanti, L., Zsoldos, E., Jensen, D. E. A., Filippini, N., Singh-Manoux, A., Kivimäki, M., Westlye, L. T., & Ebmeier, K. P. (2020). Multimodal brain-age prediction and cardiovascular risk: The Whitehall II MRI sub-study. *NeuroImage*, *222*, 117292.

de Lange, A. M. G., & Cole, J. H. (2020). Commentary: Correction procedures in brain-age prediction. *NeuroImage Clin.*, *26*, 102229.

de Lange, A. M. G., Kaufmann, T., Quintana, D. S., Winterton, A., Andreassen, O. A., Westlye, L. T., & Ebmeier, K. P. (2021). Prominent health problems, socioeconomic deprivation, and higher brain age in lonely and isolated individuals: A population-based study. *Behavioural Brain Research*, *414*, 113510.

de Lange, A. M. G., Kaufmann, T., van der Meer, D., Maglanoc, L. A., Alnæs, D., Moberget, T., Douaud, G., Andreassen, O. A., & Westlye, L. T. (2019). Population-based neuroimaging reveals traces of childbirth in the maternal brain. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 22341–22346.

De Santis, S., Gabrielli, A., Palombo, M., Maraviglia, B., & Capuani, S. (2011). Non-Gaussian diffusion imaging: A brief practical review. *Magnetic Resonance Imaging*, *29*, 1410–1416.

Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Jr., Barch, D. M., Petersen, S. E., & Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science*, *329*, 1358–1361.

Eikenes, L., Visser, E., Vangberg, T., & Håberg, A. K. (2023). Both brain size and biological sex contribute to variation in white matter microstructure in middle-aged healthy adults. *Human Brain Mapping*, *44*, 691–709.

Fan, Q., Tian, Q., Ohringer, N. A., Nummenmaa, A., Witzel, T., Tobyne, S. M., Klawiter, E. C., Mekkaoui, C., Rosen, B. R., Wald, L. L., Salat, D. H., & Huang, S. Y. (2019). Age-related alterations in axonal microstructure in the corpus callosum measured by high-gradient diffusion MRI. *NeuroImage*, *191*, 325–336.

Fieremans, E., Jensen, J. H., & Helpern, J. A. (2011). White matter characterization with diffusional kurtosis imaging. *NeuroImage*, *58*, 177–188.

Franke, K., & Gaser, C. T. (2019). Ten years of BrainAGE as a neuroimaging biomarker of brain aging: What insights have we gained? *Frontiers in Neurology*, *10*, 789.

Franke, K., Ziegler, G., Klöppel, S., & Gaser, C. (2010). Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. *NeuroImage*, *50*, 883–892.

Grady, C. L. (2012). The cognitive neuroscience of ageing. *Nature Reviews. Neuroscience*, *13*, 491–505.

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews. Neuroscience*, *52*(5), 87–96.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*, 65–70.

Hope, T. R., Selnes, P., Rektorová, I., Anderkova, L., Nemcova-Elfmarkova, N., Balážová, Z., Dale, A., Bjørnerud, A., & Fladby, T. (2019). Diffusion tensor and restriction spectrum imaging reflect different aspects of neurodegeneration in Parkinson's disease. *PLoS One*, *14*, e0217922.

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, *28*, 321–377.

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., Calabresi, P. A., Pekar, J. J., van Zijl, P. C. M., & Mori, S. (2008). Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, *39*, 336–347.

Jakobs, M., Lee, D. J., & Lozano, A. M. (2020). Modifying the progression of Alzheimer's and Parkinson's disease with deep brain stimulation. *Neuropharmacology*, *171*, 107860.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*, 782–790.

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, *53*, 1432–1440.

Jokinen, H., Ryberg, C., Kalska, H., Ylikoski, R., Rostrup, E., Stegmann, M. B., Waldemar, G., Madureira, S., Ferro, J. M., van Straaten, E. C. W., Scheltens, P., Barkhof, F., Fazekas, F., Schmidt, R., Carlucci, G., Pantoni, L., Inzitari, D., Erkinjuntti, T., & LADIS group. (2007). Corpus callosum atrophy is associated with mental slowing and executive deficits in subjects with age-related white matter hyperintensities: The LADIS study. *Journal of Neurology, Neurosurgery, and Psychiatry*, 78, 491–496.

Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D., & Alexander, D. C. (2016b). Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139, 346–359.

Kaden, E., Kruggel, F., & Alexander, D. C. (2016a). Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic Resonance in Medicine*, 75, 1752–1763.

Kantarci, K. (2014). Fractional anisotropy of the fornix and hippocampal atrophy in Alzheimer's disease. *Frontiers in Aging Neuroscience*, 6, 316.

Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., Alnæs, D., Barch, D. M., Baur-Streubel, R., Bertolino, A., Bettella, F., Beyer, M. K., Bøen, E., Borgwardt, S., Brandt, C. L., Buitelaar, J., Celius, E. G., Cervenka, S., Conzelmann, A., … Westlye, L. T. (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*, 22, 1617–1623.

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Resonance in & 2016, undefined. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Wiley Online Library*, 76, 1574–1581.

Kelly, S., Jahanshad, N., Zalesky, A., Kochunov, P., Agartz, I., Alloza, C., Andreassen, O. A., Arango, C., Banaj, N., Bouix, S., Bousman, C. A., Brouwer, R. M., Bruggemann, J., Bustillo, J., Cahn, W., Calhoun, V., Cannon, D., Carr, V., Catts, S., … Donohoe, G. (2018). Widespread white matter microstructural differences in schizophrenia across 4322 individuals: Results from the ENIGMA schizophrenia DTI working group. *Molecular Psychiatry*, 23, 1261–1269.

Kernbach, J. M., Yeo, B. T. T., Smallwood, J., Margulies, D. S., Thiebaut de Schotten, M., Walter, H., Sabuncu, M. R., Holmes, A. J., Gramfort, A., Varoquaux, G., Thirion, B., & Bzdok, D. (2018). Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 12295–12300.

Kochunov, P., Thompson, P. M., Lancaster, J. L., Bartzokis, G., Smith, S., Coyle, T., Royall, D. R., Laird, A., & Fox, P. T. (2007). Relationship between white matter fractional anisotropy and other indices of cerebral health in normal aging: Tract-based spatial statistics study of aging. *NeuroImage*, 35, 478–487.

Kwon, Y. H., Jang, S. H., & Yeo, S. S. (2014). Age-related changes of lateral ventricular width and periventricular white matter in the human brain: A diffusion tensor imaging study. *Neural Regeneration Research*, 9, 986.

Lawrence, K. E., Nabulsi, L., Santhalingam, V., Abaryan, Z., Villalon-Reina, J. E., Nir, T. M., Ba Gari, I., Zhu, A. H., Haddad, E., Muir, A. M., Laltoo, E., Jahanshad, N., & Thompson, P. M. (2021). Age and sex effects on advanced white matter microstructure measures in 15,628 older adults: A UK biobank study. *Brain Imaging and Behavior*, 15, 2813–2823.

Leonardsen, E. H., Peng, H., Kaufmann, T., Agartz, I., Andreassen, O. A., Celius, E. G., Espeseth, T., Harbo, H. F., Høgestøl, E. A., de Lange, A.-M., Marquand, A. F., Vidal-Piñeiro, D., Roe, J. M., Selbæk, G., Sørensen, Ø., Smith, S. M., Westlye, L. T., Wolfers, T., & Wang, Y. (2021). Deep neural networks learn general and clinically relevant representations of the ageing brain. *NeuroImage*, 256, 119210.

Luo, C., Li, M., Qin, R., Chen, H., Yang, D., Huang, L., Liu, R., Xu, Y., Bai, F., & Zhao, H. (2020). White matter microstructural damage as an early sign of subjective cognitive decline. *Frontiers in Aging Neuroscience*, 11, 378.

Lyall, D. M., Quinn, T., Lyall, L. M., Ward, J., Anderson, J. J., Smith, D. J., Stewart, W., Strawbridge, R. J., Bailey, M. E. S., & Cullen, B. (2022). Quantifying bias in psychological and physical health in the UK biobank imaging sub-sample. *Brain Communications*, 4, fcac1.

Marek, S., et al. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nat.*, 6037902(603), 654–660.

Mathworks. (2017). *Matlab* version 9.3.0.713579 (R2017b). IBM.

Maximov, I. I., Alnæs, D., & Westlye, L. T. (2019). Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: Effects of artefact corrections on diffusion metrics and their age associations in UK biobank. *Human Brain Mapping*, 40, 4146–4162.

Maximov, I. I., Meer, D., Lange, A. M. G., Kaufmann, T., Shadrin, A., Frei, O., Wolfers, T., & Westlye, L. T. (2021). Fast qualitY conTrol meThod foR derIved diffUsion metrics (YTTRIUM) in big data analysis: U.K. biobank 18,608 example. *Human Brain Mapping*, 42, 3141–3155.

Meeter, L. H., Kaat, L. D., Rohrer, J. D., & Van Swieten, J. C. (2017). Imaging and fluid biomarkers in frontotemporal dementia. *Nature Reviews. Neurology*, 13, 406–419.

Metzler-Baddeley, C., Mole, J. P., Sims, R., Fasano, F., Evans, J., Jones, D. K., Aggleton, J. P., & Baddeley, R. J. (2019). Fornix white matter glia damage causes hippocampal gray matter damage during age-dependent limbic decline. *Scientific Reports*, 91(9), 1–14.

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., … Smith, S. M. (2016). Multimodal population brain imaging in the UK biobank prospective epidemiological study. *Nature Neuroscience*, 19, 1523–1536.

Moeller, F. G., Hasan, K. M., Steinberg, J. L., Kramer, L. A., Dougherty, D. M., Santos, R. M., Valdes, I., Swann, A. C., Barratt, E. S., & Narayana, P. A. (2005). Reduced anterior corpus callosum white matter integrity is related to increased impulsivity and reduced discriminability in cocaine-dependent subjects: Diffusion tensor imaging. *Neuropsychopharmacology*, 30, 610–617.

Mori, S., & Wakana, S. (2005). *Zijl*, P. Van & Nagae-Poetscher.

Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., al Hamad, H., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., Ashraf, T., … Vos, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019. *The Lancet Public Health*, 7, e105–e125.

Novikov, D. S., Fieremans, E., Jespersen, S. N., & Kiselev, V. G. (2019). Quantifying brain microstructure with diffusion MRI: Theory and parameter estimation. *NMR in Biomedicine*, 32, e3998.

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magnetic Resonance in Medicine*, 79, 3172–3193.

Nowok, B., Raab, G. M., & Dibben, C. (2016). Synthpop: Bespoke creation of synthetic data in R. *Journal of Statistical Software*, 74, 1–26.

Peters, A., Sethares, C., & Moss, M. B. (2010). How the primate fornix is affected by age. *The Journal of Comparative Neurology*, 518, 3962–3980.

Pfefferbaum, A., & Sullivan, E. V. (2005). Disruption of brain white matter microstructure by excessive intracellular and extracellular fluid in alcoholism: Evidence from diffusion tensor imaging. *Neuropsychopharmacology*, 30, 423–432.

Pinaya, W. H. L., Scarpazza, C., Garcia-Dias, R., Vieira, S., Baecker, L., F da Costa, P., Redolfi, A., Frisoni, G. B., Pievani, M., Calhoun, V. D., Sato, J. R., & Mechelli, A. (2021). Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study. *Scientific Reports*, 11, 15746.

Pines, A. R., Cieslak, M., Larsen, B., Baum, G. L., Cook, P. A., Adebimpe, A., Dávila, D. G., Elliott, M. A., Jirsaraie, R., Murtha, K., Oathes, D. J., Piiwaa, K., Rosen, A. F. G., Rush, S., Shinohara, R. T., Bassett, D. S., Roalf, D. R., & Satterthwaite, T. D. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Developmental Cognitive Neuroscience*, 43, 100788.

Pluvinage, J. V., & Wyss-Coray, T. (2020). Systemic factors as mediators of brain homeostasis, ageing and neurodegeneration. *Nature Reviews. Neuroscience*, 21, 93–102.

Popescu, S. G., Glocker, B., Sharp, D. J., & Cole, J. H. (2021). Local brain-age: A U-net model. *Frontiers in Aging Neuroscience*, *13*, 838.

Raghavan, S., Reid, R. I., Przybelski, S. A., Lesnick, T. G., Graff-Radford, J., Schwarz, C. G., Knopman, D. S., Mielke, M. M., Machulda, M. M., Petersen, R. C., Jack, C. R., Jr., & Vemuri, P. (2021). Diffusion models reveal white matter microstructural changes with ageing, pathology and cognition. *Brain Communications*, *3*, fcab106.

Reisert, M., Kellner, E., Dhital, B., Hennig, J., & Kiselev, V. G. (2017). Disentangling micro from mesostructure by diffusion MRI: A Bayesian approach. *NeuroImage*, *147*, 964–975.

Richard, G., Kolskår, K., Sanders, A.-M., Kaufmann, T., Petersen, A., Doan, N. T., Sánchez, J. M., Alnæs, D., Ulrichsen, K. M., Dørum, 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*, *2018*, e5908.

Rokicki, J., Wolfers, T., Nordhøy, W., Tesli, N., Quintana, D. S., Alnæs, D., Richard, G., Lange, A. M. G., Lund, M. J., Norbom, L., Agartz, I., Melle, I., Nærland, T., Selbæk, G., Persson, K., Nordvik, J. E., Schwarz, E., Andreassen, O. A., Kaufmann, T., & Westlye, L. T. (2021). Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Human Brain Mapping*, *42*, 1714–1726.

Salih, A., Galazzo, I. B., Raisi-Estabragh, Z., Rauseo, E., Gkontra, P., Petersen, S. E., Lekadir, K., Altmann, A., Radeva, P., & Menegaz, G. (2021). Brain age estimation at tract group level and its association with daily life measures, cardiac risk factors and genetic variants. *Scientific Reports*, *111*(11), 1–14.

Senova, S., Fomenko, A., Gondard, E., & Lozano, A. M. (2020). Anatomy and function of the fornix in the context of its potential as a therapeutic target. *Journal of Neurology, Neurosurgery, and Psychiatry*, *91*, 547–559.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*, 143–155.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., & Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*, 1487–1505.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., de Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., de Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219.

Storsve, A. B., Fjell, A. M., Yendiki, A., & Walhovd, K. B. (2016). Longitudinal changes in white matter tract integrity across the adult lifespan and its relation to cortical thinning. *PLoS One*, *11*, e0156770.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Medicine*, *12*, e1001779.

Symms, M., Jäger, H. R., Schmierer, K., & Yousry, T. A. (2004). A review of structural magnetic resonance neuroimaging. *Journal of Neurology, Neurosurgery, and Psychiatry*, *75*, 1235–1244.

Thomas, A. G., Koumellis, P., & Dineen, R. A. (2011). The fornix in health and disease: An imaging review. *Radiographics*, *31*, 1107–1121.

Tu, M. C., Lo, C. P., Huang, C. F., Hsu, Y. H., Huang, W. H., Deng, J. F., & Lee, Y. C. (2017). Effectiveness of diffusion tensor imaging in differentiating early-stage subcortical ischemic vascular disease, Alzheimer's disease and normal ageing. *PLoS One*, *12*, e0175143.

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E. J., Bucholz, R., Chang, A., Chen, L., Corbetta, M., Curtiss, S. W., Penna, S. D., Feinberg, D., Glasser, M. F., Harel, N., Heath, A. C., Larson-Prior, L., Marcus, D., Michalareas, G., Moeller, S., ... WU-Minn HCP Consortium. (2012). The human connectome project: A data acquisition perspective. *NeuroImage*, *62*, 2222–2231.

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Resonance in & 2016, undefined. Diffusion MRI noise mapping using random matrix theory. *Magnetic Resonance in Medicine*, *76*, 1582–1593.

Vidal-Pineiro, D., Wang, Y., Krogsrud, S. K., Amlien, I. K., Baaré, W. F., Bartres-Faz, D., Bertram, L., Brandmaier, A. M., Drevon, C. A., Düzel, S., Ebmeier, K., Henson, R. N., Junqué, C., Kievit, R. A., Kühn, S., Leonardsen, E., Lindenberger, U., MadsenK. S., Magnussen, F., Mowinckel, A. M., & Fjell, A. (2021). Individual variations in 'brain age' relate to early-life factors more than to longitudinal brain change. *Elife*, *10*, e69995. https://doi.org/10.7554/eLife.69995

Voineskos, A. N., Farzan, F., Barr, M. S., Lobaugh, N. J., Mulsant, B. H., Chen, R., Fitzgerald, P. B., & Daskalakis, Z. J. (2010). The role of the corpus callosum in transcranial magnetic stimulation induced interhemispheric signal propagation. *Biological Psychiatry*, *68*, 825–831.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*, 600–612.

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., Grydeland, H., Tamnes, C. K., Ostby, Y., & Fjell, A. M. (2010). Life-span changes of the human brain White matter: Diffusion tensor imaging (DTI) and Volumetry. *Cerebral Cortex*, *20*, 2055–2068.

Won, E., Choi, S., Kang, J., Kim, A., Han, K. M., Chang, H. S., Tae, W. S., Son, K. R., Joe, S. H., Lee, M. S., & Ham, B. J. (2016). Association between reduced white matter integrity in the corpus callosum and serotonin transporter gene DNA methylation in medication-naive patients with major depressive disorder. *Translational Psychiatry*, *6*, e866.

Wrigglesworth, J., Ward, P., Harding, I. H., Nilaweera, D., Wu, Z., Woods, R. L., & Ryan, J. (2021). Factors associated with brain ageing–a systematic review. *BMC Neurology*, *211*(21), 1–23.

XBGoost Developers. (2021). *XGBoost documentation - introduction to boosted trees.* https://xgboost.readthedocs.io/en/latest/tutorials/model.html

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods*, *12*, 399–413.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

BIO-PSYCHO-SOCIAL FACTORS' ASSOCIATIONS WITH BRAIN AGE: A LARGE-SCALE UK BIOBANK DIFFUSION STUDY OF 35,749 PARTICIPANTS

M. Korbmacher, T. P. Gurholt, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov

Bio-psycho-social factors' associations with brain age: a large-scale UK Biobank diffusion study of 35,749 participants

Max Korbmacher[1,2,3]*, Tiril P. Gurholt[2], Ann-Marie G. de Lange[2,4,5], Dennis van der Meer[2,6], Dani Beck[2,7,8], Eli Eikefjord[1,3], Arvid Lundervold[1,3,9,10], Ole A. Andreassen[2,11], Lars T. Westlye[2,8,11] and Ivan I. Maximov[1,2]*

[1]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway, [2]Norwegian Centre for Mental Disorder Research (NORMENT), Division of Mental Health and Addiction, Oslo University Hospital, University of Oslo, Oslo, Norway, [3]Mohn Medical Imaging and Visualization Center (MMIV), Bergen, Norway, [4]Department of Psychiatry, University of Oxford, Oxford, United Kingdom, [5]LREN, Centre for Research in Neurosciences, Department of Clinical Neurosciences, Lausanne University Hospital, University of Lausanne, Lausanne, Switzerland, [6]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands, [7]Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, Norway, [8]Department of Psychology, University of Oslo, Oslo, Norway, [9]Department of Radiology, Haukeland University Hospital, Bergen, Norway, [10]Department of Biomedicine, University of Bergen, Bergen, Norway, [11]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

Brain age refers to age predicted by brain features. Brain age has previously been associated with various health and disease outcomes and suggested as a potential biomarker of general health. Few previous studies have systematically assessed brain age variability derived from single and multi-shell diffusion magnetic resonance imaging data. Here, we present multivariate models of brain age derived from various diffusion approaches and how they relate to bio-psycho-social variables within the domains of sociodemographic, cognitive, life-satisfaction, as well as health and lifestyle factors in midlife to old age (*N*=35,749, 44.6−82.8years of age). Bio-psycho-social factors could uniquely explain a small proportion of the brain age variance, in a similar pattern across diffusion approaches: cognitive scores, life satisfaction, health and lifestyle factors adding to the variance explained, but not socio-demographics. Consistent brain age associations across models were found for waist-to-hip ratio, diabetes, hypertension, smoking, matrix puzzles solving, and job and health satisfaction and perception. Furthermore, we found large variability in sex and ethnicity group differences in brain age. Our results show that brain age cannot be sufficiently explained by bio-psycho-social variables alone. However, the observed associations suggest to adjust for sex, ethnicity, cognitive factors, as well as health and lifestyle factors, and to observe bio-psycho-social factor interactions' influence on brain age in future studies.

Paper B

111

1. Introduction

Developmental trajectories of brain morphology are informative signaling markers of health. For example, significant deviations from normative morphology values can signify the presence or development of disease (Marquand et al., 2019; Remiszewski et al., 2022). Based on the idea that a general normative pattern could describe brain trajectories, the concept of brain age has been introduced. Here, different brain features are used to predict individuals' age. The difference between such predicted age and chronological age, the brain age gap (BAG), has the potential as a general health biomarker, sensitive to various neurological, neuropsychiatric, and neurodegenerative disorders (Kaufmann et al., 2019; Cole, 2020; Rokicki et al., 2021). Brain age can be derived using different imaging modalities. Structural and diffusion MRI (dMRI) have shown high prediction accuracy (e.g., Cole, 2020; Beck et al., 2022b; Chen et al., 2022; Leonardsen et al., 2022; Sone et al., 2022). Different dMRI-derived parameters allow one to describe multiple changes in WM micro-structure using various diffusion-weighted approaches. Such dMRI measures provide invaluable information about WM architecture at the micrometer scale and can be associated with macroscopic outcomes. The most popular dMRI approach, diffusion tensor imaging (DTI), is often used to describe WM organization (Basser et al., 1994). However, methodological advances and newer diffusion approaches may provide more meaningful bio-physical information (Novikov et al., 2019), thereby increasing the power of cross-validation of findings and their comparability with other clinical markers (Billiet et al., 2015; Kamagata et al., 2020; Beck et al., 2021; Wood et al., 2022).

The bio-psycho-social model (Engel, 1977) strives for a holistic perspective on medical research to understand health and disease by integrating information on biological, psychological, and social factors (e.g., Ghaemi, 2009; Wade and Halligan, 2017). Brain age can be utilized in this context as an indicator of general health (Kaufmann et al., 2019), using the different levels of observation (bio-psycho-social) to describe brain age relationships with different phenotypes. While there are some studies providing evidence for brain age associations with bio-psycho-social factors, including demographic, biomedical, lifestyle, cognitive, and behavioral factors (Cole, 2020; de Lange et al., 2021; Beck et al., 2022b; Leonardsen et al., 2022; Sone et al., 2022), it remains unclear whether brain age derived from different diffusion approaches relates differentially to sociodemographic, health, life-satisfaction, and cognitive factors (Figure 1), and what the qualities of such relationships are.

While brain age is a proxy for different health-related processes, similar to various bio-psycho-social factors, it remains largely unclear how brain age and bio-psycho-social factors associate. It is hence necessary to increase our understanding and the interpretability of brain age by observing the associations of common phenotypes with brain age. There are large differences in the usage of underlying data and machine learning approaches applied to the data for brain age predictions (Franke and Gaser, 2019). Practical effects of such differences, for example, on phenotype associations, have yet to be systematized to better interpret brain age and BAG. The bio-psycho-social approach (Engel, 1977) lends itself to categorizing phenotypes into concrete groups. Within the groups, phenotype associations with brain age in general can be considered in addition to differences in underlying data used to calculate brain age. Here, we limit our investigations to dMRI-derived brain age to examine brain age

relationships with bio-psycho-social factors specific to WM. WM has repeatedly been shown to change throughout ageing and to relate to different bio-psycho-social variables (Le Bihan and Iima 2015; Beck et al., 2021, 2022a,b). Although comparisons of single MRI modality predictions from either T1-weighted or dMRI depend on the model selection and observed parameter choice (Niu et al., 2019; Rokicki et al., 2021), models using T1-weighted and dMRI features show comparable age prediction performance (Cole, 2020; de Lange et al., 2021). However, phenotype-WM-brain-age relationships require still further examination. Using different diffusion approaches in this context will not only help extend commonly used diffusion tensor imaging by giving reference values to other brain age derived from other WM metrics but also provide a clearer understanding of WM-phenotype associations.

Diffusion MRI can describe various biological processes by providing markers of brain tissue changes across the lifespan (Beck et al., 2021). These markers are not only heritable (Elliott et al., 2018) but also indicative of health, for example, by being associated with psychiatric and neurological disorders, addiction, stroke (Le Bihan and Iima 2015), or cardiovascular health (Beck et al., 2022b). Various diffusion metrics that have previously been related to cognitive and mental health traits have also shared genetic underpinnings with cognitive and mental health phenotypes (Zhao et al., 2021). The biological underpinnings of dMRI markers become particularly apparent in WM abnormalities observed in severe mental disorders, including schizophrenia (Cetin-Karayumak et al., 2020) or bipolar disorder (Houenou et al., 2007). Furthermore, dMRI-derived brain age is higher in people showing accumulations of cardiometabolic risk and markers of adipose tissue distribution (Beck et al., 2022a,b). Such associations between phenotypes and brain age can also be observed when comparing high to low socioeconomic status (SES) groups, where low SES individuals have lower WM integrity (Pavlakis et al., 2015; Shaked et al., 2019).

Furthermore, dMRI offers both single and multi-shell approaches and various meaningful metrics describing white matter microstructure (Jensen et al., 2005; Fieremans et al., 2011; Kaden et al., 2016a,b; Reisert et al., 2017), which serve as a good basis for brain age estimations (Beck et al., 2021; Korbmacher et al., 2022) exploiting biophysically meaningful parameters of brain tissue in contrast to general measures such as grey/white volume or thickness.

While there are various diffusion models offering a plethora of metrics, most efforts have focussed on DTI which provides fractional anisotropy, which decreases, and radial, axial, and mean diffusivity, which increase over the lifetime, respectively, indicating a loss of structural integrity (Westlye et al., 2010; Behler et al., 2021). Advanced diffusion approaches also examine structural integrity, but adding further detail such as the differentiation between intra-and extra-axonal space, parametrization of extra-axonal diffusivity, and axonal bundle distribution (Jensen et al., 2005; Fieremans et al., 2011; Kaden et al., 2016a,b; Reisert et al., 2017).

Differences in brain age-phenotype relationships can be expected when varying dMRI approaches, as varying underlying dMRI approaches will also produce variability in brain age predictions (see Beck et al., 2021; Korbmacher et al., 2022), potentially due to measuring different bio-physical processes (Jensen et al., 2005; Fieremans et al., 2011; Kaden et al., 2016a,b; Reisert et al., 2017). These potential differences become important when attempting to generalize findings on brain age across the literature and setting standards for

brain age predictions. However, to what extent age predictions based on single and multi-shell dMRI approaches relate differentially to phenotypes requires further investigation. Hence, comparing dMRI-based brain age predictions can be fruitful, not only when expanding current efforts of examining brain age associations with phenotypes but also by investigating whether differences in the underlying data can influence relationships of brain age with bio-psycho-social variables.

State-of-the-art conceptualizations of health, such as the bio-psycho-social model (Engel, 1977), recommend considering various domains or levels of explanation when assessing health outcomes, such as brain age. In that sense, brain age can related to different biological, psychological and social factors. The extent of the relationships are important as they can inform on which bio-psycho-social factors lead to better compared to worse brain or general health. Beyond validating brain age as a concept, this can directly improve our understanding of health. To date, brain age is usually calculated from a large range of MRI features. The resulting brain age estimate is then usually predicted from single variables of interest while controlling for sex and age (e.g., Cole, 2020; Leonardsen et al., 2022). However, cumulative and synergy effects can be expected to partly explain health, which has, for example, been shown for cardiometabolic risk factors explaining brain age (see Beck et al., 2022a,b). Hence, we group available phenotypes that have previously been found influential for health (Figure 1) into health and lifestyle factors, representing the biological dimension of the bio-psycho-social model, respectively (Erhardt, 2009; Ning et al., 2020; Gill et al., 2021; Vidal-Pineiro et al., 2021; Beck et al., 2022a,b; Pham et al., 2022). Life satisfaction factors and cognitive factors represent the psychological dimension, and sociodemographic factors the social dimension of the bio-psycho-social model, respectively.

Generally, explaining brain age variance is required to further our understanding of brain age and its multivariate relationship with different phenotypes influencing physiology directly or indirectly. We, therefore, extend previous work by explaining variance in brain age by combining sets of bio-psycho-social variables into domains of sociodemographic, health, life satisfaction, and cognitive factors (Figure 1) to assess their associations with brain age. In addition to exploring associations of bio-psycho-social variables with dMRI-based brain age, we differentiate between diffusion approaches used for brain age predictions and exame the consistency across diffusion approaches. Previous findings revealed weak associations of various phenotypes with brain age in the UK Biobank (e.g., Smith et al., 2019; Cole, 2020). Hence, we expect only small proportions of the variance in brain age to be predicted by bio-psycho-social variables. We also hypothesize that factors directly representing or impacting physiology are more predictive of brain age than those which impact physiology only indirectly. Thus, health factors are presumed to be more predictive of brain age than sociodemographic, cognitive, and life satisfaction factors. Finally, we expect some variability in these associations to be due to the underlying diffusion approach, as different WM properties are also expected to be differentially related to phenotypes. We may move brain age closer to the clinical utility by furthering our understanding of brain age.

2. Methods

2.1. Sample characteristics

The sample used has been described elsewhere (Korbmacher et al., 2022). In brief, the UK Biobank (UKB) (Sudlow et al., 2015) diffusion MRI data consisted of $N = 42,208$ participants. We excluded subjects who withdrew their informed consent (up to 22nd of February 2022) or with an ICD-10 diagnosis from categories F, G, I, or stroke from the general health assessment

FIGURE 1
Overview of the variables used.
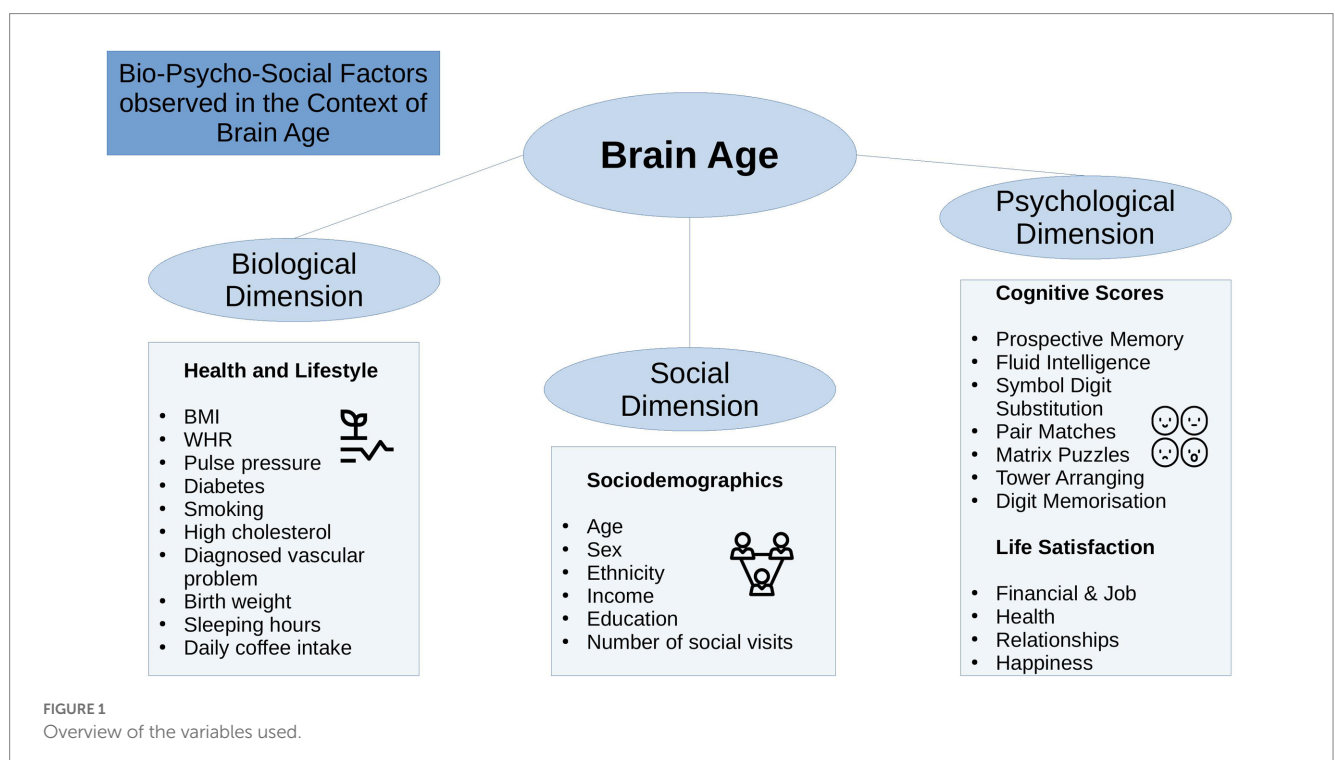
(Field 42,006; excluded: $N = 3,521$). We also excluded data that did not pass our quality control ($N = 2,938$) using the YTTRIUM method (Maximov et al., 2021). In brief, YTTRIUM converts diffusion scalar metric into 2D format using a structural similarity extension (Wang et al., 2004) of each scalar map to their mean image to create a 2D distribution of image and diffusion parameters. Quality check is based on 2 step clustering algorithm in order to identify subjects out of the main distribution. Our final sample consisted of 35,749 healthy adults. For an overview of demographics and the bio-psycho-social variables included in this study and their relationship with brain age see Table 1.

2.2. MRI acquisition, diffusion post-processing, and TBSS analysis

UKB MRI data acquisition procedures are described elsewhere (Sudlow et al., 2015; Miller et al., 2016; Alfaro-Almagro et al., 2018). Briefly, single and multi-shell data were acquired at four different locations using identical scanners: 3 T Siemens Skyra, with a standard 32-channel head coil and key diffusion parameters being MB = 3, $R = 1$, TE/TR = 92/3600 ms, PF 6/8, fat sat, b = 0 s/mm² (5x + 3× phase-encoding reversed), $b = 1,000$ s/mm² (50×), b = 2,000 s/mm² (50×) (Alfaro-Almagro et al., 2018).

We obtained access to the raw diffusion data and pre-processed the data using an optimized pipeline as described by Maximov et al. (2019). The pipeline includes corrections for noise (Veraart et al., 2016), Gibbs ringing (Kellner et al., 2016), susceptibility-induced and motion distortions, and eddy current artifacts (Andersson and Sotiropoulos, 2016). Isotropic 1 mm³ Gaussian smoothing was carried out using FSL's (Smith et al., 2004; Jenkinson et al., 2012) *fslmaths*. Employing the multi-shell data, Diffusion Tensor Imaging (DTI), Diffusion Kurtosis Imaging (DKI) (Jensen et al., 2005) and White Matter Tract Integrity (WMTI) (Fieremans et al., 2011) metrics were estimated using Matlab 2017b code.[1] Spherical mean technique SMT (Kaden et al., 2016b), and multi-compartment spherical mean technique (mcSMT) (Kaden et al., 2016a) metrics were estimated using original code[2] (Kaden et al., 2016a,b). Estimates from the Bayesian Rotational Invariant Approach (BRIA) were evaluated by the original Matlab code[3] (Reisert et al., 2017).

Previous advances observing age-dependent WM changes have largely focused on single-shell diffusion, such as DTI with DTI-derived metrics being fractional anisotropy (FA), and axial (AD), mean (MD), and radial (RD) diffusivity, all being highly sensitive to age (Westlye et al., 2010; Cox et al., 2016; Beck et al., 2021). More recently developed multi-shell diffusion approaches which extend the space of derivable diffusion metrics appear more sensitive to brain changes and sex differences (Lawrence et al., 2021), and at the same time less sensitive to motion artefacts than single-shell models (Pines et al., 2020). Newer approaches are (1) BRIA, as an alternative to not rely on fiber orientation but rotation invariant feature (Reisert et al., 2017), (2) DKI, a method tackling the problem of non-Gaussian diffusion (Jensen et al., 2005); (3) WMTI, which extends DKI by calculating

---

1 https://github.com/NYU-DiffusionMRI/DESIGNER
2 https://github.com/ekaden/smt
3 https://bitbucket.org/reisert/baydiff/src/master/

TABLE 1 Overview of the predictors used in the bio-psycho-social models.

Model	Predictors
Model 1. Baseline	Age
	Sex
	Scanner site*
Model 2. Socio-demographics	Age
	Sex
	Scanner site*
	Ethnicity
	Income
	Education
	Nb of social visits
Model 3. Cognitive Scores	Age
	Sex
	Scanner site*
	Prospective memory
	Fluid intelligence
	Symbol digit substitution
	Pair matches
	Matrix puzzles
	Tower arranging
	Digit memorization
Model 4. Life satisfaction	Age
	Sex
	Scanner site*
	Financial and job satisfaction
	Friend and family relation satisfaction
	Happiness
Model 5. Health and lifestyle	Age
	Sex
	Scanner site*
	BMI
	WHR
	Pulse pressure
	Diabetes
	Smoking
	High cholesterol
	Diagnosed vascular disorder
	Birth weight
	Sleeping hours
	Daily coffee intake
	Alcohol drinker

*Random effect.

inter and extra-axonal features (Fieremans et al., 2011); and (4) SMT (Kaden et al., 2016b) and (5) mcSMT, which factor out neurite orientation to give a better estimate of microscopic diffusion anisotropy (Kaden et al., 2016a). The selection of diffusion models was dictated by a few practical reasons. There are two conventional approaches (DTI and DKI) describing the general WM changes. As a result, these approaches are expected to be sensitive to a broad range of aging-related effects associated with WM maturation (Westlye et al., 2010; Yap et al., 2013). Advanced dMRI approaches enable more detailed quantification associated with age in a different manner (Cox et al., 2016; Beck et al., 2021). Diffusion modelling relies on

114   Paper B

biophysically motivated assumptions such as the axon bundle distribution (WMTI) or attempts to suppress such kind of parameters (SMT and SMT mc). Another modelling option are Bayesian rotation invariants (BRIA), providing multiple measures of WM but depending on efficacy of initial Bayesian simulations. All together, these approaches allow us to indirectly verify the stability and reliability of diffusion assumptions in brain-age prediction on their own and in comparison to each other, or to determine similarity among scalar metrics appearing in several diffusion approaches.

In total, we obtained 28 metrics (Supplementary Table S1) from six diffusion modeling approaches (DTI, DKI, WMTI, SMT, mcSMT, and BRIA). To normalize all metrics, we used tract-based spatial statistics (TBSS) (Smith et al., 2006) as part of FSL (Smith et al., 2004; Jenkinson et al., 2012). In brief, initially, all FSL BET-extracted (Smith, 2002) FA images were aligned to MNI space using non-linear transformation (FNIRT) (Jenkinson et al., 2012). Subsequently, we derived the mean FA image and the related mean FA skeleton. Each diffusion scalar map was projected onto the mean FA skeleton using the standard TBSS procedure. To provide a quantitative description of diffusion metrics we evaluated averaged values over the skeleton and two WM atlases, namely the Johns Hopkins University (JHU) atlas (Mori et al., 2005) and the JHU tractography atlas (Hua et al., 2008; see Supplementary Table S2 for an overview). Finally, we obtained 20 WM tracts and 48 regions of interest (ROIs) based on a probabilistic WM atlas (JHU) (Hua et al., 2008) for each of the 28 metrics, including the mean skeleton values. Altogether, we derived 1,932 features per individual [28 metrics * (48 ROIs +1 skeleton mean + 20 tracts)]; see Supplementary Table S1 for metrics and Supplementary Table S2 for regions and tracts.

2.3. Brain age predictions

We computed brain age predictions derived from 8 different models including the six diffusion approaches, their whole-brain average scores (mean multimodal), and a model combining the six diffusion approaches and their whole-Brian average scores (full multimodal). Each of the six diffusion approaches details WM features based on differing modelling assumptions and were assumed to provide unique brain age scores. Whole-brain average scores for each of the six diffusion approaches' metrics were investigated on their own to further our understanding of spacial specificity. Finally, previous results (de Lange et al., 2020b; Beck et al., 2021, 2022b) provide clear evidence of strong age prediction performance when combining diffusion metrics. We hence included a model combining all diffusion approaches' metrics and their whole-brain average scores to compare whether there are differences in multimodal to single diffusion approaches' brain-age-phenotype associations.

Brain age was predicted using the XGBoost tree-boosting algorithm (gradient boosting tree) implemented in Python (v3.7.1), being a highly effective algorithm for tabular data (Chen and Guestrin, 2016). From the total included sample ($N = 35,749$), we used 10% (N = 3,575) for hyperparameter tuning on a data set containing data from all diffusion approaches (i.e., full multimodal data with 1,932 features/parameters) using 5-fold cross-validation (after estimating an optimal hyperparameter tuning set size; Korbmacher et al., 2022). The considered hyperparameters for the randomized grid search were (1) learning rate with a range of 0.01–0.3 and steps of 0.05, (2) maximum

layers/depth with a range of 3–6 and steps of 1, and (3) number of trees with a range of 100–1,000 and steps of 50. The resulting hyperparameters (learning rate = 0.05, max layers/depth = 3, and the number of trees = 750) were then used in a 10-fold cross-validation applied to the test set ($N = 32,174$). Cross-validation was used to leverage the full sample size and to calculate the uncertainty around the estimates (for such see Korbmacher et al., 2022). The cross-validation procedure was executed using each of the six diffusion approaches' metrics, whole-brain averaged metrics for all approaches (mean multimodal model), and finally a combination of all approaches and the whole-brain average scores (full multimodal model), resulting in eight brain age models (see Supplementary Table S1 for dMRI approach-specific metrics). Each of these brain ages were used in the analyses. See Supplementary Figure S1 for an overview of the brain age models and the following modelling of these predictions from the bio-psycho-social models.

2.4. Statistical analyses

All statistical analyses were carried out using Python, version 3.7.1 and R, version 4.2.0[4] using test data set ($N = 32,174$). These analyses focused on the associations between brain age and (1) demographics, (2) social factors, (3) cognitive test scores, (4) life satisfaction, and (5) health and lifestyle factors (with weight on cardiometabolic factors). For detailed information on how variables were extracted and coded see Supplementary Table S3. First, we calculated the first principal component of bio-psycho-social factors' by grouping numeric variables of each of the 5 domains (demographics, social factors, cognitive tests scores, life satisfaction, and health and lifestyle factors), using scaling and the number of allowed components equal to the number of variables included. We then examined the first components' associations with brain age. Second, we examined to which extend multivariate models (as specified in 2.4.1) explain brain age from the factors of the five bio-psycho-social domains. Finally, we tested whether our findings would be influenced by analyzing data separately for males and females, and present bi-variate relationships between multimodal brain age and single bio-psycho-social variables.

For bi-variate relationships between bio-psycho-social factors and full multimodal brain age, we adjusted p-values for multiple comparison using Bonferroni correction, dividing the alpha-level by twenty-five ($\alpha/25$), the number of bi-variate associations observed. For multivariate relationships we divided alpha by eight ($\alpha/8$), the number of brain age models used. Furthermore, the coefficient of determination describing the proportion of variance explained (R^2) will be presented as marginal R^2, referring to variance explained by fixed effects, and conditional R^2, referring to both fixed and random effects variance explained.

2.4.1. Bio-psycho-social models explaining brain age

We used linear mixed effects models with the random intercepts at the level of scanner site to explain changes in brain age from socio-demographics, cognitive test scores, life satisfaction (self-assessment), and

---

4 www.r-project.org/

Paper B   frontiersin.org

115

health and lifestyle factors. The presented models were used in two different ways: first, with the principal component of the model-specific bio-psycho-social factors replacing the respective bio-psycho-social factors, and second using all eight brain ages from the different diffusion approaches on with the models. For an overview of the predictors in the multivariate bio-psycho-social models used see Table 1.

We established the following models to compare:

(1) A baseline model capturing the relationship of age, sex, the age-sex interaction, and scanner site with brain age. This baseline model was selected as predicted age is expected to be largely reflected by chronological age. However, also sex (e.g., Rokicki et al., 2021), and scanner site (here, Bristol, Cheadle, Newcastle, Reading) and prediction bias (e.g., Jirsaraie et al., 2022) have been shown to be influential for brain age. Using a baseline model and additional models for comparison had the goal to estimate added variance explained by the bio-psycho-social models above and beyond the baseline mode (Bollen, 1989). Additionally, predictors within these bio-psycho-social models were observed individually (bivariate compared to multivariate relationships with brain age). Model comparison to a baseline model (instead of a null model) is important in this context as brain age is sensitive to age, sex and scanner site (de Lange and Cole, 2020; Rokicki et al., 2021; Jirsaraie et al., 2022). Hence, instead of using a null model which does not contain much information, we used the following model as a reference point for further model comparison:

$$brainage = sex + age + sex * age + site$$

(2) A sociodemographic model additionally included ethnic ancestry (binary yes/no self-reported white European; for additional information sample groupings by ethnicity see Supplementary Table S4), average annual total household income before tax (coded as continuous variable 1–5, with low <£18,000 to high income >£100,000), and higher education (binary yes/no self-report of having obtained higher education) relative to the baseline model.

$$brainage = sex + age + sex * age + ethnicity \\ + income + education + site$$

(3) A cognitive model testing how non-verbal cognitive abilities add to the baseline model (overview: Fawns-Ritchie and Deary, 2020). We limited the selection of cognitive variables to non-verbal assessment measures to reduce the parameter space of cognitive variables and as non-verbal assessment scores have been found to associated with dMRI metrics throughout the lifespan (e.g., Sullivan and Pfefferbaum, 2006; Sasson et al., 2010; McPhee et al., 2019; Parikh et al., 2021). Namely, the number of matrix puzzles solved (matrixS) testing non-verbal reasoning using COGNITO Matrices, tower arranging correctly solved (towerS) testing executive function using the Delis-Kaplan Executive Function System Tower Test, prospective memory (memory) assessed with the Rivermead Behavioural Memory Test, fluid intelligence (intel) from the UKB own Fluid IQ test, digits remembered (digits) from the Symbol Digit Modalities Tests, and the mean number of incorrect pair matches (IPM) across trail A and B assessing visual declarative memory using the Wechsler Memory Scale IV Designs I and Designs II. Correlations were small to moderate ($r_{max} = 0.41$) with the variance inflation factor (VIF) indicating low levels of multicollinearity (Supplementary Figure S2).

$$brainage = sex + age + sex * age + matrixS + towerS \\ + memory + intel + digits + IPM + site$$

(4) A life satisfaction model that additionally included job satisfaction (jobS), financial satisfaction (financeS), overall health rating (healthR), health satisfaction (healthS), family relation satisfaction (famS), happiness, friend relationship satisfaction (friendS) relative to the baseline model. Some of the model features were highly correlated ($r_{max} = 0.65$), yet VIF values indicated low levels of multicollinearity (Supplementary Figure S3).

$$brainage = sex + age + sex * age + jobS \\ + financeS + healthR + healthS \\ + famS + friendS + happiness + site$$

(5) A health and lifestyle model testing how body mass index (BMI), pulse pressure (Ppressure: the difference between systolic and diastolic blood pressure), waist-to-hip-ratio (WHR), binary smoking status, binary diabetes diagnosis (both type I and II), binary high cholesterol (chol), binary diagnosed vascular problem (DVP), birth weight (Bweight), sleeping hours, and daily coffee intake (coffee) add to the baseline model, with only BMI and WHR showing a moderate correlation $r = 0.42$, but all other correlations being small $rs < 0.16$, with VIF values indicating only low levels of multicollinearity (Supplementary Figure S4).

$$brainage = sex + age + sex * age + BMI + WHR + Ppressure \\ + diabetes + smoking + chol + DVP \\ + Bweight + coffee + site$$

2.4.2. Follow-up and quality control analyses and single bio-psycho-social factor associations with multimodal brain age

Previous research showed sex differences in brain age, suggesting sex separate analyses (Rokicki et al., 2021). Hence, we conducted the analyses described in 2.4.1 separately for males and females.

$$brainage = age + site + biopsychosocialfactors$$

To examine the contributions of single bio-psycho-social variables to explaining WM brain age, linear mixed models were used to observe bio-psycho-social variable associations with brain age when controlling for age and sex with scanner site as a random factor. In other words, different from 2.4.1, we applied one model per bio-psycho-social factor. For simplicity, this analysis step only considered the best brain age predictions from the multimodal model including the metrics of all diffusion approaches (Korbmacher et al., 2022).

$$brainage = sex + age + sex * age + site \\ + singlebiopsychosocialfactor$$

Each model was then compared to a model not including the respective bio psycho social variable:

$$brainage = sex + age + sex * age + site$$

3. Results

3.1. Linear mixed effect models explaining brain age gap from bio-psycho-social factors

We ran the proposed five baseline and bio-psycho-social models with the first principal component (PC) of the numeric predictors from each of the models showing a small proportion of the variance in brain age uniquely explained by the principal components ($R^2 < 1\%$; Supplementary Table S5), with differences between these models and respective baseline models yet being highly significant (Supplementary Table S6).

When including bio-psycho-social factors instead of their PCs and comparing baseline to models 2–5, a larger proportion of both marginal or conditional variance in brain age could be uniquely explained by bio-psycho-social variables (marginal and conditional $R^2 < 0.03$; Figure 2 and Supplementary Table S7). Model comparisons showed that, with the exception of socio-demographic factors, bio-psycho-social models explained significantly more variance in brain age than the baseline model (with age, sex, and age-by-sex interaction as fixed and scanner site as random effect), irrespective of the diffusion approach used to calculate brain age ($ps < 0.01$; Figure 3). Differences between this uniquely explained marginal variance were small across diffusion approaches (Figure 2).

Across statistical and diffusion models, age was used as a control variable to correct for the mere reflection of age by brain age producing stable associations across models (1–5) for multimodal brain age (Figures 4–7). However, except for the life-satisfaction model, in contrast to the full multimodal model, the other diffusion approaches'

brain ages were negatively associated with age, giving another indication of overall poor model fit. Even more so, the effect of sex was dependent on the model, producing mixed effects with large uncertainty surrounding β-values, also in the age-by-sex interactions' associations with brain age. Overall, bio-psycho-social factors were consistently associated with brain ages from different diffusion approaches, with the exception for sex (Figures 4–7).

3.1.1. Sociodemographic factors' associations with brain age

In the model including sociodemographic factors explaining brain age (see Figure 4 for the predictors), results were mixed for the significant predictors. Sex was a significant predictor for mean DKI, DTI, and WMTI ($ps < 0.05$), the age-by-sex interaction only for BRIA ($p = 0.045$), and ethnicity only for DKI ($p = 0.012$; Figure 4). Overall, only 95% confidence intervals of β-values for age and ethnicity were not consistently overlapping, indicating differential effects of these variables on brain age based on the underlying data. All other 95% confidence intervals surrounding coefficients' β-values were overlapping across diffusion approaches, with expected strong age contributions predicting brain age.

3.1.2. Health and lifestyle factors' associations with brain age

Similarly, in the model including health and lifestyle factors explaining brain age (see Figure 5 for the predictors), significant health factors leading to higher brain age were WHR ($ps < 0.001$), pulse pressure ($ps < 0.001$), and hypertension ($ps < 0.001$). Evidence across diffusion approaches was mixed for the other predictors with smoking predicting brain age derived from BRIA, DTI, mean scores, SNT, and WMTI ($ps < 0.05$), diabetes diagnosis for all models except DKI and DTI ($ps < 0.05$), the diagnosis of at least one vascular disease for BRIA, mean scores, and mcSMT ($ps < 0.02$), and average daily cups of coffee for brain age estimates except the one based on BRIA
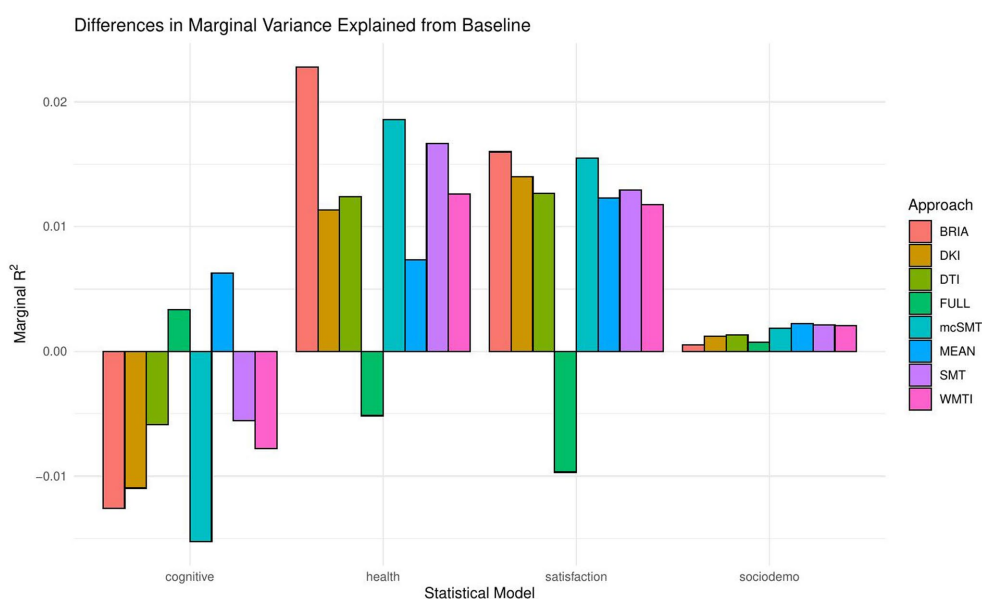
FIGURE 2
Marginal R^2 values for statistical models across diffusion approaches.

FIGURE 3
Overview of comparison of bio-psycho-social statistical models with baseline models. The figure presents χ^2 values for each of the bio-psycho-social statistical models for each diffusion approach tested against the baseline model. Note that only values of $\chi^2 > 11$ were significant ($p < 0.05$).
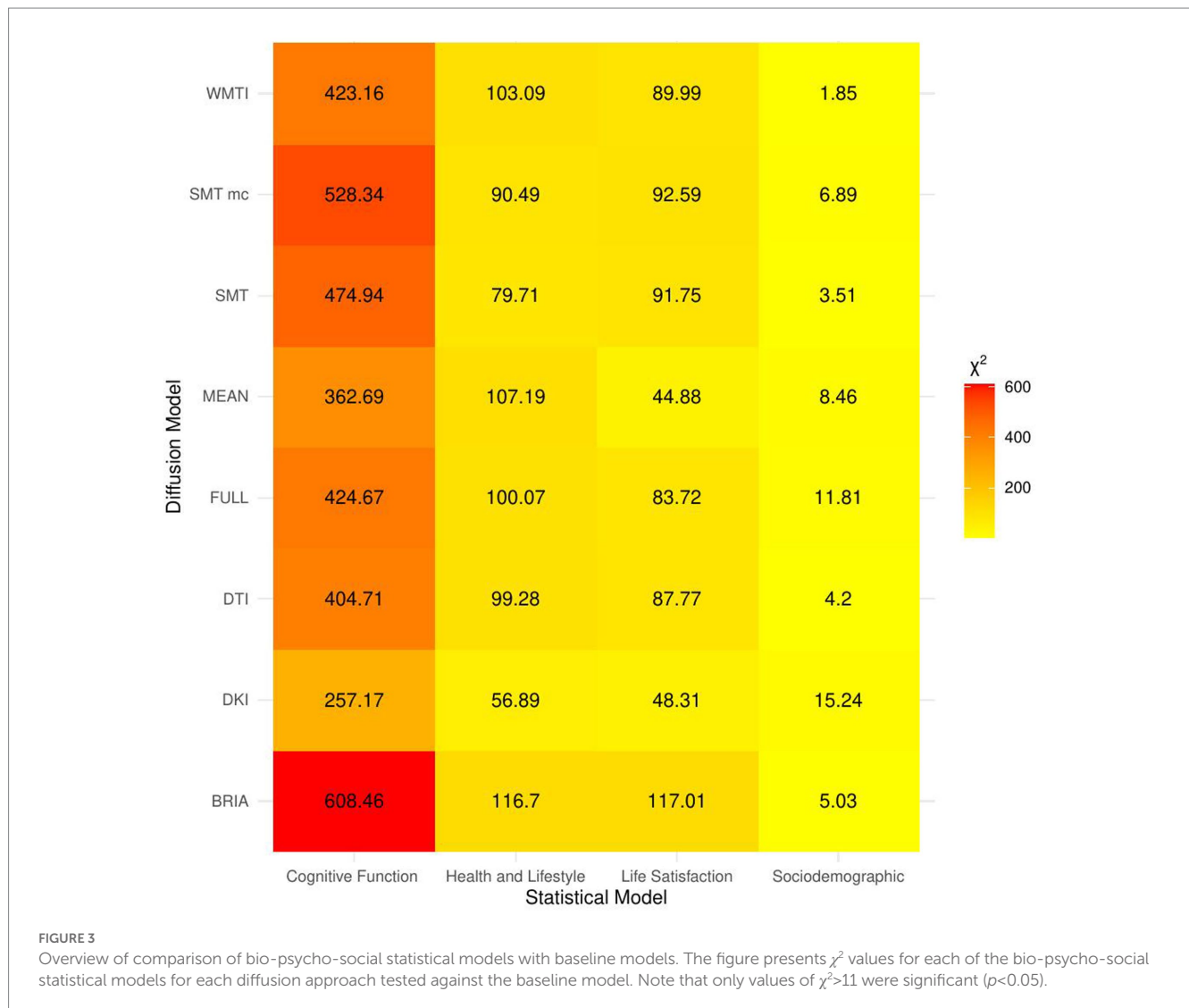
($ps < 0.01$), and the age-by-sex interaction for BRIA and mcSMT ($ps < 0.04$).

Interestingly, WHR was a stronger predictor of brain age in males than in females (Supplementary Figure S5). Practically, a WHR β_{unstd}-value of, for example $\beta = 4$ would mean that for every 0.1 step change in WHR, the brain age can be expected to increase by 0.4 years (see Supplementary Figures S6–S9 for β_{unstd}). Importantly, this association was controlled for age, as age is correlated with WHR at $r = 0.14$ and brain age at $r = 0.80$. Mean population values for WHR were found to be WHR < 1 (Molarius et al., 1999), with our sample corresponding with these estimates ($M_{WHR} = 0.871 \pm 0.088$, min = 0.534, max = 1.472) with males having a higher WHR ($M_{WHR} = 0.923 \pm 0.064$) than females ($M_{WHR} = 0.817 \pm 0.069$).

BMI was potentially non-significant due to the model construction as the highly correlated WHR (Supplementary Figure S4) was a significant predictor of brain age, and BMI alone being a significant predictor of brain age (Table 2). Finally, higher birth weight was associated with lower brain age estimated from full and mean models, as well as BRIA and WMTI ($ps < 0.02$).

Generally, 95% confidence intervals around coefficients' β-values were overlapping across models indicating no significant differences in

β-values across diffusion approaches. As a control, we ran the same model without WHR as predictor, due to its high correlation with BMI, rendering BMI as significant predictor across diffusion approaches' brain ages except the mean model ($\beta s > 0.01$, $ps < 0.004$), also showing now clearer evidence for higher brain age when smoking ($ps < 0.05$), with other predictors unchanged (Supplementary Figure S11). Furthermore, leaving out hypertension, being a substrate of blood pressure, did not lead to changes in the model (Supplementary Figure S12). For both models, variance explained is slightly reduced compared to the models including the respective variables, making the reduced models significantly different ($ps < 0.001$) from the full health models (Supplementary Tables S8, S9).

3.1.3. Life satisfaction factors' associations with brain age

When modeling brain age from life satisfaction (see Figure 6 for the predictors), self-rated health was a significant predictor of all brain age estimates except for DKI brain age ($ps < 0.05$) and health satisfaction for all brain age estimates except the mean model's brain age ($ps < 0.02$). Only the 95% confidence intervals of β-values for age do not overlap across models (with the mean model having the largest β and full model the smallest β-value for age). All other 95%

FIGURE 4
Sociodemographic model predictors' standardized beta-values with standard error. *Indicates Bonferroni-corrected *p*<0.05.

FIGURE 5
Health model predictors' standardized beta-values with standard error. *Indicates Bonferroni-corrected *p*<0.05.
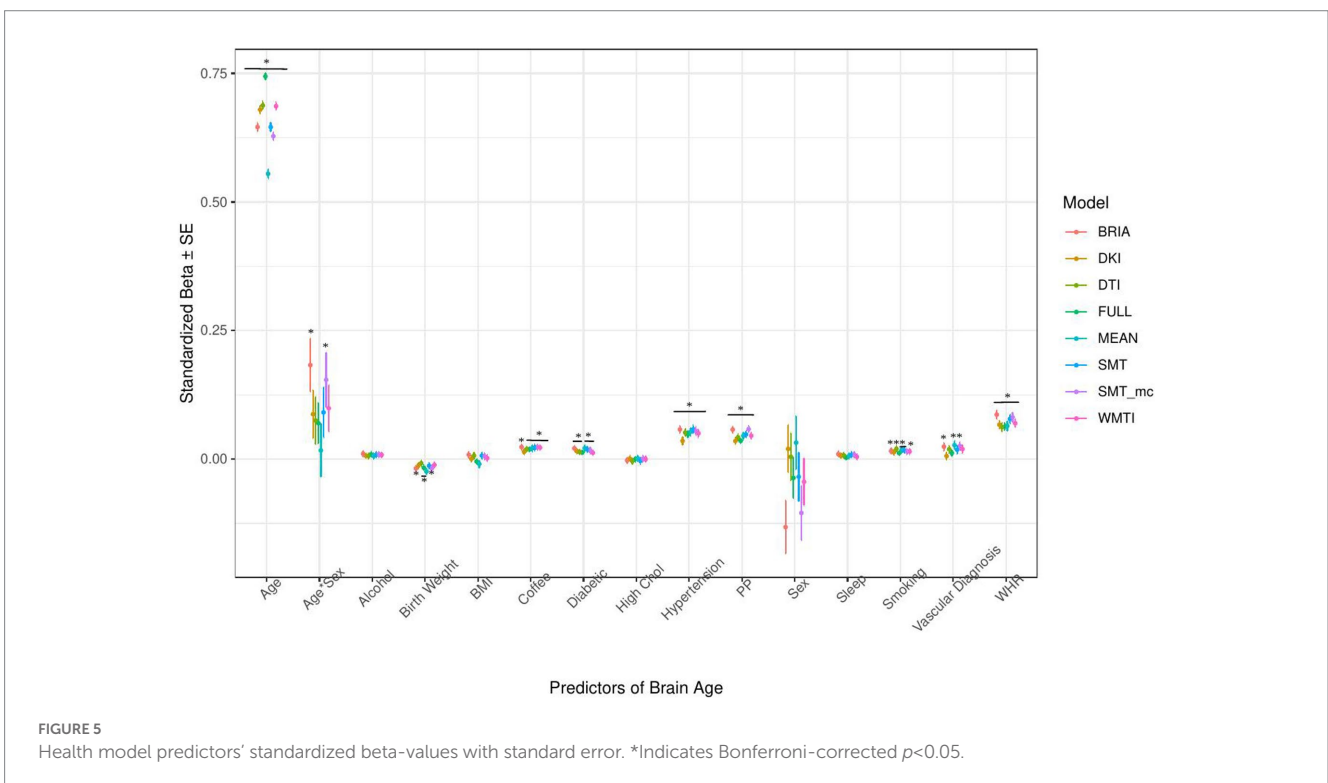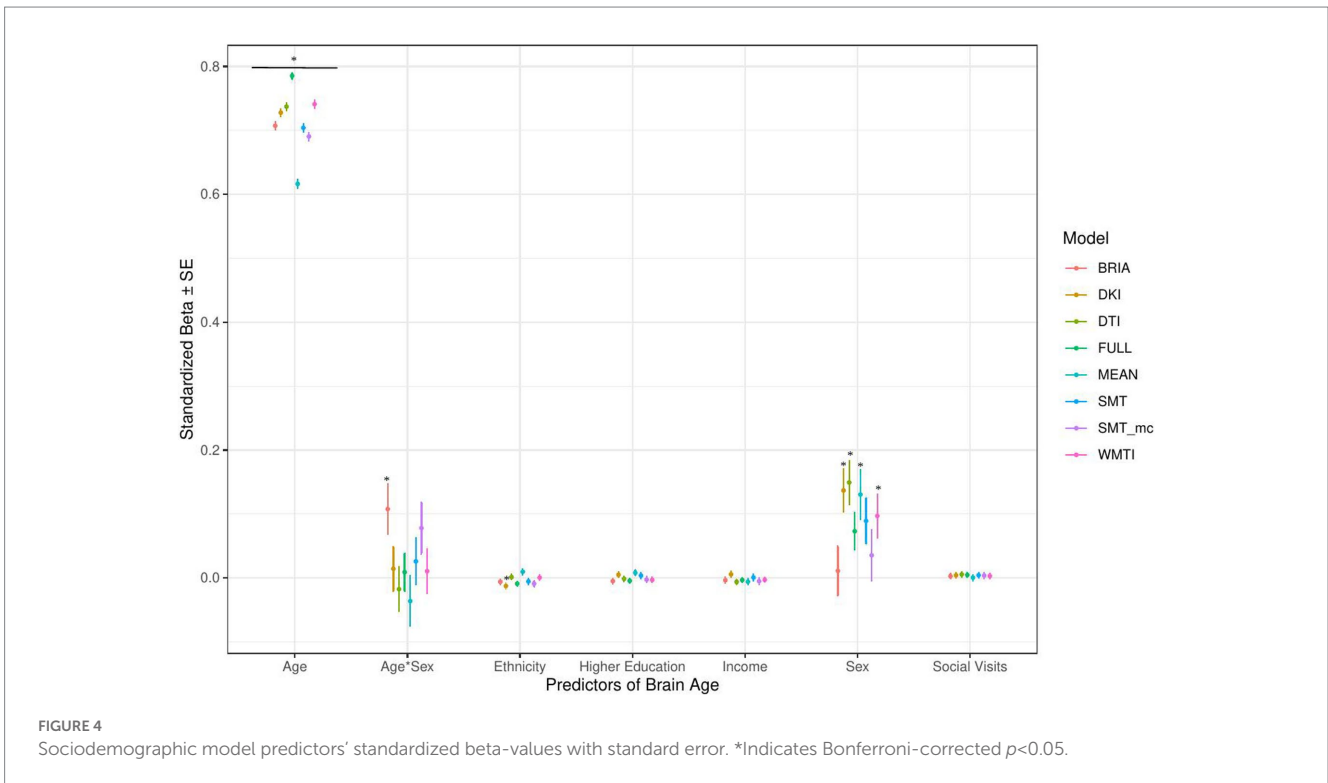
confidence intervals around coefficients' β-values overlap across models indicating no significant differences in β-values across diffusion approaches.

Perceived health is moderately correlated with health satisfaction and was left out in a control model resulting in a slightly stronger effect of

health satisfaction and significantly worse performing model ($ps<0.001$; Supplementary Figure S12 and Supplementary Tables S8, S9). Differently, when leaving out happiness as being correlated with several variables the model remains unaffected ($ps>0.23$; Supplementary Figure S13 and Supplementary Tables S8, S9).
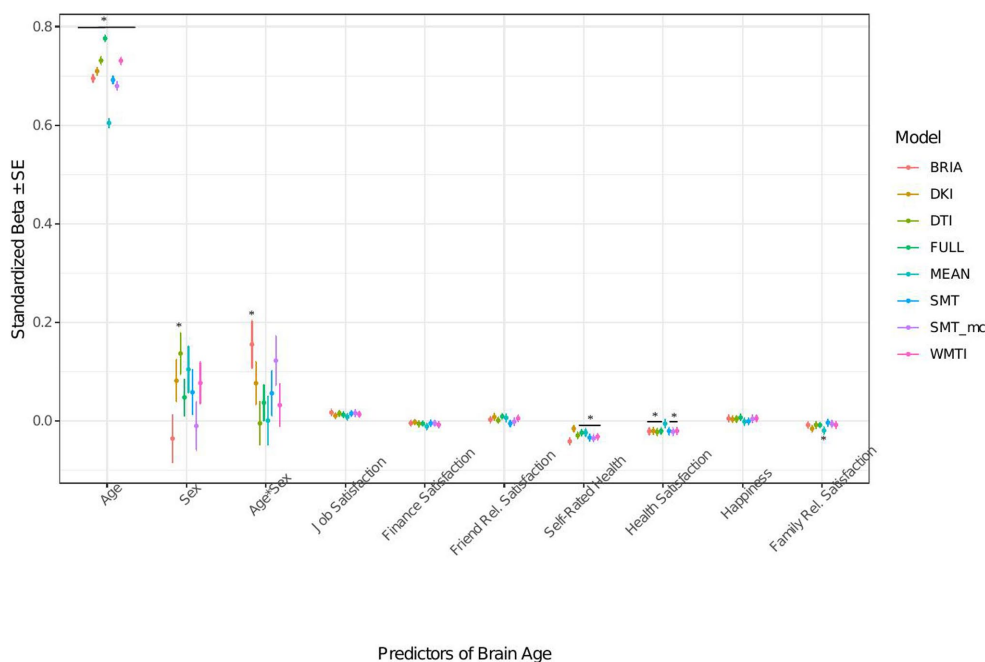
FIGURE 6
Well-being model predictors' standardized beta-values with standard error. *Indicates Bonferroni-corrected $p<0.05$.
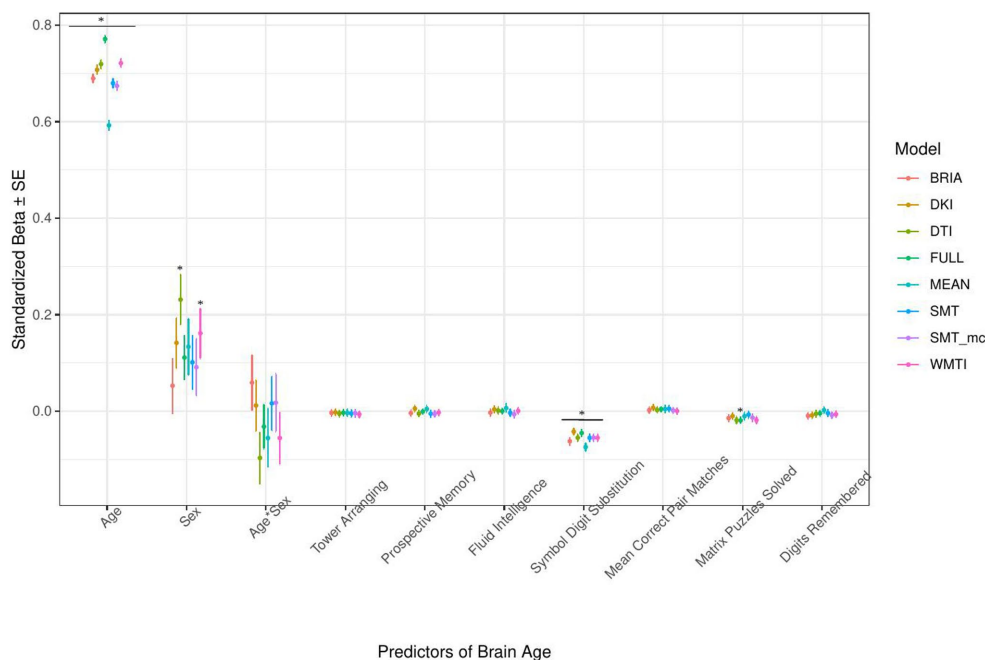
FIGURE 7
Cognition model predictors' standardized beta-values with standard error. *Indicates Bonferroni-corrected $p<0.05$.

3.1.4. Cognitive factors' associations with brain age

The only cognitive factor explaining brain age across all models was symbol digit substitution ($ps<0.001$; Figure 7). Matrix puzzles solved was only a significant predictor for the full multimodal brain age ($p=0.014$), and sex only for DTI and WMTI ($ps<0.02$).

Confidence intervals around coefficients' β-values are overlapping across models indicating no significant differences in β-values across diffusion approaches. Fluid intelligence and matrix puzzles are highly correlated and hence, matrix puzzles were left out in a quality control model, not significantly affecting the structure of most models (Supplementary Figure S14 and Supplementary Tables S8, S9).

TABLE 2 Linear models relating multimodal brain age to bio-psycho-social factors.

Variable	Level or metric	Variable value	Brain age[1]/ Age[1]	N[2]	Marginal R² diff	Log Likelihood_diff	χ² p_diff[4]	β_raw/β_std[3]	p_pred[4]
Brain age	Mean ± SD	64.470 ± 5.946		32,174					
Demographics									
Scanner site	% Cheadle	57.559	63.6/63.6	18,519					
	% Newcastle	26.403	65.3/65.2	8,495					
	% Reading	15.913	66.1/66.2	5,120					
	% Bristol	0.124	67.0/67.2	40					
Sex	% Male	47.122	65.4/65.1	17,013				−1.07/0.09	0.025
	% Female	52.878	63.7/63.9	15,161					
Age	Mean ± SD	64.473 ± 7.614	64.5/64.5	32,174				0.62/0.79	<0.001
Socio-demographics									
Ethnicity	European	96.800	64.5/64.6	31,160	-1.8×10^{-5}				
	Non-European	2.970	62.3/61.2	956	-1.8×10^{-5}	4	0.007	−0.31/−0.01	185
	Prefer not to say	(0.180)	65.6/65.6	58					
Income[5]	% less £18 k	10.363	65.6/66.1	3,310	1.4×10^{-4}				
	% £18 k-£30	24.253	65.6/66.5	7,747	1.4×10^{-4}	11	<0.001	−0.28/−0.02	0.003
	% £30 k-£52 k	27.713	64.6/64.7	8,852	1.4×10^{-4}	11	<0.001	−0.29/−0.02	<0.001
	% £52 k-100 k	21.586	62.9/61.7	6,895	1.4×10^{-4}	11	<0.001	−0.18/−0.01	0.4
	% > £100 k	6.956	61.5/59.4	2,222	1.4×10^{-4}	11	<0.001	−0.30/−0.01	0.075
	Do not know	3.040	67.2/68.0	972					
	Prefer not to say	6.086	65.6/66.8	1944					
Higher education	% Yes	49.326	64.2/63.9	15,870	5.2×10^{-5}	1	0.177	5.3×10^{-2}/−0.004	1
	% No	50.674	64.8/65.0	16,304	5.2×10^{-5}	1	0.177		
Cognitive test scores									
Matrix puzzles solved	Mean ± SD	8.011 ± 0.500	64.8/64.7	21,755	6.8×10^{-4}	27	<0.001	−0.08/−0.03	<0.001
Tower rearranging correct attempts	Mean ± SD	9.920 ± 2.123	64.8/64.7	21,587	5.0×10^{-4}	17	<0.001	−0.04/−0.02	<0.001
Prospective memory	Mean ± SD	1.067 ± 0.397	64.3/64.3	30,300	-1.7×10^{-6}	0	0.312	−0.05/−0.003	1
Fluid intelligence	Mean ± SD	6.631 ± 0.397	64.3/64.2	29,786	7.3×10^{-4}	23	<0.001	−0.07/−0.02	<0.001
Digits remembered	Mean ± SD	6.675 ± 1.540	64.9/64.8	23,070	1.6×10^{-4}	15	<0.001	−0.06/−0.02	0.002
Mean number of incorrect pair matches across trials	Mean ± SD	2.214 ± 1.279	63.9/63.8	20,770	2.9×10^{-4}	3	0.014	−0.05/0.01	0.375
Life satisfaction[6]									
Job satisfaction	Mean ± SD	4.511 ± 0.863	62.7/61.6	18,399	-5.5×10^{-6}	1	0.494	0.02/0.003	1
Financial satisfaction	Mean ± SD	4.714 ± 0.828	64.4/64.5	31,909	2.6×10^{-4}	10	<0.001	−0.10/0.003	<0.001
Health satisfaction	Mean ± SD	4.470 ± 0.766	64.5/64.5	31,911	0.001	52	<0.001	−0.26/0.003	<0.001
Overall health rating	Mean ± SD	3.030 ± 0.630	64.5/64.5	31,934	0.001	66	<0.001	−0.36/−0.04	<0.001
Family relation satisfaction	Mean ± SD	4.814 ± 0.846	64.4/64.4	31,737	2.7×10^{-4}	17	<0.001	−0.10/0.003	<0.001
Friend relationship satisfaction	Mean ± SD	4.784 ± 0.846	64.4/64.4	31,640	2.0×10^{-6}	0	0.485	−0.02/0.003	1
Happiness	Mean ± SD	4.542 ± 0.686	64.5/64.5	31,884	1.2×10^{-4}	3	0.011	−0.07/−0.008	0.275

(Continued)

Paper B 121

TABLE 2 (Continued)

Variable	Level or metric	Variable value	Brain age[1]/ Age[1]	N^2	Marginal R^2 diff	Log Likelihood$_{diff}$	$\chi^2\,p_{diff}$[4]	β_{raw}/β_{std}[3]	p_{pred}[4]
Health and lifestyle factors									
BMI	Mean ± SD	26.319 ± 4.269	64.4/64.4	31,052	5.9×10^{-4}	41	<0.001	0.04/0.03	<0.001
Pulse pressure	Mean ± SD	60.027 ± 14.540	64.3/64.3	28,184	0.002	66	<0.001	0.02/0.05	<0.001
WHR	Mean ± SD	0.872 ± 0.088	64.4/64.4	31,138	0.004	129	<0.001	4.86/0.07	<0.001
Smoking	% Yes	2.629	63.0/61.4	838	7.5×10^{-5}	4	0.005	0.34/0.01	0.15
	% No	97.374	64.5/64.5	31,033	7.5×10^{-5}	4	0.005		
Diabetes	% Yes	1.688	66.6/66.1	543	5.1×10^{-4}	22	<0.001	0.99/0.02	<0.001
	% No	98.312	64.4/64.4	31,631	5.1×10^{-4}	22	<0.001		
Hypertension	% Yes	19.680	66.7/66.9	6,332	0.002	151	<0.001	0.86/0.06	<0.001
	% No	80.320	63.9/63.9	25,842	0.002	151	<0.001		
High cholesterol	% Yes	12.202	66.9/68.0	3,926	-7.4×10^{-5}	18	<0.001	0.26/0.01	<0.001
	% No	83.798	64.1/64	28,248	-7.4×10^{-5}	18	<0.001		
Vascular diagnosis	% Yes	22.726	66.7/67.1	7,312	0.002	136	<0.001	0.78/0.06	<0.001
	% No	87.274	63.8/63.7	24,862	0.002	136	<0.001		
Birth weight (kg)	Mean ± SD	3.358 ± 0.619	63.7/63.3	19,409	3.0×10^{-4}	10	<0.001	−0.18/−0.02	<0.001
Daily coffee intake (cups)	Mean ± SD	2.065 ± 1.815	64.5/64.5	31,973	3.0×10^{-4}	20	<0.001	0.06/0.02	<0.001

Linear models were used to observe each of the bio-psycho-social variables of interest in a model of [brain age ~ age + sex + age*sex + bio-psycho-social variable] with scanner site as a random factor. The model was then compared to a baseline model of [brain age ~ age + sex + age*sex] with scanner site as a random factor, and hence, R^2 values refer to variance uniquely contributed by the single bio-psycho-social variable when added to the baseline model, referring to the marginal R^2. For sex and age, fixed effects, and site, the random effect are being presented from the baseline model. [1]These values differ because of missing data in the associated bio-psycho-social factors and will be influenced by the model's age bias. [2]Sample sizes differ due to missing data in the bio-psycho-social variables. [3]β_{raw}/β_{std} are the raw and standardized β values of the bio-psycho-social variables of interest. [4]p_{pred} refers to the Bonferroni-corrected (multiplied by the number of bi-variate tests = 25) p-values of the predictors/bio-psycho-social variables' β, p_{diff} refers to the differences between baseline model and the respective model including the bio-psycho-social variable of interest using χ^2 test. [5]Likert-type scales were applied for self-rating scales ranging from 1 "extremely unhappy" to 6 "extremely happy" with the exception of overall health ranging from 1 "poor" to 4 "excellent."

3.1.5. Follow-up: quality control and bivariate relationships of multimodal brain age and bio-psycho-social factors

Due to the strong variability in sex β-values across models (Figures 2, 4–6), we also ran the described analyses separately for males and females showing some differences in model performance. For example, bio-psycho-social models explained a differential of between 1 and 4% of conditional variance for males (Supplementary Table S10) and differences in contributions of the different models' predictors, predictors' β-values being generally higher for males (Supplementary Figure S2). Overall, quality checks show small levels of multicollinearity, and that each predictor contributes individual to the models (Supplementary Figures S2–S10 and Supplementary Tables S8, S9), supporting assumptions about the robustness of the utilized models, as well as that simply adding all variables together saturates the model leading to lower model performance than at baseline across brain ages based on different diffusion approaches with a differential in marginal $R^2 = 3.38\%$.

Finally, for a better understanding of bivariate relationships, Table 2 gives an overview of brain age calculated from combined single and multi-shell diffusion data in relation to the observed bio-psycho-social factors. Strongest standardized associations when adding single factors to a model explaining brain age from age were found for WHR ($\beta_{std} = 0.07$, $p < 0.001$), PP ($\beta_{std} = 05$, $p < 0.001$), and overall health rating ($\beta_{std} = -0.04$, $p < 0.001$), and health satisfaction ($\beta_{std} = -0.03$, $p < 0.001$). Strongest brain age group differences were

found for sex ($\beta_{std} = -0.09$, $p = 0.001$), diabetes ($\beta_{std} = 0.02$, $p < 0.001$), and hypertension ($\beta_{std} = 0.06$, $p < 0.001$).

4. Discussion

We assessed the influence of various bio-psycho-social variables on brain age estimated from different diffusion approaches (and their combinations). As predicted, linear mixed effects models showed that bio-psycho-social variables uniquely explain a small proportion of brain age variability consistently across models, and estimates overlap for most predictors. Health and lifestyle factors were most indicative of brain age. However, differences in brain age variance explained between bio-psycho-social models and diffusion approaches were small. Significant predictors of brain age were job satisfaction, health satisfaction, WHR (and to a lesser extent BMI when excluding WHR as a predictor), diabetes, hypertension, any vascular diagnosis, daily coffee consumption, smoking, birth weight, matrix puzzles, and symbol digit substitution performance. Our findings indicate that brain age estimates derived from different diffusion approaches relate similarly to the examined bio-psycho-social factors. This is an important finding as it reveals that different WM characteristics share common aging associations, which are detailed by bio-psycho-social factor associations. The presented diffusion approaches are based on different theoretical assumptions for deriving a set of WM features. For example, DTI and DKI metrics are usually quite sensitive to a

broad range of WM changes due to their integrative nature of the scalar metrics (Basser et al., 1994; Jensen et al., 2005), i.e., DTI's FA or DKI's MK allow one to detect and localize the WM changes but not to explain their origins. In turn, dMRI approaches such as SMTmc or BRIA offer several metrics potentially allowing us to bind WM architecture with their predictive power (Kaden et al., 2016a,b; Reisert et al., 2017). For example, the intra-axonal water fraction appearing in both models might correlate with axonal density and axon diameter (Jelescu et al., 2020). Consequently, the metric provides information about WM maturation associated with aging leading to similar associations with age, aging and aging-related variables as DTI/DKI models. This encourages the usage of both conventional and advanced diffusion approaches when examining the relationship of bio-psycho-social factors and WM. Particularly, the application of dMRI approaches with more accurate assumptions around biophysical processes such as a ratio between intra-and extra-axonal diffusivities, permeability and other features offers various opportunities to investigate aging and associated diseases.

4.1. Explaining brain age from bio-psycho-social factors

Recent research has made a strong case for the conjunction effects of various bio-psycho-social factors in explaining general health (Lehman et al., 2017). Applied to brain age, for example, cardiometabolic effects have been shown to influence brain age (Beck et al., 2022a,b). However, assessments of how much of the variance explained in brain age above and beyond age, sex, age-by-sex interaction, and scanner site have not been described in the literature. We find close-to-zero added brain age variance explained by models including single bio-psycho-social variables (Table 2). Principal components of the health and lifestyle, life satisfaction, socio-demographics, and cognitive ability variables also added only small levels of brain age variance explained to the baseline model. A comparably larger proportion of brain age variance ($R^2 < 4\%$) is uniquely explained by health and lifestyle, life satisfaction, socio-demographics, and cognitive ability variables underlying the principal components (Figure 2 and Supplementary Table S7). These results suggest to include the different bio-psycho-social variables as predictors in order to explain brain age and the full covariance structure rather than using components which reduces the covariance matrix.

Health and lifestyle factors explained most brain age variance when added to the baseline model, followed by life satisfaction, and sociodemographic factors. Adding cognitive scores to the baseline model decreased brain age variance explained by the model (Figure 2). This suggests that biological and psychological factors are more influential than demographic factors. In turn, the observed bio-psycho-social factors are not independent of each other. Thus, we assume that bio-psycho-social factors contribute to explanations of brain age conjunctively. Additionally, we revealed that the added variance explained was small across models. A potential reason for small added R^2 values might lay in multiple confounder effects and heterogeneity in effects across covariate levels (Table 2 fallacy, Westreich and Greenland, 2013). Importantly, the added brain age variance explained is not just an effect of adding predictors randomly to the model, which rather decreases the variance explained, as shown

when adding all bio-psycho-social variables to the model. Hence, it seems more sensible to employ models incorporating several compared to single domain-specific variables to explain brain age. However, our results also indicate that a large part of the variance in brain age cannot be explained by our proposed bio-psycho-social models. Whether this unexplained variance is due to actual biologically founded individual differences, or the characteristics of brain age, for example, how the metric is being estimated (de Lange et al., 2022), remains unclear. BAG might also be rather static and indicated by constants such as genetic architecture and birth weight (Vidal-Pineiro et al., 2021). This would explain the smaller influence of more variable bio-psycho-social variables. Strong deviations from the norm, for example, due to atrophy will also have a strong influence on brain age (Kaufmann et al., 2019). Hence, for diseases impacting brain structure, brain age can be a useful indicator of health status (Kaufmann et al., 2019). Potentially, the health and lifestyle factors which are most likely to impact brain structure are therefore also more predictive of brain age than other bio-psycho-social variables (Figures 4–7). While our models failed to explain larger proportions of the variance of brain age, there are various interesting phenotype associations within these models which will be discussed in the following.

4.1.1. The importance of age, sex, and ethnicity

Usually, age, sex, and at times, scanner site, are used as covariates for brain age-phenotype associations as they are expected to influence various phenotypes (Jirsaraie et al., 2022). As brain age reflects chronological age, age also explains most of the brain age variance (Figures 4–7). We also find that the effects of sex and the sex-age interaction were highly variable across diffusion models predicting brain age with sex and the sex-age interaction being mostly non-significant predictors across diffusion models (Figures 4–7). Nevertheless, brain age does significantly differ between sexes (Sanford et al., 2022; Subramaniapillai et al., 2022), and we cannot exclude sex difference in WM microstructure. These relationships might also lead to differences in WM brain ages between sexes. Furthermore, models were more predictive of bio-psycho-social factors in males than females (Supplementary Table S2 and Supplementary Figure S2). Where the influence of sex changes based on the model construction, while potentially also influencing the model (Figures 4–7 and Supplementary Figures S5–S8). Some of the observed sex differences might be based on anatomical features, such as higher intracranial volume in males and different sex-specific aging (Eikenes et al., 2022). Brain age was differentially sensitive to ethnicity dependent on the approach it was calculated on (Figures 4–7), with these differences being influenced by sex (Supplementary Figure S2). A previous study showed that being a UK immigrant might influence brain age estimates (Leonardsen et al., 2022). Potentially, genetic contributions to brain age both estimated from T1-weighted (Ning et al., 2020; Vidal-Pineiro et al., 2021) and dMRI data (Salih et al., 2021) also have a connection with the mentioned brain age differences by sex and ethnicity. However, the causal structure of sex and ethnicity differences in brain age estimates requires further investigation.

Previous research has shown the effects of sex on metrics derived from conventional and advanced diffusion approaches, such as BRIA, DKI, DTI, NODDI, RSI, SMT, SMT mc, and WMTI (Beck et al., 2021; Eikenes et al., 2022). While a systematic assessment of sex-related effects on diffusion metrics from both conventional and advanced

dMRI approaches from voxel-to-whole-brain averages over the lifespan is yet to be established, different studies presented sex-related developmental trajectories in the structural connectome in children (Ingalhalikar et al., 2014), and sex related WM changes during aging (Hsu et al., 2008). Furthermore, sex differences in aging reflected in WM microstructure can be expected due to menopause and cascading biological processes, affecting both brain and body systems in various ways (Barth and de Lange, 2020; Mosconi et al., 2021; Lohner et al., 2022). Hence, developmental trajectories differing between males and females can be expected which makes sex-separated analyses useful to providing important additional information (e.g., as in Subramaniapillai et al., 2022). To which extend this applies to ethnicity requires further research. Hence, further research is required to delineate the underlying causal structure of sex and ethnicity to explain their highly variable associations with brain age.

4.1.2. Health and lifestyle factors

Interestingly, while the health and lifestyle factors models explained only a small proportion of the brain age variance, most of its predictors were significant. Furthermore, these predictors are generally only weakly correlated (Supplementary Figure S5), but when added in conjunction explaining more variability in brain age than on their own (compare Table 2 and Figure 5). To a certain degree, this is not surprising, due to dependencies between these predictors. For example, WHR, being the strongest predictor of brain age (see Figure 5), shows a clear relationship with pulse pressure (Supplementary Figure S5). For the extreme cases, this is expressed in a well-established relationship between obesity and hypertension (Kotsis et al., 2010) or any vascular diagnosis (Mathew et al., 2008). This is reflected in brain age, where minimum and maximum values show that there is an expected difference of up to 4 years in brain age between those with lowest compared to highest WHR, or a 2.4-years brain age difference between mean and maximum WHR. Interestingly, blood pressure is expected to increase with age, and higher blood pressure is positively associated with BAG (Cherbuin et al., 2021). However, these effects were not exclusively driven by hypertension but across the spectrum of measured blood pressure values (Cherbuin et al., 2021). This was supported by our findings showing both an effect of pulse pressure and hypertension on brain age. These effects are not surprising, as hypertension has been suggested as one of the most important risk factors for various cerebrovascular complications such as cerebral small vessel disease and resulting cognitive impairments (Meissner, 2016; Forte et al., 2019).

Another aspect of high WHR and BMI is obesity increasing diabetes risk (Kahn et al., 2006). While the evidence for the direction of the effect of diabetes is mixed (Franke et al., 2013; Cole et al., 2018; Sone et al., 2022), we find participants with diabetes to show higher brain age than those without diabetes (Table 1 and Figure 5). Several complications within the central nervous system have been associated with diabetes, including morphological, electrophysiological, and cognitive changes, often in the hippocampus (Wrighten et al., 2009), just as WM lesions and altered metabolite ratios (van der Harten et al., 2006; Biessels and Reijmer, 2014), supporting the idea of higher brain age among those with diabetes. But also generally, the increase in risk of cardiovascular disease by WHR is mediated by BMI, systolic blood pressure, diabetes, lipids, and smoking (Gill et al., 2021). In relation to the brain, higher WHR has been generally associated with lower gray matter volume (Hamer and Batty, 2019;

Gurholt et al., 2021), and higher WM brain age (Beck et al., 2022a,b; Subramaniapillai et al., 2022). Hence, to which extent high WHR accelerates brain aging requires further investigation, which might be particularly informative when observed in combination with other health and lifestyle variables (Hamer and Batty, 2019) and sex (Subramaniapillai et al., 2022).

Negative health consequences of smoking (Erhardt, 2009) are reflected in smoker's cortex being thinner (Gurholt et al., 2021), and smokers' brains being 1.5 years older on average than non-smokers' brains (Table 2). Smoking is a known risk factor for cardiovascular health significantly increasing its mortality and inducing various negative downstream effects on health (Erhardt, 2009), with negative impacts on the reward system (Le Foll et al., 2022), repeatedly shown in rats (e.g., Gozzi et al., 2006; Kenny and Markou, 2006; Cao et al., 2013). It can hence be expected that both general and brain health are influenced by smoking, making it an important control variable in assessing brain age.

The findings for coffee on the other hand are mixed, suggesting coffee consumption to be generally positive for cardiovascular health and decreasing the risk of Parkinson's disease, stroke, and Alzheimer's (Nehlig, 2016). The consumption of higher doses of caffeine is, however, associated with smaller brain volume and an increased risk of dementia (Pham et al., 2022). Practically, the direct effect of the number of daily cups of coffee consumed is small in our study. It would require on average 10 cups of coffee daily for an increase of 0.6 years of brain age, fitting the observations made by Pham et al. (2022). It also remains unclear whether the effect of coffee consumption on brain age is rather mediated by third variables such as poor sleep and mental health downstream effects which show direct negative effects on health (Distelberg et al., 2017). Additionally, there are vulnerable groups in which caffeine can cause adverse effects such as people with hypertension (Higdon and Frei, 2006). We conclude that health and lifestyle factors function in synergy in influencing brain age.

4.1.3. Health perception and satisfaction, and job satisfaction

We find significant assignations of self-rated health, friendship/relationship satisfaction, and job satisfaction with brain age. Self-assessments and self-rated scores are some of the fastest and easiest assessments. Yet, their reliability is under constant scrutiny, particularly when assessing health outcomes (e.g., Crossley and Kennedy, 2002; Reychav et al., 2019). In our study, self-rated overall health was a significant predictor of brain age, suggesting that asking participants about their health can be a useful preliminary assessment of different aspects of health. Self-rated health was additionally moderately correlated with health perception (Supplementary Figure S4), indicating both variables measure, to a certain degree, the same underlying phenomenon. However, self-rated health brain age associations were stronger and more variable across diffusion approaches' brain ages (Figure 6). These associations support the idea of brain age is not only indicative of brain health, but also overall health (Kaufmann et al., 2019).

Lastly, there was a trend of individuals' job satisfaction being associated with brain age (Figure 6). Conceptually, this would not be surprising as associations between wealth and health (e.g., Adler and Ostrove, 1999) as well as job (e.g., Faragher et al., 2013) and

financial satisfaction and health (e.g., Hsieh, 2001) have already been investigated. However, in the case of our study, higher job satisfaction was also indicative of higher brain age. Potential reasons are speculative but might reflect the tendency of people engaged in their jobs to work long hours which has previously been related with various negative mental and physical health outcomes (Lim et al., 2010; Bannai and Tamakoshi, 2014). Nevertheless, the underlying mechanisms of the associations between these single items in their relationship with brain age require further investigation.

4.1.4. Cognitive scores

Cognitive scores' impact on brain age might be small in the current study, yet still important in general (Table 2). This might be due to the selection of the observed cognitive test scores, with many more possible tests to be included which are potentially more indicative of brain age, such as IQ (Elliott et al., 2021). Another opportunity lies in assessing associations of cognitive performance and brain age in clinical groups. For example, brain age has been found to be explanatory of symbol digit modality test scores in multiple sclerosis suggesting brain age as a biomarker for cognitive dysfunction (Denissen et al., 2022). Similar to such findings, we find a similarly sized effect of symbol digit substitution test scores in our healthy aging data (Figure 7). Associations of cognitive performance and brain age are also sensitive to sex. For example, the number of solved matrix puzzles showing an effect when analyzing males and females data together seemed to be a predictor of brain age only in females when analyzing females from males data separately (Supplementary Figure S2). The quality of these differences requires further investigation.

4.2. Variability in brain age-phenotype relationships

Imaging phenotypes derived from diffusion UKB data contribute to a small additional proportion of the variability in the obtained results. However, the presented comparison of R^2 differences (Figure 3) underestimates the effects of single bio-psycho-social factors, and has to be interpreted with care, with cognitive function, life satisfaction, and health and lifestyle factors significantly adding to the baseline model (Figure 3). Yet, the used brain age estimation model might also introduce variability in brain age phenotype associations. Problematically, model evaluation metrics such as R^2, MAE, or RMSE depend additionally on cohort-and study-specific data characteristics making brain age model comparison across the literature not straightforward (de Lange et al., 2022). Additionally, there are differences between models trained on voxel-level compared to region-averaged data. Deep learning models using voxel-level data reach age predictions errors as low as MAE=2.14 years in midlife to late adulthood (Peng et al., 2021) or MAE=3.90 years across the lifespan (Leonardsen et al., 2022) while explaining large proportions of variance in age ($R^2 > 0.90$), whereas models trained on regional and global average measures predict age usually with larger error, MAE>3.6 years, and/or lower variances explained $R^2 < 0.75$ (de Lange et al., 2020a,b; Beck et al., 2021, 2022b; Rokicki et al., 2021; Korbmacher et al., 2022). However, Niu et al. (2019) showed that with different shallow and deep machine learning algorithms (ridge regression, support vector regressor, Gaussian process

regressor, deep neural networks) high prediction accuracies ($R^2 > 0.75$, MAE<1.43) could be reached when using multimodal regional average data using a young sample with narrow age range. Nonetheless, the same database (UKB) is able to provide similar patterns of detected associations between brain age and used phenotypes by applying different samples, modalities, and methods to calculate brain age. For example, diabetes diagnosis, diagnosed vascular problems or place of birth (see Figure 4 in Leonardsen et al., 2022), hip circumference, trail-making tasks, and matrix pattern completion were significantly associated with brain age (see Table 5 in Cole, 2020). However, it remains unclear whether the differences in the findings are due to analysis degree of freedom, sample characteristics, or actual bio-physical manifestations. For instance, the underlying data used for brain age estimation can be based on different modalities, e.g., dMRI metrics, as in the present work, versus T1-weighted images in Cole (2020) and Leonardsen et al. (2022). We can assume that WM-derived brain age associations with bio-psycho-social factors are relatively stable across diffusion approaches (see Figures 2, 4–6). We used four mixed models grouping (a) demographics, (b) cognitive, (c) life satisfaction, and (d) health and lifestyle variables to predict brain age. In contrast, Cole (2020) predicted bias-adjusted brain age from simple linear models with sex, age, and age^2 as covariates, and Leonardsen et al. (2022) observed similar associations for uncorrected brain age predicted from the respective phenotype and age and sex as covariates. However, bio-psycho-social variables are likely to interact in a complex pattern when explaining variables such as brain age. If we add only single bio-psycho-social variables, such as waist-to-hip-ratio, to a baseline model and then compare the two models, the differences in variance explained are small. Adding blocks of meaningfully related variables leads to stronger increases in Brian age variance explained (compare Table 2 and Supplementary Table S7). In summary, there are various sources of variability in brain age prediction. Phenotype associations could encompass not only the underlying data but also researchers' degree of freedom such as data selection, processing, and analysis.

5. Conclusion and future directions

Bio-psycho-social factors contribute similarly to explaining WM brain age across conventional and advanced diffusion MRI approaches when arranged as cognitive scores, life satisfaction, health and lifestyle factors, but not socio-demographics. Focusing on single predictors, health and lifestyle factors, WHR, birth weight, diabetes, hypertension, and related diagnoses, as well as smoking status and coffee consumption, were more predictive of brain age than cognitive and life satisfaction measures. Apart from health satisfaction and self-ratings, we found relationships of life satisfaction variables with brain age to be non-significant. Of the cognitive scores, only the digit substitution task performance was a significant predictor, which might be relevant in samples from midlife to old age. Furthermore, the influence of sex and ethnicity is largely variable suggesting the usage of sensible control mechanisms, such as separate analyses or exclusions in case of strongly imbalanced samples. We recommend future study designs taking observable interactions between the different bio-psycho-social effects into account. A potentially helpful guiding principle in the search for bio-psycho-social variables affecting brain age could be to focus on measures which are directly or indirectly related to or reflect pathology.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving human participants were reviewed and approved by National Health Service National Research Ethics Service (ref 11/NW/0382). The patients/participants provided their written informed consent to participate in this study.

Author contributions

MK: study design, software, formal analysis, visualizations, project administration, writing –original draft, and writing – review and editing. TG: writing – review and editing. A-ML and DM: software, writing – review and editing. AL: funding acquisition. EE: writing – review and editing, and funding acquisition. DB: writing – review and editing. OA: writing – review and editing, and funding acquisition. LW: writing – review and editing, and funding acquisition. IM: supervision, study design, data pre-processing and quality control, writing – review and editing, and funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

OA has received a speaker's honorarium from Lundbeck and is a consultant to Coretechs.ai.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1117732/full#supplementary-material

References

Adler, N. E., and Ostrove, J. M. (1999). Socioeconomic status and health: what we know and what we don't. *Ann. N. Y. Acad. Sci.* 896, 3–15. doi: 10.1111/j.1749-6632.1999.tb08101.x

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N. K., Andersson, J. L., Griffanti, L., Douaud, G., et al. (2018). Image processing and quality control for the first 10,000 brain imaging datasets from UK biobank. *NeuroImage* 166, 400–424. doi: 10.1016/j.neuroimage.2017.10.034

Andersson, J. L., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019

Bannai, A., and Tamakoshi, A. (2014). The association between long working hours and health: a systematic review of epidemiological evidence. *Scand. J. Work Environ. Health* 40, 5–18. doi: 10.5271/sjweh.3388

Barth, C., and de Lange, A. M. G. (2020). Towards an understanding of women's brain aging: the immunology of pregnancy and menopause. *Front. Neuroendocrinol.* 58:100850. doi: 10.1016/j.yfrne.2020.100850

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66, 259–267. doi: 10.1016/S0006-3495(94)80775-1

Beck, D., de Lange, A. M. G., Alnæs, D., Maximov, I. I., Pedersen, M. L., Leinhard, O. D., et al. (2022a). Adipose tissue distribution from body MRI is associated with cross-sectional and longitudinal brain age in adults. *NeuroImage* 33:102949. doi: 10.1016/j.nicl.2022.102949

Beck, D., de Lange, A. M. G., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., et al. (2021). 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: 10.1016/j.neuroimage.2020.117441

Beck, D., de Lange, A. M. G., Pedersen, M. L., Alnæs, D., Maximov, I. I., Voldsbekk, I., et al. (2022b). Cardiometabolic risk factors associated with brain age and accelerate brain ageing. *Hum. Brain Mapp.* 43, 700–720. doi: 10.1002/hbm.25680

Behler, A., Kassubek, J., and Müller, H. P. (2021). Age-related alterations in DTI metrics in the human brain—consequences for age correction. *Front. Aging Neurosci.* 13:682109. doi: 10.3389/fnagi.2021.682109

Biessels, G. J., and Reijmer, Y. D. (2014). Brain changes underlying cognitive dysfunction in diabetes: what can we learn from MRI? *Diabetes* 63, 2244–2252. doi: 10.2337/db14-0348

Billiet, T., Vandenbulcke, M., Mädler, B., Peeters, R., Dhollander, T., Zhang, H., et al. (2015). Age-related microstructural differences quantified using myelin water imaging and advanced diffusion MRI. *Neurobiol. Aging* 36, 2107–2121. doi: 10.1016/j.neurobiolaging.2015.02.029

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons

Cao, J., Wang, J., Dwyer, J. B., Gautier, N. M., Wang, S., Leslie, F. M., et al. (2013). Gestational nicotine exposure modifies myelin gene expression in the brains of adolescent rats with sex differences. *Transl. Psychiatry* 3:e247. doi: 10.1038/tp.2013.21

Cetin-Karayumak, S., Di Biase, M. A., Chunga, N., Reid, B., Somes, N., Lyall, A. E., et al. (2020). White matter abnormalities across the lifespan of schizophrenia: a

126   Paper B

harmonized multi-site diffusion MRI study. *Mol. Psychiatry* 25, 3208–3219. doi: 10.1038/s41380-019-0509-y

Chen, T., and Guestrin, C. (2016). Xgboost: a scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785–794)

Chen, C. L., Kuo, M. C., Chen, P. Y., Tung, Y. H., Hsu, Y. C., Huang, C. W. C., et al. (2022). Validation of neuroimaging-based brain age gap as a mediator between modifiable risk factors and cognition. *Neurobiol. Aging* 114, 61–72. doi: 10.1016/j.neurobiolaging.2022.03.006

Cherbuin, N., Walsh, E. I., Shaw, M., Luders, E., Anstey, K. J., Sachdev, P. S., et al. (2021). Optimal blood pressure keeps our brains younger. *Front. Aging Neurosci.* 13:694982. doi: 10.3389/fnagi.2021.694982

Cole, J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiol. Aging* 92, 34–42. doi: 10.1016/j.neurobiolaging.2020.03.014

Cole, J. H., Ritchie, S. J., Bastin, M. E., Hernández, V., Muñoz Maniega, S., Royle, N., et al. (2018). Brain age predicts mortality. *Mol. Psychiatry* 23, 1385–1392. doi: 10.1038/mp.2017.62

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., et al. (2016). Ageing and brain white matter structure in 3,513 UK biobank participants. *Nat. Commun.* 7, 1–13. doi: 10.1038/ncomms13629

Crossley, T. F., and Kennedy, S. (2002). The reliability of self-assessed health status. *J. Health Econ.* 21, 643–658. doi: 10.1016/S0167-6296(02)00007-3

de Lange, A. M. G., Anatürk, M., Rokicki, J., Han, L. K., Franke, K., Alnæs, D., et al. (2022). Mind the gap: performance metric evaluation in brain-age prediction. *Hum. Brain Mapp.* 43, 3113–3129. doi: 10.1002/hbm.25837

de Lange, A. M. G., Anatürk, M., Suri, S., Kaufmann, T., Cole, J. H., Griffanti, L., et al. (2020a). Multimodal brain-age prediction and cardiovascular risk: the Whitehall II MRI sub-study. *NeuroImage* 222:117292. doi: 10.1016/j.neuroimage.2020.117292

de Lange, A. M. G., Barth, C., Kaufmann, T., Maximov, I. I., van der Meer, D., Agartz, I., et al. (2020b). Women's brain aging: effects of sex-hormone exposure, pregnancies, and genetic risk for Alzheimer's disease. *Hum. Brain Mapp.* 41, 5141–5150. doi: 10.1002/hbm.25180

de Lange, A. M. G., and Cole, J. H. (2020). Commentary: correction procedures in brain-age prediction. *NeuroImage* 26:102229. doi: 10.1016%2Fj.nicl.2020.102229

de Lange, A. M. G., Kaufmann, T., Quintana, D. S., Winterton, A., Andreassen, O. A., Westlye, L. T., et al. (2021). Prominent health problems, socioeconomic deprivation, and higher brain age in lonely and isolated individuals: a population-based study. *Behav. Brain Res.* 414:113510. doi: 10.1016/j.bbr.2021.113510

Denissen, S., Engemann, D. A., De Cock, A., Costers, L., Baijot, J., Laton, J., et al. (2022). Brain age as a surrogate marker for cognitive performance in multiple sclerosis. *Eur. J. Neurol.* 29, 3039–3049. doi: 10.1111/ene.15473

Distelberg, B. J., Staack, A., Elsen, K. D. D., and Sabaté, J. (2017). The effect of coffee and caffeine on mood, sleep, and health-related quality of life. *J. Caffeine Res.* 7, 59–70. doi: 10.1089/jcr.2016.0023

Eikenes, L., Visser, E., Vangberg, T., and Håberg, A. K. (2022). Both brain size and biological sex contribute to variation in white matter microstructure in middle-aged healthy adults. *Hum. Brain Mapp.* 44, 691–709. doi: 10.1002/hbm.26093

Elliott, M. L., Belsky, D. W., Knodt, A. R., Ireland, D., Melzer, T. R., Poulton, R., et al. (2021). Brain-age in midlife is associated with accelerated biological aging and cognitive decline in a longitudinal birth cohort. *Mol. Psychiatry* 26, 3829–3838. doi: 10.1038/s41380-019-0626-7

Elliott, L. T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K. L., Douaud, G., et al. (2018). Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* 562, 210–216. doi: 10.1038/s41586-018-0571-7

Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science* 196, 129–136. doi: 10.1126/science.847460

Erhardt, L. (2009). Cigarette smoking: an undertreated risk factor for cardiovascular disease. *Atherosclerosis* 205, 23–32. doi: 10.1016/j.atherosclerosis.2009.01.007

Faragher, E. B., Cass, M., and Cooper, C. L. (2013). "The relationship between job satisfaction and health: a meta-analysis" in *From stress to wellbeing*. ed. C. L. Cooper, vol. 1 (London: Palgrave Macmillan), 254–271.

Fawns-Ritchie, C., and Deary, I. J. (2020). Reliability and validity of the UK Biobank cognitive tests. *PLoS One* 15:e0231627. doi: 10.1371/journal.pone.0231627

Fieremans, E., Jensen, J. H., and Helpern, J. A. (2011). White matter characterization with diffusional kurtosis imaging. *NeuroImage* 58, 177–188. doi: 10.1016/j.neuroimage.2011.06.006

Forte, G., De Pascalis, V., Favieri, F., and Casagrande, M. (2019). Effects of blood pressure on cognitive performance: a systematic review. *J. Clin. Med.* 9:34. doi: 10.3390/jcm9010034

Franke, K., and Gaser, C. (2019). Ten years of BrainAGE as a neuroimaging biomarker of brain aging: what insights have we gained? *Front. Neurol.* 10:789. doi: 10.3389/fneur.2019.00789

Franke, K., Gaser, C., Manor, B., and Novak, V. (2013). Advanced BrainAGE in older adults with type 2 diabetes mellitus. *Front. Aging Neurosci.* 5:90. doi: 10.3389/fnagi.2013.00090

Ghaemi, S. N. (2009). The rise and fall of the biopsychosocial model. *Br. J. Psychiatry* 195, 3–4. doi: 10.1192/bjp.bp.109.063859

Gill, D., Zuber, V., Dawson, J., Pearson-Stuttard, J., Carter, A. R., Sanderson, E., et al. (2021). Risk factors mediating the effect of body mass index and waist-to-hip ratio on cardiovascular outcomes: Mendelian randomization analysis. *Int. J. Obes.* 45, 1428–1438. doi: 10.1038/s41366-021-00807-4

Gozzi, A., Schwarz, A., Reese, T., Bertani, S., Crestan, V., and Bifone, A. (2006). Region-specific effects of nicotine on brain activity: a pharmacological MRI study in the drug-naive rat. *Neuropsychopharmacology* 31, 1690–1703. doi: 10.1038/sj.npp.1300955

Gurholt, T. P., Kaufmann, T., Frei, O., Alnæs, D., Haukvik, U. K., van der Meer, D., et al. (2021). Population-based body–brain mapping links brain morphology with anthropometrics and body composition. *Transl. Psychiatry* 11, 295–212. doi: 10.1038/s41398-021-01414-7

Hamer, M., and Batty, G. D. (2019). Association of body mass index and waist-to-hip ratio with brain structure: UK biobank study. *Neurology* 92, e594–e600. doi: 10.1212/WNL.0000000000006879

Higdon, J. V., and Frei, B. (2006). Coffee and health: a review of recent human research. *Crit. Rev. Food Sci. Nutr.* 46, 101–123. doi: 10.1080/10408390500400009

Houenou, J., Wessa, M., Douaud, G., Leboyer, M., Chanraud, S., Perrin, M., et al. (2007). Increased white matter connectivity in euthymic bipolar patients: diffusion tensor tractography between the subgenual cingulate and the amygdalo-hippocampal complex. *Mol. Psychiatry* 12, 1001–1010. doi: 10.1038/sj.mp.4002010

Hsieh, C. M. (2001). Correlates of financial satisfaction. *Int. J. Aging Hum. Dev.* 52, 135–153. doi: 10.2190/9YDE-46PA-MV9C-2JRB

Hsu, J. L., Leemans, A., Bai, C. H., Lee, C. H., Tsai, Y. F., Chiu, H. C., et al. (2008). Gender differences and age-related white matter changes of the human brain: a diffusion tensor imaging study. *Neuroimage* 39, 566–577. doi: 10.1016/j.neuroimage.2007.09.017

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., et al. (2008). Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage* 39, 336–347. doi: 10.1016/j.neuroimage.2007.07.053

Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., et al. (2014). Sex differences in the structural connectome of the human brain. *Proc. Natl. Acad. Sci.* 111, 823–828. doi: 10.1073/pnas.1316909110

Jelescu, I. O., Palombo, M., Bagnato, F., and Schilling, K. G. (2020). Challenges for biophysical modeling of microstructure. *J. Neurosci. Methods* 344:108861. doi: 10.1016/j.jneumeth.2020.108861

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., and Smith, S. M. (2012). FSL. *NeuroImage* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., and Kaczynski, K. (2005). Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn Reson Med.* 53, 1432–1440. doi: 10.1002/mrm.20508

Jirsaraie, R. J., Kaufmann, T., Bashyam, V., Erus, G., Luby, J. L., Westlye, L. T., et al. (2022). Benchmarking the generalizability of brain age models: challenges posed by scanner variance and prediction bias. *Hum. Brain Mapp.* 44, 1118–1128. doi: 10.1002/hbm.26144

Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D., and Alexander, D. C. (2016a). Multi-compartment microscopic diffusion imaging. *NeuroImage* 139, 346–359. doi: 10.1016/j.neuroimage.2016.06.002

Kaden, E., Kruggel, F., and Alexander, D. C. (2016b). Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magn. Reson. Med.* 75, 1752–1763. doi: 10.1002/mrm.25734

Kahn, S. E., Hull, R. L., and Utzschneider, K. M. (2006). Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 444, 840–846. doi: 10.1038/nature05482

Kamagata, K., Andica, C., Hatano, T., Ogawa, T., Takeshige-Amano, H., Ogaki, K., et al. (2020). Advanced diffusion magnetic resonance imaging in patients with Alzheimer's and Parkinson's diseases. *Neural Regen. Res.* 15, 1590–1600. doi: 10.4103/1673-5374.276326

Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., et al. (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat. Neurosci.* 22, 1617–1623. doi: 10.1038/s41593-019-0471-7

Kellner, E., Dhital, B., Kiselev, V. G., and Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn. Reson. Med.* 76, 1574–1581. doi: 10.1002/mrm.26054

Kenny, P. J., and Markou, A. (2006). Nicotine self-administration acutely activates brain reward systems and induces a long-lasting increase in reward sensitivity. *Neuropsychopharmacology* 31, 1203–1211. doi: 10.1038/sj.npp.1300905

Korbmacher, M., de Lange, A. M., van der Meer, D., Beck, D., Eikefjord, E. N., Lundervold, A., et al. (2022). Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain age. *bio Rxiv*. doi: 10.1101/2022.09.29.510029

Kotsis, V., Stabouli, S., Papakatsika, S., Rizos, Z., and Parati, G. (2010). Mechanisms of obesity-induced hypertension. *Hypertens. Res.* 33, 386–393. doi: 10.1038/hr.2010.9

Lawrence, K. E., Nabulsi, L., Santhalingam, V., Abaryan, Z., Villalon-Reina, J. E., Nir, T. M., et al. (2021). Age and sex effects on advanced white matter microstructure measures in 15, 628 older adults: a UK Biobank study. *Brain Imaging Behav.* 15, 2813–2823. doi: 10.1007/s11682-021-00548-y

Le Bihan, D., and Iima, M. (2015). Diffusion magnetic resonance imaging: what water tells us about biological tissues. *PLoS biology* 13:e1002203. doi: 10.1371/journal.pbio.1002203

Le Foll, B., Piper, M. E., Fowler, C. D., Tonstad, S., Bierut, L., Lu, L., et al. (2022). Tobacco and nicotine use. *Nat. Rev. Dis. Primers.* 8, 1–16. doi: 10.1038/s41572-022-00346-w

Lehman, B. J., David, D. M., and Gruber, J. A. (2017). Rethinking the biopsychosocial model of health: understanding health as a dynamic system. *Soc. Personal. Psychol. Compass* 11:e12328. doi: 10.1111/spc3.12328

Leonardsen, E. H., Peng, H., Kaufmann, T., Agartz, I., Andreassen, O. A., Celius, E. G., et al. (2022). Deep neural networks learn general and clinically relevant representations of the ageing brain. *NeuroImage* 256:119210. doi: 10.1016/j.neuroimage.2022.119210

Lim, N., Kim, E. K., Kim, H., Yang, E., and Lee, S. M. (2010). Individual and work-related factors influencing burnout of mental health professionals: a meta-analysis. *J. Employ. Couns.* 47, 86–96. doi: 10.1002/j.2161-1920.2010.tb00093.x

Lohner, V., Pehlivan, G., Sanroma, G., Miloschewski, A., Schirmer, M. D., Stöcker, T., et al. (2022). Relation between sex, menopause, and white matter hyperintensities: the Rhineland study. *Neurology* 99, e935–e943. doi: 10.1212/WNL.0000000000200782

Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., and Beckmann, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Mol. Psychiatry* 24, 1415–1424. doi: 10.1038/s41380-019-0441-1

Mathew, B., Francis, L., Kayalar, A., and Cone, J. (2008). Obesity: effects on cardiovascular disease and its diagnosis. *J. Am. Board Family Med.* 21, 562–568. doi: 10.3122/jabfm.2008.06.080080

Maximov, I. I., Alnæs, D., and Westlye, L. T. (2019). Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: effects of artefact corrections on diffusion metrics and their age associations in UK Biobank. *Hum. Brain Mapp.* 40, 4146–4162. doi: 10.1002/hbm.24691

Maximov, I. I., van Der Meer, D., de Lange, A. M. G., Kaufmann, T., Shadrin, A., Frei, O., et al. (2021). Fast quality control method for derived diffusion metrics (YTTRIUM) in big data analysis: UK Biobank 18,608 example. *Hum. Brain Mapp.* 42, 3141–3155. doi: 10.1002/hbm.25424

McPhee, G. M., Downey, L. A., and Stough, C. (2019). Effects of sustained cognitive activity on white matter microstructure and cognitive outcomes in healthy middle-aged adults: a systematic review. *Ageing Res. Rev.* 51, 35–47. doi: 10.1016/j.arr.2019.02.004

Meissner, A. (2016). Hypertension and the brain: a risk factor for more than heart disease. *Cerebrovasc. Dis.* 42, 255–262. doi: 10.1159/000446082

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., et al. (2016). Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* 19, 1523–1536. doi: 10.1038/nn.4393

Molarius, A., Seidell, J. C., Sans, S., Tuomilehto, J., and Kuulasmaa, K. (1999). Waist and hip circumferences, and waist-hip ratio in 19 populations of the WHO MONICA project. *Int. J. Obesity* 23, 116–125. doi: 10.1038/sj.ijo.0800772

Mori, S., Wakana, S., Van Zijl, P. C., and Nagae-Poetscher, L. M. (2005). *MRI atlas of human white matter*. Amsterdam: Elsevier.

Mosconi, L., Berti, V., Dyke, J., Schelbaum, E., Jett, S., Loughlin, L., et al. (2021). Menopause impacts human brain structure, connectivity, energy metabolism, and amyloid-beta deposition. *Sci. Rep.* 11, 10867–10816. doi: 10.1038/s41598-021-90084-y

Nehlig, A. (2016). Effects of coffee/caffeine on brain health and disease: what should I tell my patients? *Pract. Neurol.* 16, 89–95. doi: 10.1136/practneurol-2015-001162

Ning, K., Zhao, L., Matloff, W., Sun, F., and Toga, A. W. (2020). Association of relative brain age with tobacco smoking, alcohol consumption, and genetic variants. *Sci. Rep.* 10:10. doi: 10.1038/s41598-019-56089-4

Niu, X., Zhang, F., Kounios, J., and Liang, H. (2019). Improved prediction of brain age using multimodal neuroimaging data. *Hum. Brain Mapp.* 41, 1626–1643. doi: 10.1002/hbm.24899

Novikov, D. S., Fieremans, E., Jespersen, S. N., and Kiselev, V. G. (2019). Quantifying brain microstructure with diffusion MRI: theory and parameter estimation. *NMR Biomed.* 32:e3998. doi: 10.1002/nbm.3998

Parikh, M. N., Chen, M., Braimah, A., Kline, J., McNally, K., Logan, J. W., et al. (2021). Diffusion MRI microstructural abnormalities at term-equivalent age are associated with neurodevelopmental outcomes at 3 years of age in very preterm infants. *Am. J. Neuroradiol.* 42, 1535–1542. doi: 10.3174/ajnr.A7135

Pavlakis, A. E., Noble, K., Pavlakis, S. G., Ali, N., and Frank, Y. (2015). Brain imaging and electrophysiology biomarkers: is there a role in poverty and education outcome research? *Pediatr. Neurol.* 52, 383–388. doi: 10.1016/j.pediatrneurol.2014.11.005

Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., and Smith, S. M. (2021). Accurate brain age prediction with lightweight deep neural networks. *Med. Image Anal.* 68:101871. doi: 10.1016/j.media.2020.101871

Pham, K., Mulugeta, A., Zhou, A., O'Brien, J. T., Llewellyn, D. J., and Hyppönen, E. (2022). High coffee consumption, brain volume and risk of dementia and stroke. *Nutr. Neurosci.* 25, 2111–2122. doi: 10.1080/1028415X.2021.1945858

Pines, A. R., Cieslak, M., Larsen, B., Baum, G. L., Cook, P. A., Adebimpe, A., et al. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Dev. Cogn. Neurosci.* 43:100788. doi: 10.1016/j.dcn.2020.100788

Reisert, M., Kellner, E., Dhital, B., Hennig, J., and Kiselev, V. G. (2017). Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. *NeuroImage* 147, 964–975. doi: 10.1016/j.neuroimage.2016.09.058

Remiszewski, N., Bryant, J. E., Rutherford, S. E., Marquand, A. F., Nelson, E., Askar, I., et al. (2022). Contrasting case-control and normative reference approaches to capture clinically relevant structural brain abnormalities in patients with first-episode psychosis who are antipsychotic naive. *JAMA Psychiat.* 79, 1133–1138. doi: 10.1001/jamapsychiatry.2022.3010

Reychav, I., Beeri, R., Balapour, A., Raban, D. R., Sabherwal, R., and Azuri, J. (2019). How reliable are self-assessments using mobile technology in healthcare? The effects of technology identity and self-efficacy. *Comput. Hum. Behav.* 91, 52–61. doi: 10.1016/j.chb.2018.09.024

Rokicki, J., Wolfers, T., Nordhøy, W., Tesli, N., Quintana, D. S., Alnæs, D., et al. (2021). Multimodal imaging improves brain age prediction and reveals distinct abnormalities in patients with psychiatric and neurological disorders. *Hum. Brain Mapp.* 42, 1714–1726. doi: 10.1002/hbm.25323

Salih, A., Boscolo Galazzo, I., Raisi-Estabragh, Z., Rauseo, E., Gkontra, P., Petersen, S. E., et al. (2021). Brain age estimation at tract group level and its association with daily life measures, cardiac risk factors and genetic variants. *Sci. Rep.* 11, 20563–20514. doi: 10.1038/s41598-021-99153-8

Sanford, N., Ge, R., Antoniades, M., Modabbernia, A., Haas, S. S., Whalley, H. C., et al. (2022). Sex differences in predictors and regional patterns of brain age gap estimates. *Hum. Brain Mapp.* 43, 4689–4698. doi: 10.1002/hbm.25983

Sasson, E., Doniger, G. M., Pasternak, O., and Assaf, Y. (2010). Structural correlates of memory performance with diffusion tensor imaging. *NeuroImage* 50, 1231–1242. doi: 10.1016/j.neuroimage.2009.12.079

Shaked, D., Leibel, D. K., Katzel, L. I., Davatzikos, C., Gullapalli, R. P., Seliger, S. L., et al. (2019). Disparities in diffuse cortical white matter integrity between socioeconomic groups. *Front. Hum. Neurosci.* 13:198. doi: 10.3389/fnhum.2019.00198

Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155. doi: 10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23, S208–S219. doi: 10.1016/j.neuroimage.2004.07.051

Smith, S. M., Vidaurre, D., Alfaro-Almagro, F., Nichols, T. E., and Miller, K. L. (2019). Estimation of brain age delta from brain imaging. *NeuroImage* 200, 528–539. doi: 10.1016/j.neuroimage.2019.06.017

Sone, D., Beheshti, I., Shinagawa, S., Niimura, H., Kobayashi, N., Kida, H., et al. (2022). Neuroimaging-derived brain age is associated with life satisfaction in cognitively unimpaired elderly: a community-based study. *Transl. Psychiatry* 12, 25–26. doi: 10.1038/s41398-022-01793-5

Subramaniapillai, S., Suri, S., Barth, C., Maximov, I. I., Voldsbekk, I., van der Meer, D., et al. (2022). Sex-and age-specific associations between cardiometabolic risk and white matter brain age in the UK Biobank cohort. *Hum. Brain Mapp.* 43, 3759–3774. doi: 10.1002/hbm.25882

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., et al. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 12:e1001779. doi: 10.1371/journal.pmed.1001779

Sullivan, E. V., and Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neurosci. Biobehav. Rev.* 30, 749–761. doi: 10.1016/j.neubiorev.2006.06.002

van der Harten, B., de Leeuw, F. E., Weinstein, H. C., Scheltens, P., and Biessels, G. J. (2006). Brain imaging in patients with diabetes: a systematic review. *Diabetes Care* 29, 2539–2548. doi: 10.2337/dc06-1637

Veraart, J., Fieremans, E., and Novikov, D. S. (2016). Diffusion MRI noise mapping using random matrix theory. *Magn. Reson. Med.* 76, 1582–1593. doi: 10.1002/mrm.26059

Vidal-Pineiro, D., Wang, Y., Krogsrud, S. K., Amlien, I. K., Baaré, W. F., Bartres-Faz, D., et al. (2021). Individual variations in 'brain age' relate to early-life factors more than to longitudinal brain change. *eLife* 10:e69995. doi: 10.7554/eLife.69995

Wade, D. T., and Halligan, P. W. (2017). The biopsychosocial model of illness: a model whose time has come. *Clin. Rehabil.* 31, 995–1004. doi: 10.1177/0269215517709890

Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612. doi: 10.1109/TIP.2003.819861

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjørnerud, A., Due-Tønnessen, P., Engvig, A., et al. (2010). Lifespan changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb. Cortex* 20, 2055–2068. doi: 10.1093/cercor/bhp280

Westreich, D., and Greenland, S. (2013). The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am. J. Epidemiol.* 177, 292–298. doi: 10.1093/aje/kws412

Wood, D. A., Kafiabadi, S., Al Busaidi, A., Guilhem, E., Montvila, A., Lynch, J., et al. (2022). Accurate brain-age models for routine clinical MRI examinations. *NeuroImage* 249:118871. doi: 10.1016/j.neuroimage.2022.118871

Wrighten, S. A., Piroli, G. G., Grillo, C. A., and Reagan, L. P. (2009). A look inside the diabetic brain: contributors to diabetes-induced brain aging. *Biochim. Biophys. Acta* 1792, 444–453. doi: 10.1016/j.bbadis.2008.10.013

Yap, Q. J., Teh, I., Fusar-Poli, P., Sum, M. Y., Kuswanto, C., and Sim, K. (2013). Tracking cerebral white matter changes across the lifespan: insights from diffusion tensor imaging studies. *J. Neural Transm.* 120, 1369–1395. doi: 10.1007/s00702-013-0971-7

Zhao, B., Zhang, J., Ibrahim, J. G., Luo, T., Santelli, R. C., Li, Y., et al. (2021). Large-scale GWAS reveals genetic architecture of brain white matter microstructure and genetic overlap with cognitive and mental health traits (n=17,706). *Mol. Psychiatry* 26, 3943–3955. doi: 10.1038/s41380-019-0569-z

Paper B   **129**

CONSIDERATIONS ON BRAIN AGE PREDICTIONS FROM REPEATEDLY SAMPLED DATA ACROSS TIME

M. Korbmacher, M.-Y. Wang, R. Eikeland, R. Buchert, O. A. Andreassen, T. Espeseth, E. Leonardsen, L. T. Westlye, I. I. Maximov, K. Specht

ORIGINAL ARTICLE

Brain and Behavior
Open Access   WILEY

Considerations on brain age predictions from repeatedly sampled data across time

Max Korbmacher[1,2,3] ![ORCID] | Meng-Yun Wang[3,4] | Rune Eikeland[3,4] | Ralph Buchert[5] |
Ole A. Andreassen[2,6] ![ORCID] | Thomas Espeseth[7,8] | Esten Leonardsen[2,7] |
Lars T. Westlye[2,7] ![ORCID] | Ivan I. Maximov[1,2] | Karsten Specht[3,4,9]

[1]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway

[2]Norwegian Centre for Mental Disorders Research (NORMENT), Oslo University Hospital & Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[3]Mohn Medical Imaging and Visualisation Center (MMIV), Bergen, Norway

[4]Department of Biological and Medical Psychology, University of Bergen, Bergen, Norway

[5]Department of Diagnostic and Interventional Radiology and Nuclear Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[6]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

[7]Department of Psychology, University of Oslo, Oslo, Norway

[8]Department of Psychology, Oslo New University College, Oslo, Norway

[9]Department of Education, UiT The Arctic University of Norway, Tromsø, Norway

Correspondence
Max Korbmacher, Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway.
Email: Max.Korbmacher@hvl.no

Funding information
H2020 European Research Council, Grant/Award Number: 847776; Norges Forskningsråd, Grant/Award Numbers: 223273, 276044; Helse Sør-Øst RHF, Grant/Award Number: 2022080

Abstract

Introduction: Brain age, the estimation of a person's age from magnetic resonance imaging (MRI) parameters, has been used as a general indicator of health. The marker requires however further validation for application in clinical contexts. Here, we show how brain age predictions perform for the same individual at various time points and validate our findings with age-matched healthy controls.

Methods: We used densely sampled T1-weighted MRI data from four individuals (from two densely sampled datasets) to observe how brain age corresponds to age and is influenced by acquisition and quality parameters. For validation, we used two cross-sectional datasets. Brain age was predicted by a pretrained deep learning model.

Results: We found small within-subject correlations between age and brain age. We also found evidence for the influence of field strength on brain age which replicated in the cross-sectional validation data and inconclusive effects of scan quality.

Conclusion: The absence of maturation effects for the age range in the presented sample, brain age model bias (including training age distribution and field strength), and model error are potential reasons for small relationships between age and brain age in densely sampled longitudinal data. Clinical applications of brain age models should consider of the possibility of apparent biases caused by variation in the data acquisition process.

1 | BACKGROUND: WHAT IS BRAIN AGE AND WHAT IS IT GOOD FOR?

Brain age refers to the estimation of a person's age from magnetic resonance imaging (MRI) parameters (Franke & Gaser, 2019). This has been done using either neural networks on 3D data (Leonardsen et al., 2022) or tabular data containing region-averaged metrics (Korbmacher et al., 2023; Vidal-Pineiro et al., 2021). Brain age becomes particularly interesting when assuming that lifespan changes in the brain follow normative patterns and that deviations from such patterns might be indicative of disease or disease development (Marquand et al., 2019; Kaufmann et al., 2019). An elevated predicted compared with chronological age in adults may be indicative of psychiatric, neurodegenerative, and neurological disorders (Kaufmann et al., 2019) and poorer health, for example measured by various cardiometabolic risk factors (Beck et al., 2022; Korbmacher et al., 2022). Hence, brain age is a promising developing biomarker of general brain health (Franke & Gaser, 2019).

However, revealing connections between brain age and structural and functional brain architecture is needed to fully understand the biological underpinnings of brain age and its potential clinical implications (Vidal-Pineiro et al., 2021). Furthermore, large cross-sectional samples are often used, which could obscure effects of predictive power of brain age by confounders, in particular, differences in MRI acquisition (Jirsaraie et al., 2022). Hence, contributions of individual differences to brain age estimates require a closer examination. With the aim of assessing the effects of automated MRI scan quality control (QC) metrics on brain age predictions, we used a pretrained deep neural network model (Leonardsen et al., 2022) to predict brain ages from densely sampled T1-weighted MRI data from three individuals (BBSC1–3) scanned in total $N_{BBSC} = 103$ times over a 1-year interval (Wang et al., 2022), and an independent data set including one individual (FTHP1) scanned $N_{FTHP} = 557$ times over a 3-year interval. We first observed within-subject prediction error and correlations between chronological and predicted age, revealing small, nonsignificant correlations and larger prediction errors than previously shown in between-subjects analyses. We then tested associations of QC metrics on brain age using linear random intercept models showing potential associations between QC parameters and brain age as well as associations between acquisition parameters and brain age. Finally, we validate the findings in cross-sectional data and investigate differences in the variability in predictions between longitudinal and cross-sectional datasets.

2 | RESULTS AND DISCUSSION

2.1 | Weak correlation between brain age and age

Crude within-subject correlations between age and brain age revealed differing directionalities of slopes across subjects, with only the FTHP1 correlation being statistically significant ($r = 0.38$, 95% CI [0.24; 0.51], $p < .001$; Figure 1).

This is likely due to the small age range and short interscan intervals, as illustrated by differences in model-innate error for the different subjects (Table 1) compared with error statistics across age groups ($MAE_{test} = 2.47$, $MAE_{external} = 3.90$, as presented in Leonardsen et al., 2022).

Additionally, the ages of BBSC1–3 fall into some of the least represented parts of the training data age distribution in the underlying model for the brain age predictions (see Leonardsen et al., 2022) which might contribute to explaining some of the prediction differences beyond model error and intraindividual age range across scanning sessions.

Yet, when using age-matched healthy controls from the cross-sectional TOP and NCNG samples (see *Materials and Methods* section) using BBSC and FTHP longitudinal participants' mean ages ± 5 years (presented in Table 1), correlations between age and brain age estimates were significant for age matches (representing subsamples of TOP and NCNG samples; Table 2).

Interestingly, we also find systematically underestimated brain ages across subjects (Figure 1) with the underestimations being stronger for a field strength of 3T than 1.5T for FTHP1 (Table 1), and as compared with age-matched cross-sectional data (Table 2). While longitudinal brain age predictions were more closely related to age at 3T MRI ($r_{partial} = 0.38$, 95% CI [0.24, 0.51], $p < .001$) than at 1.5T MRI ($r_{partial} = 0.06$, 95% CI [−0.04, 0.16], $p = .239$; Figure 2), the prediction error was smaller at 1.5T (Table 1), with these findings being robust to exclusions of back-to-back repeat scans acquired in the same session without repositioning of the head (Supplement 1). When using the out-of-sample test sets TOP and NCNG cross-sectional data, we find highly corresponding relationships between age and brain age at 1.5T ($r = 0.98$, 95% CI [0.97, 0.98], $p < .001$) and 3T ($r = 0.92$, 95% CI [0.91, 0.93], $p < .001$), but higher prediction error at 3T for age-matched subjects (Table 2) and overall (Supplement 3).
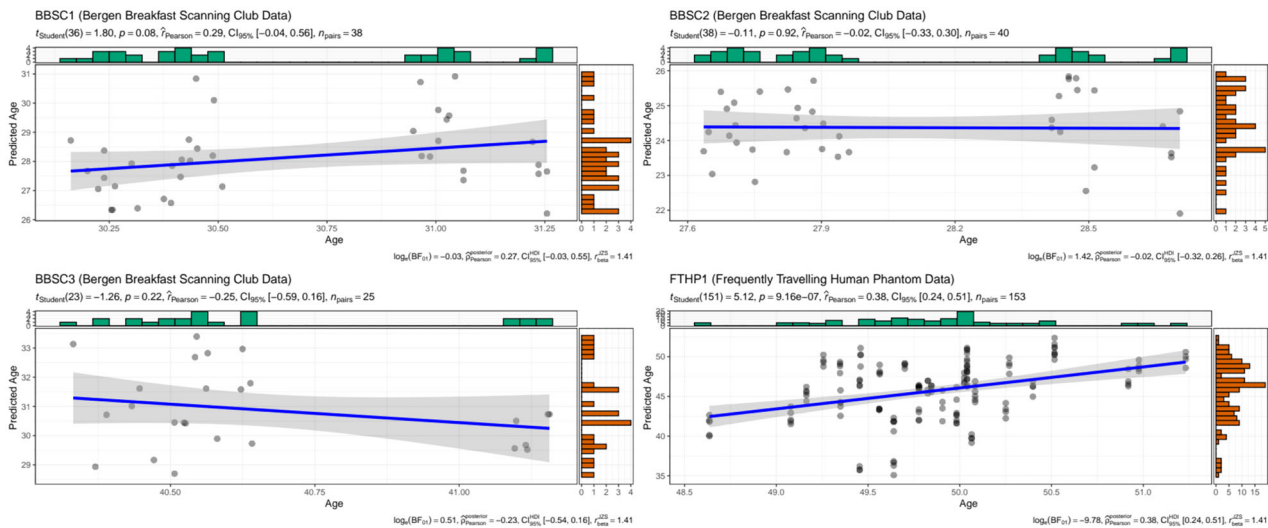
FIGURE 1 Intraindividual correlations between brain age and chronological age at 3T for BBSC1–3 and FTHP1. Dot color was gray, with overlapping dots presented as darker.

TABLE 1 Age, predicted age, brain age gap (BAG), and prediction error by subject and field strength.

Subject	Field strength	N observations	Mean age	SD age	Mean prediction	SD prediction	Mean BAG	SD BAG	MAE	RMSE
BBSC1	3T	38	30.66	0.38	28.13	1.25	−2.52	1.20	2.55	2.79
BBSC2	3T	40	28.09	0.38	24.37	0.95	−3.72	1.03	3.72	3.85
BBSC3	3T	25	40.66	0.28	30.87	1.37	−9.79	1.46	9.79	9.89
FTHP1	3T	153	49.86	0.54	45.71	3.70	−4.15	3.52	4.31	5.44
FTHP1	1.5T	394	49.64	0.46	48.39	2.52	−1.25	2.54	2.15	2.83

The presented data refer to the longitudinal, densely sampled data of few individuals.
BAG, brain age gap; MAE, mean absolute error; RMSE, root mean squared error. BAG is calculated as the difference between predicted age and age.

TABLE 2 Correlations between age-matching cross-sectional subsamples' ages and brain age estimates.

Matched subject	Field strength	$N_{subjects}$	Pearson's r [95% CI]*	Mean age	SD age	Mean prediction	SD prediction	Mean BAG	SD BAG	MAE	RMSE
BBSC1	3T	279	0.56 [0.47, 0.64]	30.64	2.74	28.34	4.10	−2.30	3.42	3.33	4.12
BBSC2	3T	269	0.62 [0.54, 0.69]	28.81	2.83	26.75	3.96	−2.05	3.13	3.02	3.74
BBSC3	3T	248	0.44 [0.34, 0.54]	40.71	2.95	37.86	5.21	−2.85	4.71	4.52	5.50
FTHP1	3T	113	0.71 [0.60, 0.79]	48.60	3.04	44.68	5.93	−3.91	4.34	4.59	5.84
FTHP1	1.5T	49	0.79 [0.65, 0.88]	49.61	3.22	51.98	4.40	2.38	2.71	2.91	3.58

Matched subject refers to the longitudinally sample subjects presented in Table 1. Mean ages for the respective subjects with an interval of five years were used to sample from the cross-sectional validation set consisting of 3T and 1.5T data from TOP and NCNG samples. BAG, brain age gap; MAE, mean absolute error; RMSE, root mean squared error. BAG is calculated as the difference between predicted age and age.
*All $p < .001$.

This emphasizes the importance of treating predictions for age groups which are underrepresented in the training sample and differences in field strength with care. In that sense, the observed within-subjects variability associated with acquisition- or scanner-specific effects might be used to estimate the minimum size of true within-subject changes (e.g., due to disease) to be detected with

a given power. Previous findings outlined the influence of scanner site on brain age predictions and scan quality (Jirsaraie et al., 2022; Leonardsen et al., 2022) indicated by the Euler number (Rosen et al., 2018). Lower quality scans lead to lower prediction errors. We hence hypothesize that there might be additional reasons for inaccuracies in brain age predictions caused by factors beyond the characteris-
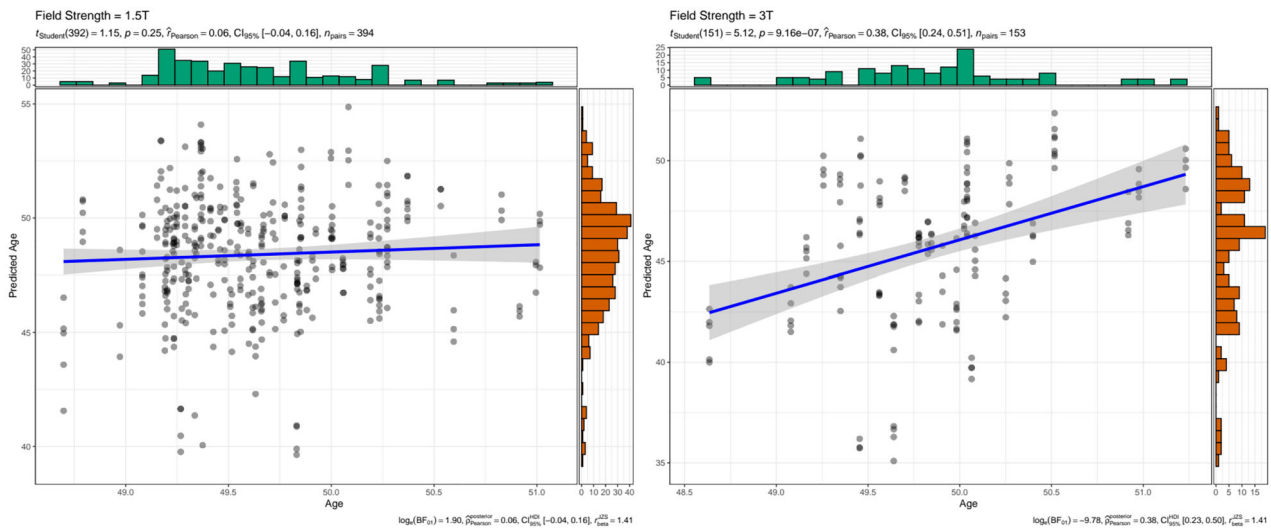
FIGURE 2 Intraindividual correlations between brain age and chronological age at 1.5T and 3T for FTHP1. Dot color was gray, with overlapping dots presented darker.

tics of the brain age model, in particular scan quality and acquisition parameters.

2.2 | Scan quality and acquisition: possible reasons for inaccurate brain age predictions?

We used linear random intercept models at the participant level to examine associations of individual QC metrics (see Figure 3; *Materials and Methods*) and brain age, while controlling for age in BBSC1–3. Entropy-focus criterion (EFC, $\beta_{std} = -0.489$, $p_{Holm} < .001$) and the foreground–background energy ratio (FBER, $\beta_{std} = 0.456$, $p_{Holm} < .001$) were significant predictors of brain age. In a separate analysis of FTHP1, scanned at different sites using different scanning parameters, we included scanner site, field strength, and slice thickness as random factors, rendering none of the QC metrics significant after correcting for multiple testing ($p_{Holm} = 1$).

Follow-up analyses in FTHP1 focused on examining acquisition parameters. We observed individual fixed effects of field strength, manufacturer, and slice thickness in one model each, while keeping scanner site and the other acquisition parameters as random effects at the level of the intercept, revealing only significant associations of field strength ($\beta = -1.141$, $p_{Holm} < .001$) with brain age.

For validation, we replicate this finding in healthy controls from the TOP and NCNG (see *Materials and Methods* section). We found differences in BAG at different field strengths ($\beta = -3.547$, $p < .001$), with Mean$_{BAG-1.5T} = 1.357 \pm 3.285$ and Mean$_{BAG-3T} = -2.19 \pm 4.06$ using the entire out-of-sample test data, with this difference being attenuated when regressing out age ($\beta = -5.318$, $p < .001$). When age-matching FTHP1 and including only the $N = 162$ participants aged 50 ± 5 years ($N = 49$ scanned at 1.5T), the effect of field strength appears stronger ($\beta = -6.294$, $p < .001$), with Mean$_{BAG-1.5T} = 2.38 \pm 2.71$ and Mean$_{BAG-3T} = -3.92 \pm 4.35$, yet

smaller when regressing out the age-effect ($\beta = -1.942$, $p < .001$). In the case of age-matching, also correlations between age and brain age are stronger at 1.5T compared with 3T (Table 2). This was also true when using the entire cross-sectional data (combining TOP and NCNG data), yet correlations between age and brain age were more similar at 1.5T ($r = 0.98$, 95% CI [0.97, 0.98], $p < .001$) and 3T ($r = 0.92$, 95% CI [0.91, 0.93], $p = .004$).

While our findings indicate an association between QC parameters EFC and FBER and brain age in all BBSC subjects when controlling for age and constant scanning parameters and scanner site, no QC parameters were significantly associated with brain age after adjustments for multiple comparisons in FTHP1. Based on that, one could speculate that scan quality impacts brain age predictions when participant ages are sampled from under-represented age groups within the prediction model. For example, Jirsaraie et al. (2022) showed that neural networks' reliability of brain age predictions was lowest at the ends of the age distributions across scanning sites, and predictions were less consistent when image quality was low. Furthermore, QC metrics might be sensitive to individual differences, and vary across scanner sites. FTHP1 results also suggest a strong effect of field strength on brain age. This indicates overall that brain age estimates are potentially dependent on intraindividual variables in addition to acquisition parameters and other scanner site-specific covariates. While we cannot generalize from the obtained single-subject results (FTHP1) on field strength, the additional analyses on external datasets support the effect of field strength congruent with Jirsaraie et al.'s (2022) findings of lower prediction errors at 1.5T compared with 3T. This was expressed in our analyses as generally higher brain age estimates at 1.5T compared with 3T, and higher prediction errors at 3T in both cross-sectional and longitudinal data. Finally, we show that prediction error in longitudinal data can be much higher than anticipated from cross-sectional estimates, without the presence of mental or physical disorder (see BBSC3 in Table 1; compare Tables 2 and Supplement 3).
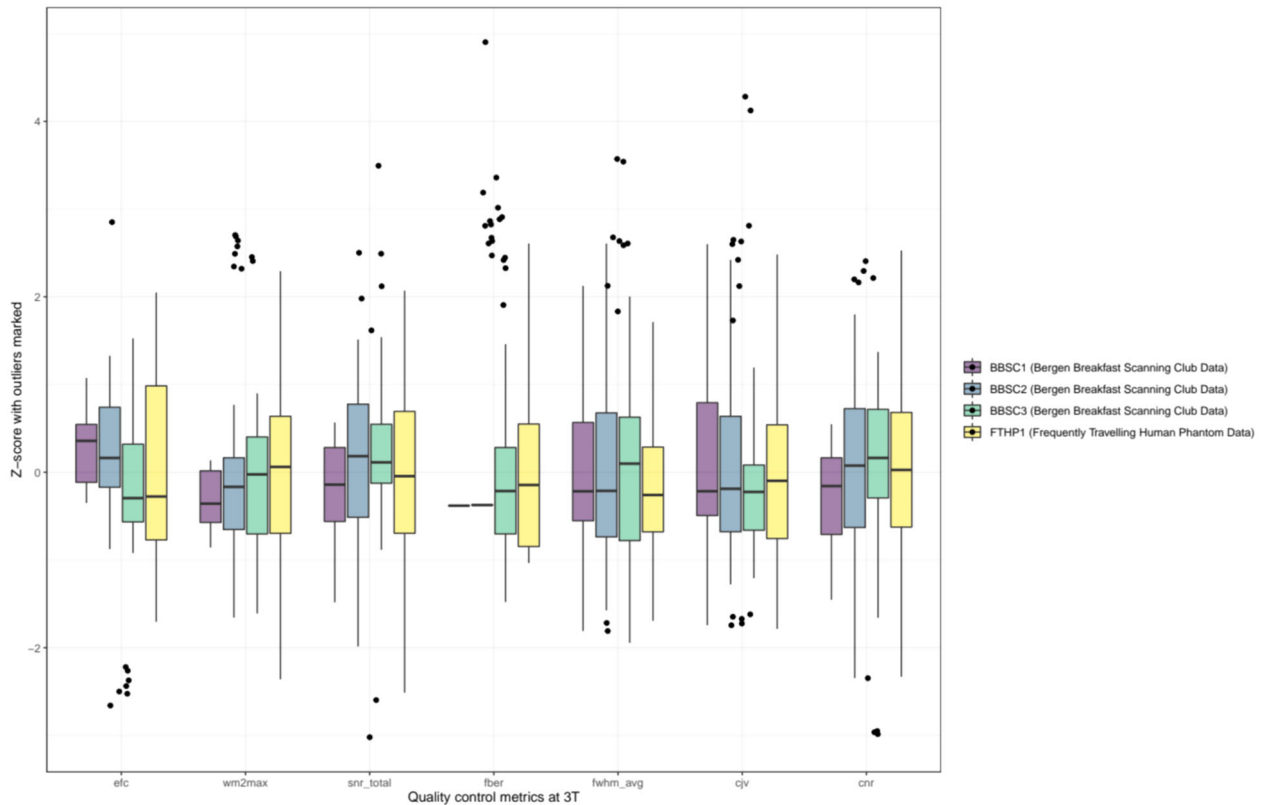
FIGURE 3 Standardized quality control metrics at 3T per subject. For an overview of scan quality control metrics at 1.5T (only applicable for FTHP1), see Supplement 2.

A potential approach for future brain age modelling could be to employ multiple, more specific models which are better tuned to individual differences, developmental trajectories, and scan quality. Such models could for example be trained on data with a smaller age range and a single field strength. Dependent on these parameters, brain age predictions can then be made by a model selected based on the available scan and group the individual belongs to.

3 | CONCLUSION

Variability in brain age predictions complicate the metric's clinical usage, for example, as a (pre-) diagnostic tool. We presented small correlations between age and brain age when repeatedly sampling T1-weighted MRI data from the same individual in a short period of time (1–3 years). Reasons might lay in the absence of maturation effects for the age range in the presented sample, brain age model bias (including a bimodal or trimodal age training distribution) and model error. While limited, our results suggest an influence of field strength and mixed evidence for scan quality on brain age. Individual differences and the processing of such in the brain age model, might lead to variability in associations between brain age and QC metrics. The presented testing of an established brain age model on multiple single-subject short-timespan retesting data is a stricter test than the usual use-case and does not invalidate MRI group differences. However, intraindivid-

ual differences contributing to brain age require further attention in order to advance brain age as a clinical tool.

4 | MATERIALS AND METHODS

4.1 | Participants

We used two datasets for the analyses which had received ethics approval with all participants consenting formally previously (Opfer et al., 2022; Wang et al., 2022, 2023). The first dataset was the Bergen Breakfast Scanning Club (BBSC) dataset (Wang et al., 2022, 2023), including three male subjects (BBSC2:start-age$_{BBSC2}$ = 27, BBSC1:start-age$_{BBSC1}$ = 30, and BBSC3:start-age$_{BBSC3}$ = 40) who were scanned over the period of circa 1 year with a summer break in the middle of the scanning period (Wang et al., 2022). This resulted in a total number of N_{BBSC} = 103 scans, relatively equally distributed across subjects (N_{BBSC1} = 38, N_{BBSC2} = 40, N_{BBSC3} = 25). The second dataset was the frequently travelling human phantom (FTHP) MRI dataset (Opfer et al., 2022), including one male subject (FTHP1:start-age$_{FTHP}$ = 48) with 157 imaging sessions at 116 locations, resulting in a total of N_{FTHP} = 557 MRI volumes. Of these, we excluded N = 6 volumes based on errors in the processing pipeline, resulting in a final sample for the main analyses of N_{FTHP} = 551. For QC (Supplement 1), we removed another N_{FTHP} = 25 volumes which were repeat-sequences run at the

same scanner and time without changing head position or acquisition parameters, resulting in a final sample for the supplemental analyses of $N_{FTHP} = 526$.

Finally, as additional validation data, we selected healthy controls from two of the cross-sectional out-of-sample test datasets described in Leonardsen et al. (2022): locally collected data (TOP; Tønnesen et al., 2018) and the Norwegian Cognitive NeuroGenetics sample (NCNG; Espeseth et al., 2012), as these provided most MRI scans on healthy controls. Together these datasets include a total of $N = 209$ scans of healthy controls at 1.5T (Mean$_{age}$ = 54.66 ± 15.51), and $N = 856$ scans of healthy controls at 3T (Mean$_{age}$ = 32.93 ± 10.55).

4.2 | Image acquisition and preprocessing

T1-weighted volumes of BBSC1–3 were acquired with a spin echo sequence (TE = 2.95 ms, TR = 6.88 ms, FA = 12°, TI = 450, 188 slices, slice thickness = 1mm, in-plane resolution = 1 mm × 1 mm, FOV = 256mm, isotropic voxel size = 1 mm^3) at a 3T GE system with 32-channel head coil (see Wang et al., 2022, 2023). T1-weighted volumes of FTHP1 were acquired at different scanners with various different scanning parameters (see Opfer et al., 2022 or https://www.kaggle.com/datasets/ukeppendorf/frequently-traveling-human-phantom-fthp-dataset). All imaging sites involved in the scanning of FTHP1 were informed that the scan was acquired for the purpose of MRI-based volumetry. Furthermore, all FTHP sites were asked to use acquisition parameters in accordance with the ADNI recommendations for magnetization prepared rapid gradient-echo (MP-RAGE) MRI for volumetric analyses. Thus, the range of FTHP acquisition parameters is representative of MRI-based volumetry in everyday clinical routine at nonacademic sites. However, the scan quality might be higher than during average clinical assessments, as only few scans were affected by motion artifacts (relatively young healthy subject). TOP data (Tønnesen et al., 2018) including only healthy controls were acquired at 3T on a GE 3T Signa HDxT (TR = 7.8 ms, TE = 2.956 ms, FA = 12°; one subset with HNS coil, one subset with 8HRBRAIN coil), and a GE 3T Discovery GE750 (TR = 8.16 ms, TE = 3.18 ms, FA = 12°). NCNG data (Espeseth et al., 2012) were acquired at a 1.5T Siemens Avanto scanner using two 3D MP-RAGE T1-weighted sequences (TR = 2400 ms, TE = 3.61 ms, TI = 1000 ms, FA = 8°, with 160 sagittal slices (1.3 × 1.3 × 1.2 mm)).

Before prediction, the volumes were automatically processed using Freesurfer version 5.3 (Fischl, 2012) and FSL version 6.0 (Jenkinson et al., 2012; Smith et al., 2004), both being widely used open-source software packages (see for overview of advantages and disadvantages compared with other packages: Man et al., 2015) which were validated in clinical and nonclinical samples (Clerx et al., 2015; Fischl, 2012; Jenkinson et al., 2012; Smith et al., 2004). The processing procedure included skull-stripping as part of Freesurfer's recon-all pipeline, linearly orienting to MNI152 space (6 degrees of freedom) using FSL's linear registration, and excess border removal. While linear registration in FSL is sensitive to atrophy and high levels of noise (Dadar et al., 2018), this does not apply for the current quality controlled

data including only healthy controls. As Freesurfer's skull stripping algorithm can include errors (Falkovskiy et al., 2016; Waters et al., 2019), the images were manually checked for accuracy. A step-by-step processing tutorial including necessary code can be found at https://github.com/estenhl/pyment-public.

4.3 | Brain age estimation

We applied a fully convolutional neural network (Gong et al., 2021; Peng et al., 2021) trained on 53,542 minimally processed MRI T1-weighted whole-brain images from individuals aged 3–95 years collected at a variety of scanning sites (both 1.5 and 3T field strength) (SFCN-reg detailed in Leonardsen et al., 2022) to estimate participants' ages directly from the MRI using Python v3.9.13. The model was tested in both clinical and nonclinical samples (Leonardsen et al., 2022) and presented high accuracy and test–retest reliability compared with other brain age models (Dörfel et al., 2023).

4.4 | QC metrics

QC metrics were extracted for each T1-weighted volume by using the automated MRIQC tool version 22.0.6 (Esteban et al., 2017). Of these metrics, we used those which are calculated for the whole brain or volume, being (1) noise measures: contrast-to-noise ratio, signal-to-noise ratio, coefficient of joint variation of gray and white matter, (2) measures based on information theory EFC and foreground–background energy ratio (FBER), (3) white-matter to maximum intensity (WM2MAX), and (4) other measures: full-width half-maximum (FWHM).

4.5 | Statistical analyses

All statistical analyses were conducted using R (v4.1.2). First, correlations of brain age with chronological age and additionally commonly used error metrics for brain age predictions (mean absolute error and root mean squared error) were assessed on a participant level. We further investigated associations between brain age and age in FTHP1 (from the Frequently Travelling Human Phantom dataset) when partialling out scanner site and field strength, as these were expected to influence prediction accuracy.

Further analyses focused on associations between QC metrics and brain age as well as acquisition parameters and brain age. We decided to observe each single independent variable of interest in a dedicated model, as model diagnostics indicated potential multicollinearity in models including multiple QC metrics. Furthermore, random effect models were chosen due to the possibility to account for variances being dependent on different grouping variables, such as ID, scanner site, field strength, and slice thickness.

Hence, linear random intercept models at the participant level were used to examine associations of individual QC metrics and brain age,

while controlling for age in the BBSC dataset, by running one model for each QC metric. Similarly, for dataset 2, we predicted each QC metric as a fixed effect in addition to the fixed effect of age in a single model. However, we used different random effects, namely, scanner site, field strength, and slice thickness, as dataset 2 contained only FTHP1.

We also examined single individual acquisition parameters in the FTHP dataset (including only one subject FTHP1) as fixed effects in addition to the fixed age effect. The acquisition parameters of interest were field strength, manufacturer, and slice thickness. Acquisition parameters not used as fixed effects were used as random effect at the level of the intercept in addition to scanner site. All p-values were adjusted for multiple testing using Holm correction, marked with p_{Holm}. Standardized β-values (β_{std}) for predictors were used for comparability across β-weights by scaling QC metrics for each subject individually.

Finally, as a validation step, we estimated brain ages for healthy controls in NCNG and TOP datasets and correlated the estimates with age for the entire sample, subjects which were age-matched to the longitudinal, densely sampled individuals mean age ± 5 years. This provided a baseline understanding for differences in inter and intra subject brain age variability. In a second step, brain age gap (BAG) was examined by field strength and scanner site in the validation sample.

AUTHOR CONTRIBUTION

Max Korbmacher: Study design; software; formal analysis; visualizations; project administration; writing—original draft; writing—review and editing. **Meng-Yun Wang**: Writing—review and editing. **Rune Eikeland**: Writing—review and editing. **Ralph Buchert**: Writing—review and editing. **Ole A. Andreassen**: Writing—review and editing; funding acquisition. **Thomas Espeseth**: Writing—review and editing; funding acquisition. **Esten Leonardsen**: Writing—review and editing. **Lars T. Westlye**: Writing—review and editing; funding acquisition. **Ivan I. Maximov**: Writing—review and editing. **Karsten Specht**: Writing—review and editing; funding acquisition.

ACKNOWLEDGMENTS

CONFLICT OF INTEREST STATEMENT

O. O. A. has received a speaker's honorarium from Lundbeck and is a consultant to Coretechs.ai.

DATA AVAILABILITY STATEMENT

Data processing pipeline and weights for the trained convolutional neural network can be found at https://github.com/estenhl/pyment-public. Processed tabular data and analysis code are made available at https://github.com/MaxKorbmacher/BBSC_Brain Age.

ORCID

Max Korbmacher https://orcid.org/0000-0002-8113-2560
Ole A. Andreassen https://orcid.org/0000-0002-4461-3568
Lars T. Westlye https://orcid.org/0000-0001-8644-956X

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/brb3.3219

REFERENCES

Beck, D., Lange, A.-M. G., Pedersen, M. L., Alnæs, D., Maximov, I. I., Voldsbekk, I., Richard, G., Sanders, A.-M., Ulrichsen, K. M., Dørum, E. S., Kolskår, K. K., Høgestøl, E. A., Steen, N. E., Djurovic, S., Andreassen, O. A., Nordvik, J. E., Kaufmann, T., & Westlye, L. T. (2022). Cardiometabolic risk factors associated with brain age and accelerate brain ageing. *Human Brain Mapping*, 43(2), 700–720. https://doi.org/10.1002/hbm.25680

Clerx, L., Gronenschild, E. H. B. M., Echavarri, C., Verhey, F., Aalten, P., & Jacobs, H. I. L. (2015). Can FreeSurfer compete with manual volumetric measurements in Alzheimer's disease? *Current Alzheimer Research*, 12(4), 358–367. https://doi.org/10.2174/1567205012666150324174813

Dadar, M., Fonov, V. S., & Collins, D. L, & Alzheimer's Disease Neuroimaging Initiative. (2018). A comparison of publicly available linear MRI stereotaxic registration techniques. *Neuroimage*, 174, 191–200. https://doi.org/10.1016/j.neuroimage.2018.03.025

Dörfel, R. P., Arenas-Gomez, J. M., Fisher, P. M., Ganz, M., Knudsen, G. M., Svensson, J., & Plaven-Sigray, P. (2023). Prediction of brain age using structural magnetic resonance imaging: A comparison of accuracy and test-retest reliability of publicly available software packages. *BioRxiv*, 2023-01. https://doi.org/10.1101/2023.01.26.525514 bioRxiv

Espeseth, T., Christoforou, A., Lundervold, A. J., Steen, V. M., Le Hellard, S., & Reinvang, I. (2012). Imaging and cognitive genetics: The Norwegian Cognitive NeuroGenetics sample. *Twin Research and Human Genetics*, 15(3), 442–452. https://doi.org/10.1017/thg.2012.8

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS ONE*, 12(9), e0184661. https://doi.org/10.1371/journal.pone.0184661

Falkovskiy, P., Maréchal, B., Yan, S., Jin, Z., Quian, T., O'Brien, K., & Roche, A. (2016). Quantitative comparison of MP2RAGE skull-stripping strategies. In *ISMRM 2016, ISMRM 24rd Annual Meeting & Exhibition, SMRT 25th Annual Meeting*.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, 62(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Franke, K., & Gaser, C. (2019). Ten years of BrainAGE as a neuroimaging biomarker of brain aging: What insights have we gained? *Frontiers in Neurology*, 10, 789. https://doi.org/10.3389/fneur.2019.00789

Gong, W., Beckmann, C. F., Vedaldi, A., Smith, S. M., & Peng, H. (2021). Optimising a simple fully convolutional network for accurate brain age prediction in the pac 2019 challenge. *Frontiers in Psychiatry*, 12, 627996. https://doi.org/10.3389/fpsyt.2021.627996

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *Neuroimage*, 62(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jirsaraie, R. J., Kaufmann, T., Bashyam, V., Erus, G., Luby, J. L., Westlye, L. T., Davatzikos, C., Brach, D. M., & Sotiras, A. (2022). Benchmarking the generalizability of brain age models: Challenges posed by scanner variance and prediction bias. *Human Brain Mapping*, 44, 1118–1128. https://doi.org/10.1002/hbm.26144

Kaufmann, T., Van Der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., Alnæs, D., Barch, D. M., Baur-Streubel, R., Bertolino, A., Bettella, F., Beyer, M. K., Bøen, E., Borgwardt, S., Brandt, C. L., Buitelaar, J., Celius, E. G., Cervenka, S., Conzelmann, A., … Westlye, L. T. (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nature Neuroscience*, 22(10), 1617–1623. https://doi.org/10.1038/s41593-019-0471-7

Korbmacher, M., De Lange, A. M., Van Der Meer, D., Beck, D., Eikefjord, E., Lundervold, A., Andreassen, O. A., Westlye, L. T., & Maximov, I. I. (2023). Brain-wide associations between white matter and age highlight the role

of fornix microstructure in brain ageing. *Human Brain Mapping*, *44*(10), 4101–4119. https://doi.org/10.1002/hbm.26333

Korbmacher, M., Gurholt, T. P., De Lange, A.-M. G., Van Der Meer, D., Beck, D., Eikefjord, E., Lundervold, A., Andreassen, O. A., Westlye, L. T., & Maximov, I. I. (2022). Bio-psycho-social factors' associations with brain age: A large-scale UK Biobank diffusion study of 35,749 participants. *Frontiers in Psychology*, *14*, 1117732. https://doi.org/10.3389/fpsyg.2023.1117732

Leonardsen, E. H., Peng, H., Kaufmann, T., Agartz, I., Andreassen, O. A., Celius, E. G., Espeseth, T., Harbo, H. F., Høgestøl, E. A., Lange, A.-M. D., Marquand, A. F., Vidal-Piñeiro, D., Roe, J. M., Selbæk, G., Sørensen, Ø., Smith, S. M., Westlye, L. T., Wolfers, T., & Wang, Y. (2022). Deep neural networks learn general and clinically relevant representations of the ageing brain. *Neuroimage*, *256*, 119210. https://doi.org/10.1016/j.neuroimage.2022.119210

Man, M. Y., Ong, M. S., Mohamad, M. S., Deris, S., Sulong, G., Yunus, J., & Harun, F. K. C. (2015). A review on the bioinformatics tools for neuroimaging. *The Malaysian Journal of Medical Science*, *22*, 9..

Marquand, A. F., Kia, S. M., Zabihi, M., Wolfers, T., Buitelaar, J. K., & Beckmann, C. F. (2019). Conceptualizing mental disorders as deviations from normative functioning. *Molecular Psychiatry*, *24*(10), 1415–1424. https://doi.org/10.1038/s41380-019-0441-1

Opfer, R., Krüger, J., Spies, L., Ostwaldt, A.-C., Kitzler, H H., Schippling, S., & Buchert, R. (2022). Automatic segmentation of the thalamus using a massively trained 3D convolutional neural network: Higher sensitivity for the detection of reduced thalamus volume by improved inter-scanner stability. *European Radiology*, *33*, 1852–1861. https://doi.org/10.1007/s00330-022-09170-y

Peng, H., Gong, W., Beckmann, C. F., Vedaldi, A., & Smith, S. M. (2021). Accurate brain age prediction with lightweight deep neural networks. *Medical Image Analysis*, *68*, 101871. https://doi.org/10.1016/j.media.2020.101871

Rosen, A. F. G., Roalf, D. R., Ruparel, K., Blake, J., Seelaus, K., Villa, L. P., Ciric, R., Cook, P. A., Davatzikos, C., Elliott, M. A., Garcia De La Garza, A., Gennatas, E D., Quarmley, M., Schmitt, J. E, Shinohara, R. T., Tisdall, M. D, Craddock, R. C., Gur, R. E., Gur, R. C., & Satterthwaite, T. D. (2018). Quantitative assessment of structural image quality. *Neuroimage*, *169*, 407–418. https://doi.org/10.1016/j.neuroimage.2017.12.059

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M, & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Tønnesen, S., Kaufmann, T., Doan, N. T., Alnæs, D., Córdova-Palomera, A., Meer, D. V. D., Rokicki, J., Moberget, T., Gurholt, T. P., Haukvik, U. K., Ueland, T., Lagerberg, T. V., Agartz, I., Andreassen, O. A., & Westlye, L. T. (2018). White matter aberrations and age-related trajectories in patients with schizophrenia and bipolar disorder revealed by diffusion tensor imaging. *Scientific Reports*, *8*(1), 14129. https://doi.org/10.1038/s41598-018-32355-9

Vidal-Pineiro, D., Wang, Y., Krogsrud, S. K., Amlien, I. K., Baaré, W. F., Bartres-Faz, D., Bertram, L., Brandmaier, A. M., Drevon, C. A., Düzel, S., Ebmeier, K., Henson, R. N., Junqué, C., Kievit, R. A., Kühn, S., Leonardsen, E., Lindenberger, U., Madsen, K. S., Magnussen, F., … Fjell, A. (2021). Individual variations in 'brain age'relate to early-life factors more than to longitudinal brain change. *Elife*, *10*, e69995. https://doi.org/10.7554/eLife.69995

Wagen, A. Z., Coath, W., Keshavan, A., James, S.-N., Parker, T. D., Lane, C. A., Buchanan, S. M., Keuss, S. E., Storey, M., Lu, K., Macdougall, A., Murray-Smith, H., Freiberger, T., Cash, D. M., Malone, I. B., Barnes, J., Sudre, C. H., Wong, A., Pavisic, I. M., … Schott, J. M. (2022). Life course, genetic, and neuropathological associations with brain age in the 1946 British Birth Cohort: A population-based study. *The Lancet Healthy Longevity*, *3*(9), e607–e616. https://doi.org/10.1016/S2666-7568(22)00167-2

Wang, M.-Y., Korbmacher, M., Eikeland, R., & Specht, K. (2022). Deep brain imaging of three participants across 1 year: The Bergen breakfast scanning club project. *Frontiers in Human Neuroscience*, *16*, 1021503. https://doi.org/10.3389/2Ffnhum.2022.1021503

Wang, M. Y., Korbmacher, M., Eikeland, R., & Specht, K. (2023). The Bergen Breakfast Scanning Club dataset: A deep brain imaging dataset. *BioRxiv*, https://doi.org/10.1101/2023.05.30.542072 bioRxiv

Waters, A. B., Mace, R. A., Sawyer, K. S., & Gansler, D. A. (2019). Identifying errors in Freesurfer automated skull stripping and the incremental utility of manual intervention. *Brain Imaging and Behavior*, *13*, 1281–1291. https://doi.org/10.1007/s11682-018-9951-8

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Korbmacher, M., Wang, M.-Y., Eikeland, R., Buchert, R., Andreassen, O. A., Espeseth, T., Leonardsen, E., Westlye, L. T., Maximov, I. I., & Specht, K. (2023). Considerations on brain age predictions from repeatedly sampled data across time. *Brain and Behavior*, e3219. https://doi.org/10.1002/brb3.3219

BRAIN ASYMMETRIES FROM MID- TO LATE LIFE AND HEMISPHERIC BRAIN AGE

M. Korbmacher, A. M. de Lange, D. van der Meer, D. Beck, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov

TITLE
Brain asymmetries from mid- to late life and hemispheric brain age

AUTHORS

Max Korbmacher[1,2,3*], Dennis van der Meer[2,4], Dani Beck[2,5,6], Ann-Marie de Lange[2,7,8], Eli Eikefjord[1,3], Arvid Lundervold[1,3,9,10], Ole A. Andreassen[2,11], Lars T. Westlye[2,6,11], Ivan I. Maximov[1,2*]

[1]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway.
[2]NORMENT Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway.
[3]Mohn Medical Imaging and Visualization Centre (MMIV), Bergen, Norway.
[8]Department of Psychiatry, University of Oxford, Oxford, UK.
[7]LREN, Centre for Research in Neurosciences - Department of Clinical Neurosciences, CHUV and University of Lausanne, Lausanne, Switzerland.
[4]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands.
[5]Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, Norway.
[6]Department of Psychology, University of Oslo, Oslo, Norway.
[9]Department of Radiology, Haukeland University Hospital, Bergen, Norway.
[10]Department of Biomedicine, University of Bergen, Bergen, Norway.
[11]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway.

*Corresponding authors. E-mails: max.korbmacher@hvl.no; ivan.maximov@hvl.no;

1

ABSTRACT

The human brain demonstrates structural and functional asymmetries which have implications for ageing and mental and neurological disease development. We used a set of magnetic resonance imaging (MRI) metrics derived from structural and diffusion MRI data in $N{=}48{,}040$ UK Biobank participants to evaluate age-related differences in brain asymmetry. Most regional grey and white matter metrics presented asymmetry, appearing higher at a higher age. Informed by these results, we conducted *hemispheric brain age* (HBA) predictions from left/right multi-modal MRI metrics. HBA was concordant to conventional brain age predictions, using metrics from both hemispheres, but offers a supplemental general marker of brain asymmetry when setting left/right HBA into relationship with each other. In contrast to WM brain asymmetries, left/right discrepancies in HBA are lower at higher ages. The findings emphasise the value of further investigating the role of brain asymmetries in brain ageing and disease development.

1 INTRODUCTION

There are various structural and functional differences in brain architecture between the left and right hemispheres[1–6]. Microstructural brain characteristics, such as white matter (WM) pathways or intra- and extra-neurite water organisation, might underlie the brain's functional lateralisation[7]. Functional network difference has been associated with handedness[8]. Both structural and functional brain asymmetry exhibit clinical importance as there are differences in brain asymmetry between healthy controls and various disease groups, including neurodegenerative diseases such as Alzheimer's disease[9, 10], Parkinson's disease[11], and psychiatric disease such as obsessive–compulsive disorder[4, 12, 13] and schizophrenia[14]. In that context and particularly relevant from a lifespan-perspective, cortical thickness asymmetry decreases throughout ageing, with this alteration being potentially accelerated in the development of neurodegenerative disorders such as Alzheimer's Disease[9]. Similarly, some studies suggest lower WM microstructure asymmetry at higher ages, indicated by intra-axonal water fraction[15], fractional anisotropy, or the apparent diffusion coefficient[16]. Additional investigations into brain asymmetries' age-dependencies can provide a more comprehensive understanding of the influence of asymmetries on ageing and disease development.

Brain age is a developing integrative marker of brain health, particularly sensitive to neurodegenerative diseases[17, 18]. Brain age refers to the predicted age in contrast to chronological age and is based on a set of scalar metrics derived from brain scans such as MR. To date, brain age has often been estimated using a global brain parametrisation such as the averaged scalar measures over particular anatomical regions or the whole brain[17–21]. Hence, we refer to these whole-brain age predictions as global brain age (GBA). However, while brain age has been calculated for different brain regions[18, 22–24], the use of hemisphere-specific data is usually not being considered as a potential source of additional information. Yet, one study presents hemisphere-specific

2

and region-specific brain ages containing useful clinical information about post-stroke cognitive improvement[22].

Previous results show that brain age prediction depends on the specific features used[25–27], rendering for example modality as important. Yet, the influence of hemispheric differences or brain asymmetry on the age predictions remains unclear. However, previously outlined brain asymmetries[1–6] might be informative for age predictions. One way of leveraging brain asymmetries into simple metrics is to estimate separate brain ages for each hemisphere (HBA) and to then compare the estimates. It remains unclear whether predictions from a single hemisphere lead to less accurate predictions due to the inclusion of less data and a potential attenuation of noise. At the same time, in the case of diffusion MRI (dMRI), different model-based diffusion features yield highly concordant brain age predictions, also when varying the number of included features[21]. Finally, although the evidence is mixed on the influence of handedness on brain asymmetry[28–31], differences in handedness are potentially reflected in brain structure, which would in turn influence age predictions differently when obtained from the left or right hemisphere only. Hence, handedness requires further examination as potential confounding effect when assessing asymmetry.

HBA, a new brain age measure, may propose more sensitive brain health markers than GBA, as age predictions can be compared between hemispheres to infer the integrity of each hemisphere and give a general estimate of brain asymmetry. Brain asymmetries are commonly observed using the Laterality Index (LI)[32]. However, different ways of estimating asymmetry can introduce variability in its dependency with age[33], and covariates of brain age require further investigation[34, 35]. To extend the existing brain age conceptualisation of using features across the whole brain and to maximise interpretability, we restrict brain age predictions to region-averaged and global features and not asymmetries of these features. Additionally, differences in the models' abilities to predict age from WM microstructure features derived from dMRI compared to T_1-weighted features (volume, surface area, thickness) need to be ruled out in order to validate both GBA and HBA.

Hence, in the present work, we tested first the preregistered hypotheses (written study and analysis plan prior data inspection and analyses[36, 37]) that the GBA and HBA depend on the used MRI modality (Hypothesis 1), disentangling whether the different grey matter (GM) and WM metrics and the degree of their asymmetry influences brain age predictions. We furthermore tested whether there was an effect of hemisphere (Hypothesis 2) and handedness (Hypothesis 3) on brain age predictions. Exploratory analyses included (a) revealing hemispheric differences between GM and WM features, (b) examining LI associations with age, including the LI of the brain features as well as left and right brain ages, and (c) testing the consistency of brain age-covariate associations (specifically, health-and-lifestyle factors, as these were previously associated with brain age[20, 26, 38–41]).

3

RESULTS

Hemispheric differences and age sensitivity for GM and WM features

Two-tailed paired samples t-tests showed that a significant proportion of the GM and WM features differed between hemispheres with medium effect sizes. Among the significant 793 of 840 dMRI feature asymmetries (94.4%, $p<.05$, with Cohen's $|\bar{d}_{dMRI}|$ = 0.57±0.44). The largest differences were found for DTI FA in the inferior longitudinal fasciculus ($d = 3.64$), and cingulum ($d = 1.95$), and for AD in superior longitudinal fasciculus.

Effects sizes of the significant hemispheric differences of the 115 of 117 T_1-weighted features (98.3%), were similar: mean $|\bar{d}_{T_1}| = 0.53$±0.41, and the largest asymmetries were found for the surface area of the transverse-temporal region ($d = 1.81$), frontal pole ($d = 1.76$), and pars orbitalis ($d = 1.74$; see Supplementary Table 10 for T_1-weighted and dMRI features with strongest hemispheric differences).

LRTs comparing a baseline model predicting age from sex and scanner site compared to a model where the respective smooth of the metric was added (Eq. 3 and 4) indicated most features as age-sensitive (231 of the 234 (98.72%) of the T_1-weighted features; 1601 of the 1680 (95.53%) dMRI features). Age-sensitivity was strongly expressed in both significant T_1-weighted features ($\bar{F}_{T_1} = 1{,}168.90$±993.59), as well as significant dMRI metrics ($\bar{F}_{dMRI} = 1{,}208.97$±943.52) with strongest age-sensitivity observed for left superior temporal thickness, left/right overall thickness, left/right hippocampus volume, and right inferior parietal thickness and multiple WMM metrics in the right anterior limb of the internal capsule, the left/right fornix-striaterminalis pathway, left/right anterior corona radiata and inferior fronto-occipital fasciculus ($F > 3{,}000$; for top features see Supplementary Table 2).

Results were similar when comparing linear models to the baseline model (Eq. 2 and 4): 1448 of the 1680 (86.19%) dMRI metrics, and 228 of the 234 (97.44%) of the T_1-weighted features were age-sensitive ($\bar{F}_{T_1} = 3{,}426.89$±2,947.11, $\bar{F}_{dMRI} = 2{,}378.46$±2,357.80), with the features with the strongest age-sensitivity resembling LRT results of non-linear models (for top features see Supplementary Table 3).

Considering only left/right averages identified only DTI-AD, and WMTI axial and radial extra-axonal diffusivity to not differ between hemispheres ($p > .05$). Furthermore, all features were age-sensitive when GAMs ($p < 3.4 \times 10^{-64}$; yet for linear models, BRIA-vCSF and WMTI-axEAD, as well as right DTI-AD and left WMTI-radEAD were not age sensitive (Supplementary Tables 4, 5). Furthermore, the age-relationships for most of the left/right averages were similar across hemispheres (Figure 1, both for crude and adjusted values: Supplementary Figure 1, and for linear and non-linear models: Supplementary Figure 4). However, differences in dMRI metrics were observed for the ends of the distribution including individuals aged younger than 55 ($N = 5{,}307$) and older than 75 ($N = 3{,}480$).

4

GM and WM feature asymmetry

Using LRTs comparing GAMs to a baseline model 53 (45.30%) of the 117 T_1-weighted and 733 of the 840 (87.26%) dMRI $|LI|$ features as age sensitive ($p<$.05). Using LRTs on linear effects identified 53 (45.30%) of the 117 T_1-weighted and 678 of the 840 (80.71%) dMRI $|LI|$ features as age sensitive ($p<$.05).

In the following we constrain analyses to linear models and present partial derivatives / slopes as a measure of effect size, allowing for simple comparisons across age-relationships as model fit indices AIC and BIC of linear models and GAMs suggested on average no differences across both T_1-weighted ($p_{adj\ AIC} = .759$; $p_{adj\ BIC} = 1$) and diffusion-weighted features ($d_{AIC} = 0.510, p_{adj\ AIC} = .020$; $p_{adj\ BIC} = .126$).

The absolute feature asymmetries were higher later in life ($\bar{\beta}_{dMRI} = 0.05 \pm 0.07$; $\bar{\beta}_{T_1} = 0.03 \pm 0.06$, $|\bar{\beta}_{multimodal}| = 0.05 \pm 0.07$, here only $p_{adj} < .05$ selected; Supplementary Figure 2)-3).

The strongest adjusted relationships between the respective features' asymmetries and age were found for dMRI metrics ($|\bar{\beta}_{dMRI}| = 0.08 \pm 0.05$, $|\bar{\beta}_{T_1}| = 0.05 \pm 0.03$; Figure 2), particularly outlining asymmetry increases in the cingulate gyrus ($\beta_{BRIA-microRD} = 0.25$, $\beta_{BRIA-microFA} = 0.22$, $\beta_{DTI-MD} = 0.20$, $\beta_{BRIA-microADC} = 0.19$), and decrease in the cerebral peduncle ($\beta_{SMTmc-extratrans} = -0.20$, $\beta_{SMT-trans} = -0.19$, $\beta_{BRIA-Vextra} = -0.14$) and superior longitudinal temporal fasciculus ($\beta_{BRIA-microAX} = -0.17$, $\beta_{SMT-long} = -0.17$, $\beta_{BRIA-DAXextra} = -0.16$).

For T_1-weighted metrics, larger structures' $|LI|$ were most sensitive to age, with the strongest negative associations including the inferior lateral ($\beta = -0.16$) and lateral ventricles ($\beta = -0.09$), pallidum ($\beta = -0.11$) volumes, rostro-middle thickness ($\beta = -0.11$), thalamus volume ($\beta = -0.07$) and enthorinal area ($\beta = -0.05$). Largest positive age-associations were were shown for accumbens area ($\beta = 0.13$), WM surface area ($\beta = 0.13$) and volume ($\beta = 0.11$), amygdala ($\beta = 0.11$), caudal anterior cingulate thicknes ($\beta = 0.11$), cortex volume ($\beta = 0.10$), caudate volume ($\beta = 0.10$), and cerebellar WM volume ($\beta = 0.09$), in addition to several temporal and limbic areas (Figure 2).

No influences of handedness, but sex-specific differences in the influence of hemisphere and modality on brain age estimates

Model performance metrics indicated that most accurately age predictions were accomplished using multimodal MRI data based on left, right, and both hemispheres (Table 1), with obtained HBA and GBA being strongly correlated with each other for similar models (Figure 3).

LMERs did not indicate a difference between modalities (Hypothesis 1) when comparing brain ages estimated from dMRI to multimodal MRI ($p = .623$), and dMRI to T_1-weighted MRI ($p = .452$). There were also no differences in brain age estimates between hemispheres ($p = .413$, Hypothesis 2). Moreover, LRTs indicated no significant difference between models when adding handedness ($\chi^2 = 4.19$, $p = .123$, $df = 2$) or handedness-hemisphere interaction and handedness ($\chi^2 = 7.32$, $p = .120$, $df = 4$; see Eqs. 5-6). Notably, a follow-up analysis adding the sex-hemisphere and sex-modality interaction terms indicates higher right than left brain age in females

5

compared to male higher left than right brain age ($\beta = 0.22$ $years, p < .001$), as well as sex differences in brain age estimates comparing diffusion to T$_1$-weighted ($\beta = -0.89 years, p < .001$) or multimodal MRI ($\beta = -0.54$ $years, p < .001$).

However, sex-based stratification for males showed lower multimodal ($\beta = 0.23$ years, $p < .001$), and T$_1$-weighted ($\beta = 0.41$ years, $p < .001$) than diffusion MRI derived brain age, and higher right than left brain age ($\beta = -0.09$ years, $p < .001$). Females showed opposite effects of higher multimodal ($\beta = -0.28$ years, $p < .001$), and T$_1$-weighted ($\beta = -0.49$ years, $p < .001$) than diffusion MRI derived brain age, and lower right than left brain age ($\beta = 0.13$ years, $p < .001$). Neither handedness nor the handedness-hemisphere interactions were significant for either sex in LMERs and LRTs ($p > .05$).

Lower brain age asymmetry at higher ages

To test whether asymmetries between hemisphere-specific brain age predictions are lower at higher age, $|LI_{HBA}|$, was associated with age. $|LI_{HBA}|$ showed negative unadjusted associations with age for T$_1$-weighted ($r = -0.069$, $p < .001$), dMRI ($r = -0.121$, $p < .001$), and multimodal models ($r = -0.121$, $p < .001$). The associations were similar when using LMEs adjusting for sex and the random intercept site (T$_1$-weighted: $\beta = -0.069$, $p < .001$, dMRI: $\beta = -0.115$, $p < .001$, multimodal: $\beta = -0.117, p < .001$). LRTs indicate the age-sensitivity of LI_{HBA} (T$_1$-weighted: $\chi^2 = 173.42$, $p < .001$, dMRI: $\chi^2 = 488.74$, $p < .001$, multimodal: $\chi^2 = 506.08$, $p < .001$).

These results held also true when stratifying by sex, for unadjusted ($r_{dMRI\ males} = -0.134$, $r_{dMRI\ females} = -0.104, r_{T1\ males} = -0.134$, $r_{T1\ females} = -0.048$, $r_{multimodal\ males} = -0.134$, $r_{multimodal\ females} = -0.111$), and adjusted associations ($\beta_{dMRI\ males} = -0.134$, $\beta_{dMRI\ females} = -0.099$, $\beta_{T1\ males} = -0.134$, $\beta_{T1\ females} = -0.045$, $\beta_{multimodal\ males} = -0.134$, $\beta_{multimodal\ females} = -0.106$), with χ^2 tests suggesting age sensitivity (all $p < .001$).

HBA and GBA and health-and-lifestyle factors

We further investigated the pattern of relationships with general health-and-lifestyle phenotypes across HBAs (Figure 4). Relationships between brain ages from single and both hemispheres were similar within modalities, but varied slightly between modalities (Figure 4). These results were robust to sex stratifications. Yet, while males' brain age was sensitive to high cholesterol, hip circumference, smoking and weight, this was not the case for females' brain age (Supplementary Figure 7).

Sex stratified hemispheric differences and age sensitivity for GM and WM features

For further insights into sex differences, we repeated the presented analyses on hemispheric differences and features' age-sensitivity stratifying by sex. Two-tailed paired samples t-tests assessing regional differences between hemispheres showed similar results between sexes, which are also comparable to cross-sex results. Most features differed between hemispheres for both males and females (T$_1$-weighted: 98.3% for both

6

sexes, dMRI$_{males}$: 96%, dMRI$_{females}$: 95%), and effect sizes were similar ($|\bar{d}_{T_1males}| =$ 0.54 $\pm$ 0.42, $|\bar{d}_{T_1females}| =$ 0.53 $\pm$ 0.42, $|\bar{d}_{dMRI~males}| =$ 0.57 $\pm$ 0.41, $|\bar{d}_{dMRI~females}|$ = 0.60 $\pm$ 0.47).

Also the strongest effects were similar across sexes: strongest differences in T$_1$-weighted features in males were observed for frontal pole ($d_{T_1males} = 1.82$) and pars orbitalis ($d_{T_1males} = 1.78$) surface area, and for females in the area of the transverse temporal area ($d_{T_1females} = 1.89$) and the frontal pole ($d_{T_1females} = 1.73$). Strongest WM differences were observed for both sexes in inferior longitudinal fasciculus ($d_{dMRI~males} = 3.44$, $d_{dMRI~females} = 3.91$), and superior lonitudinal temporal fasciculus ($d_{dMRI~males} = 2.09$, $d_{dMRI~females} = 2.40$; Supplementary Table 6).

LRTs comparing a baseline model predicting age from sex and scanner site compared to a model where the respective smooth of the metric was added (Eq. 3 and 4) indicated most features as age-sensitive (230 of the 234 (98.29%) of the T$_1$-weighted features (both sexes); 1,557 and 1564 of the 1,680 (92.68% and 93.10%) dMRI features for males and females, respective). Age-sensitivity was strongly expressed in both significant T$_1$-weighted features ($\bar{F}_{T_1males} = 640.80 \pm 521.33$; $\bar{F}_{T_1females} = 578.61 \pm 500.79$), as well as significant dMRI metrics ($\bar{F}_{dMRI~males} = 586.38 \pm 450.68$, $\bar{F}_{dMRI~females} = 674.61 \pm 499.58$).

Similar to the results including both sexes, the strongest T$_1$-weighted feature age-sensitivity was observed for left superior temporal thickness, left/right hippocampus volume for both sexes, and right inferior parietal thickness only for females. Concerning dMRI features, sex stratification reflects the findings accounting for sex, outlining the fornix-striaterminalis pathway, anterior corona radiata and inferior fronto-occipital fasciculus, yet adding the anterior limb of the internal capsule and the anterior thalamic radiation. Unique to non-linear models, also the lateral ventricle volume was lined out as highly age sensitive (all $F > 1,666$; for top features see Supplementary Table 7.

Results were similar when comparing linear models to the baseline model (Eq. 2 and 4): 1,557 and 1,564 of the 1680 (92.68%, 93.01%) dMRI metrics, and 226 and 224 of the 234 (96.58%, 95.73%) of the T$_1$-weighted features were age-sensitive for males and females, respectively ($\bar{F}_{T_1males} = 1,767.60 \pm 1,474.69$; $\bar{F}_{T_1females} = 1,712.73 \pm 1,488.97$; $\bar{F}_{dMRI~males} = 1,198.85 \pm 1,135.84$, $\bar{F}_{dMRI~females} = 1,297.51 \pm 1,257.02$), with the features with the strongest age-sensitivity resembling LRT results of non-linear models (for top features see Supplementary Table 8). Considering only left and right hemispheric averages, t-tests indicated that all features differed between hemispheres for males ($p < 3.1 \times 10^{-9}$). In females, WMTI radEAD and axEAD as well as DTI AD did not differ between hemispheres ($p > 0.05$), but all other metrics differing between hemispheres ($p < 1.5 \times 10^{-36}$).

Considering all regional features, LRTs on GAMs (Eq. 4, 3) indicated that all features were age-sensitive ($p < 5.1 \times 10^{-71}$). LRTs on linear models (Eq. 2, 4) indicated that right hemisphere BRIA-vCSF and left microRD were not age sensitive ($p_{adj} > 0.05$) in males. In females, additionally, left DTI-RD and GM thickness as well as left and right WMTI-axEAD were not age-sensitive. All other metrics were age sensitive ($p < 2.7 \times 10^{-11}$). Hemispheric features' age-relationships showed similar intercepts and slopes across sexes, except DTI-AD, WMTI-radEAD and WMTI-axEAD (Supplementary Figure 4-5).

7

Sex differences in GM and WM feature asymmetry

Sex-stratified analyses indicate most dMRI $|LI|$ features to be age sensitive (dMRI$_{males}$ = 64.29%, dMRI$_{females}$ = 69.52%), but less T_1-weighted features ($T_{1\ males}$ = 47.86%, $T_{1\ females}$ = 38.46%) when using non-linear models. Linear models showed similar results (dMRI$_{males}$ = 60.95%, dMRI$_{females}$ = 64.05%; $T_{1\ males}$ = 44.44%, $T_{1\ females}$ = 37.61%). Comparing linear to non-linear models using paired samples t-tests suggests no differences model fit indicated in AIC or BIC scores for both males and females in T_1-weighted and diffusion features' asymmetry ($p > 0.05$). Hence, linear model outcomes are presented below. Similar to models including both sexes, when stratifying for sex, $|LI|$ for diffusion and T_1-weighted feature were positively associated with age ($\bar{\beta}_{dMRI\ male} = 0.05 \pm 0.08$, $\bar{\beta}_{dMRI\ female} = 0.05 \pm 0.08$, $\bar{\beta}_{T1\ male} = 0.03 \pm 0.06$, $\bar{\beta}_{T1\ female} = 0.03 \pm 0.06$).

The strongest adjusted relationships for diffusion features were found in the cingulate gyrus tract ($\beta_{males\ BRIA-microRD} = 0.25$, $\beta_{males\ BRIA-microFA} = 0.22$, $\beta_{females\ BRIA-microRD} = 0.25$, $\beta_{males\ BRIA-microFA} = 0.21$) and in the cerebral peduncle ($\beta_{males\ SMTmc-extratrans} = -0.19$, $\beta_{males\ SMT-trans} = -0.18$, $\beta_{females\ SMTmc-extratrans} = -0.21$, $\beta_{females\ SMT-trans} = -0.20$, $\beta_{females\ BRIA-Vextra} = -0.18$). Strongest age associations with T_1-weighted asymmetries were found for the area of the accumbens ($\beta_{males} = 0.14, \beta_{females} = 0.12$) and WM surface ($\beta_{males} = 0.13, \beta_{females} = 0.12$), with strongest inverse relationships observed for inferior lateral ventricles ($\beta_{males} = -0.17$, $\beta_{females} = -0.14$) and pallidum ($\beta_{males} = -0.11$, $\beta_{females} = -0.12$).

DISCUSSION

In the present work we investigated a new way of utilizing brain age to differentiate between hemispheres, and performed a detailed assessment of brain asymmetry associations with age. As a baseline, we showed that most grey and white matter features were age-sensitive and differed between hemispheres with relatively large effect sizes. Brain asymmetry was age-sensitive, and overall higher at higher ages. In contrast, asymmetry in hemispheric brain age was lower at higher ages. The strongest relationship of age and absolute brain asymmetry was identified in larger GM and WM regions, as well subcortical structures, including the limbic system, the ventricles, cingulate and cerebral peduncle WM.

Brain age predictions exhibited concordant accuracy within modalities for left, right, and both hemispheres, and concordant associations with health-and-lifestyle factors also when sex-stratifying. The predictions did not differ statistically between hemispheres, modalities, or handedness groups when considering both sexes together. However, sex-stratified analyses revealed significant opposing effects between sexes for hemisphere and modality but not handedness. There are multiple reasons for the observed higher brain age in females' right hemisphere compared to males' higher brain age of the left hemisphere, in addition to modality-specific differences. First, male and female brain structure differs, resulting in sex-specific regional variations in brain age estimates[42]. Second, body and brain ageing trajectories differ between sexes, for example, outlined by sex-dependent importance of cardiometabolic risk

8

factors[43].Hence, the tendency of males' predicted brain age being lower using T_1-weighted and multimodal in contrast to diffusion-derived brain ages, with these trends reversed in females, might also reflect stronger brain age associations with cardiometabolic risk factors in males (Supplementary Figure 7), which have been demonstrated earlier for WM features and WM brain age [38, 39]. In that sense, HBA allows to assess the structural integrity of each hemisphere individually, and to set brain ages from the two hemispheres in relationship to each other providing a general marker of asymmetry. Despite brain asymmetries overall increasing, the asymmetries between left/right HBA were smaller at a higher age. At higher ages, both hemispheres might hence become overall more comparable, despite ageing-related changes.

We found that the majority of regional and hemisphere-averaged MRI features differed between hemispheres. Both features and asymmetries were age-sensitive indicating that the investigation of asymmetries are useful across ages and MRI modalities.

Interestingly, hemisphere-averaged features' age-associations and HBA of the same modality were similar between hemispheres (Figure 1), and the hemisphere was not a significant predictor of brain age estimated from a particular hemisphere, when analysing data from both sexes together. However, when sex-stratifying, modality and hemisphere were significant predictors, suggesting that HBA captures both brain asymmetries as well as biological sex-differences which become apparent when using multimodal MRI. These results outline the importance of considering sex-differences in brain age analyses.

Several studies present evidence for asymmetries in WM[6, 44–47] and GM[4, 9, 48–50]. In contrast to these previous studies, for the first time, we examine various metrics supplying information on both WM and GM in a large sample. While we find various differences between hemispheres, age relationships of T_1-weighted and dMRI features were similar between hemispheres using hemispheric averages, also when stratifying by sex. Spatially finer-grained examinations revealed more specific patterns of asymmetry in T_1-weighted features, such as GM thickness[9], and dMRI features[44]. This is also shown in the present study by stronger age-effects for specific regional asymmetries compared to asymmetries in hemispheric averages. Age-MRI metric relationships depend, however, on the selected metric, the sample, and the sampling (cross-sectional or longitudinal)[51, 52]. For example, previous evidence from T_1-weighted MRI indicates no differences in GM volume between hemispheres[53], but hemispheric differences of cortical thickness and surface area across ageing[4, 9].

The presented age charts of MRI metrics in the current work (Figure 1, Supplementary Figure 1) provide similar trends to those reported in previous studies observing global age dependencies[19, 21, 54–56]. Yet, the stratification between hemispheres when presenting brain features' age dependence is a novel way of presenting brain charts.

We found asymmetries based on GM and WM brain scalar measures. Unimodal studies with smaller, younger samples presented age-dependence of the brain asymmetry during early WM development[47] and adult cortical thickness[9], other T_1-derived metrics[33], and functional network development[5], showed lower asymmetry at higher

9

ages. In contrast to HBA asymmetries, brain asymmetries do generally not support the notion of lower but instead of higher brain asymmetry later in life. Different study design choices, such as temporal and spatial levels might provide supplemental information into the age-dependence of brain asymmetries, for example, by further investigating longitudinal and voxel-level asymmetries.

We extended previous findings by providing a comprehensive overview of brain asymmetry associations throughout mid- to late life including both GM and WM. Our findings indicate that when considering various metrics, older brains generally appear less symmetric than younger brains in the current sample mid- to late life sample, whereas brain age appears more symmetric in older brains.

Notably, we identified strong associations between specific brain regions' asymmetry and age. The strongest age-associations of asymmetries were observed for subcortical, ventricle-near structures. The general age-sensitivity of such structures[21, 57, 58] might be a reason for the observed age-associations in asymmetries, and hence pointing towards one hemisphere being stronger affected by degradation effects, or even the involvement of such regions in psychiatric and neurodegenerative disorders[40, 54, 57, 59–64]. For example, the hippocampus, a prominent limbic structure, presents relatively high levels of adult neurogenesis, which might potentially explain repeated findings of the region's associations with psychiatric disorders and disorder states such as depression, anxiety, schizophrenia, addiction, and psychosis[65, 66], and neurdegenerative disorders, especially Alzheimer's Disease[67], but also ageing in general[68]. Some of the strongest age-relationship for T_1-derived asymmetries were observed in accumbens, ventricles and pallidum. In turn, a series of dMRI approaches was sensitive to asymmetry in the cingulum tract, which is higher in late-life and peduncle asymmetry which appears lower in late-life. In particular radial diffusivity metrics, such SMT-trans, SMTmc-extratrans, and BRIA-microRd, and fractional anisotropy indicated by BRIA-microFA were sensitive to age-dependencies of these asymmetries. Although speculative, this observation could indicate a relationship between asymmetry and axonal properties during ageing, such as myelination, density, or diameter, in the cingulum, with yet a more general marker (BRIA-microFA) of anisotropy asymmetry increasing at advanced age. However, limitations of the different diffusion metrics, such as the inability to account for axonal swelling, infection, or crossing fibres[69], aggravates the interpretation of such asymmetry changes. Overall, asymmetries' age-dependencies in subcortical, limbic and ventricle-near areas are not surprising, considering that the cingulum and cerebral peduncle WM, and middle temporal GM area also presented some of the strongest asymmetries across the sample (Supplementary Table 10).

Both GM volume, surface, and thickness show asymmetries across studies[1, 3, 4, 9, 53]. We identified the lowest degree of asymmetry linked to higher ages in the ventricular and pallidum volumes. The strongest positive age-relationships for T_1-weighted features were observed for accumbens and WM surface area, as well as limbic structures such as amygdala, hippocampus, and cingulate. Limbic structures have previously been outlined as highly age-sensitive[21, 57, 58, 68]. Higher asymmetry-levels might speak to asymmetric atrophy in these limbic regions, potentially explaining several ageing-related effects [9]. However, lifespan changes in ventricular

10

volume asymmetry in relation to symptom and disorder expression requires additional investigations.

Cingulum WM microstructure has been reported to differ between hemispheres[70–72]. Abnormalities in cingulum asymmetry have been linked to schizophrenia[73–75] and epilepsy[76, 77], and Alzheimer's disease[58]. Additionally, the cingulum tract was associated with the anti-depressant effects of deep brain stimulation in treatment-resistant depression[78]. Recent evidence points out strongest polygenic risk associations for several psychiatric disorders in addition to Alzheimer's Disease with longitudinal WM in the cerebral peduncle [57]. Future research could assess regional asymmetries to evaluate such metrics' value for diagnostics and treatment in a range of brain disorders.

Overall, most absolute MRI feature asymmetries were positively related to age, with brain age asymmetries showing inverse age-relationships. However, for both WM and GM this process was observed to be spatially distributed. Metric-specific changes might indicate accelerated and pathological ageing [9], which urges to examine different WM and GM metrics across temporal and spatial resolutions and in clinical samples.

Informed by the presented brain asymmetries and their age-dependence, we suggest HBA, indicating the structural integrity of each hemisphere when compared to the chronological age. Moreover, HBA provides a general marker of asymmetry, when setting left/right HBA in relationship to each other. While this added information to conventional GBA is promising, first, the degree to which HBA captures GBA predictions, had to be assessed. This investigation included (1) direct comparisons of HBA and GBA models and their predictions, (2) the influence of covariates of brain age including MRI modality, hemisphere, handedness, and the hemisphere-handedness interaction effect, and (3) a comparison of health-and-lifestyle phenotype-associations with HBA and GBA. Overall, HBA and GBA were highly similar across these dimensions, yet different between hemispheres and modalities within males and females, with these differences contrasting each other. This renders HBA sensitive to potential underlying biological processes which only become apparent when assessing males and females separately. Additionally, different modalities might be sensitive to a range of biological phenomena in terms of brain age, such as dMRI brain age which is correlated with diabetes only in males. In that sense, a further route of investigation could be to establish sex-specific uni- and multimodal brain age models (which account for sex differences in brain morphology and its developmental trajectories). The influence of hemisphere and sex on how these models relate to biological phenomena can then be assessed.

Congruently with previous research which combined MRI modalities[27], we found higher prediction accuracy for multimodal compared to unimodal predictions for both HBA and GBA. Our results extend previous findings on conventional brain age by not only estimating brain age from different MRI modalities, but also for each hemisphere separately. HBA could hold potential in clinical samples by informing about the consistency between the two hemispheres' brain age predictions. Particularly diseases or conditions which affect a single hemisphere, such as unilateral stroke or trauma, might then be sensitively detected, and the integrity of the unaffected hemisphere can be assessed by observing the congruence of HBA[22]. Larger discrepancies between

11

HBAs of the same individual might act as a marker of hemisphere-specific brain health imbalance, which may indicate potential pathology.

While this study provides initial explorations of asymmetries and HBA, our findings remain limited to the examined sample (imaging subset of the UKB), and limited by generational effects within the sample. The UKB contains individuals born in different decades, which influences individual predispositions for brain health through various factors such as the living environment[79] or education[80], representing various potential confounding effects. Additional bias might have been introduced by the sample characteristics and sampling procedure. The UKB consists of nearly exclusively white UK citizens, limiting the generalisability beyond white Northern Europeans and US Americans in their midlife to late life. The volunteer-based sampling procedure might additionally have introduced bias, reducing generalisability to the UK population[81], with the imaging sample of the UKB showing an additional positive health bias (better physical and mental health) over the rest of the UKB sample[82], rendering this sub-sample as even less representative of the total UK population.

In conclusion, we identified asymmetries throughout the brain from midlife to late-life. These asymmetries appear higher later in life across GM and WM. The inverse age-relationships of brain asymmetry are reflected by the difference in left/right hemispheric brain age, which is smaller at higher ages. We furthermore identify various regional asymmetries which do not only show age-dependence but which have also been related to various clinical diagnoses. The identified age-relationships of asymmetries provide future opportunities to better understand ageing and disease development.

METHODS

Sample characteristics

We obtained UK Biobank (UKB) data[83], including $N = 48,040$ T_1-weighted datasets, $N = 39,637$ dMRI datasets, resulting in $N = 39,507$ joined/multimodal datasets after exclusions were applied. Participant data were excluded when consent had been withdrawn, an ICD-10 diagnosis from categories F (presence of mental and behavioural disorder), G (disease of the nervous system), I (disease of the circulatory system), or stroke was present, and when datasets were not meeting quality control standards using the YTTRIUM method[84] for dMRI datasets and Euler numbers were larger than 3 standard deviations below the mean for T_1-weighted data[85]. In brief, YTTRIUM[84] converts the dMRI scalar metric into 2D format using a structural similarity[86, 87] extension of each scalar map to their mean image in order to create a 2D distribution of image and diffusion parameters. These quality assessments are based on a 2-step clustering algorithm applied to identify subjects located outside of the main distribution.

Data were collected at four sites, with the T_1-weighted data collected in Cheadle (58.41%), Newcastle (25.97%), Reading (15.48%), and Bristol (0.14%). Of these data, 52.00% were females, and the participants age range was from 44.57 to 83.71, mean $= 64.86 \pm 7.77$, median $= 65.38 \pm 8.79$. DMRI data were available from four sites: Cheadle (57.76%), Newcastle (26.12%), Reading (15.98%), and Bristol (0.14), with

12

52.19% female, and an age range of 44.57 to 82.75, mean $= 64.63 \pm 7.70$, median $= 65.16 \pm 8.73$. The multimodal sample ($N = 39,507$) was 52.22% female, with an age range of 44.57 to 82.75, mean $= 64.62 \pm 7.70$, median $= 65.15 \pm 8.73$. Information on sex was acquired from the UK central registry at recruitment, but in some cases updated by the participant. Hence the sex variable may contain a mixture of the sex the UK National Health Service (NHS) had recorded for the participant as well as self-reported sex.

MRI acquisition and post-processing

UKB MRI data acquisition procedures are described elsewhere[83, 88, 89]. The raw T_1-weighted and dMRI data were processed accordingly. Namely, the dMRI data passed through an optimised pipeline[84]. The pipeline includes corrections for noise[90], Gibbs ringing[91], susceptibility-induced and motion distortions, and eddy current artifacts[92]. Isotropic 1 mm^3 Gaussian smoothing was carried out using FSL's[93, 94] *fslmaths*. Employing the multi-shell data, Diffusion Tensor Imaging (DTI)[95], Diffusion Kurtosis Imaging (DKI)[96] and White Matter Tract Integrity (WMTI)[97] metrics were estimated using Matlab 2017b code (https://github.com/NYU-DiffusionMRI/DESIGNER). Spherical mean technique (SMT)[98], and multi-compartment spherical mean technique (SMTmc)[99] metrics were estimated using original code (https://github.com/ekaden/smt)[98, 99]. Estimates from the Bayesian Rotational Invariant Approach (BRIA) were evaluated by the original Matlab code (https://bitbucket.org/reisert/baydiff/src/master/)[100].

T_1-weighted images were processed using Freesurfer (version 5.3)[101] automatic *recon-all* pipeline for cortical reconstruction and subcortical segmentation of the T_1-weighted images (http://surfer.nmr.mgh.harvard.edu/fswiki)[102].

In total, we obtained 28 WM metrics from six diffusion approaches (DTI, DKI, WMTI, SMT, SMTmc, BRIA; see for overview in Supplement 9). In order to normalise all metrics, we used Tract-based Spatial Statistics (TBSS)[103], as part of FSL[93, 94]. In brief, initially all brain-extracted[104] fractional anisotropy (FA) images were aligned to MNI space using non-linear transformation (FNIRT)[94]. Following, the mean FA image and related mean FA skeleton were derived. Each diffusion scalar map was projected onto the mean FA skeleton using the TBSS procedure. In order to provide a quantitative description of diffusion metrics we used the John Hopkins University (JHU) atlas[105], and obtained 30 hemisphere-specific WM regions of interest (ROIs) based on a probabilistic WM atlas (JHU)[106] for each of the 28 metrics. For T_1-weighted data, we applied the Desikan-Killiany Atlas[107]. Altogether, 840 dMRI features were derived per individual [28 metrics $\times$ (24 ROIs + 6 tracts)] for each hemisphere, and 117 T_1-weighted features (surface area, volume, thickness for each of the 34 regions; 3 whole-brain gray matter averages, and 2 averages of white matter surface area and volume) for each hemisphere.

Brain Age Predictions

Brain age was predicted using the XGBoost algorithm[108] implemented in Python (v3.7.1). We used six data subsets to predict brain age split in the following manner:

13

1) right hemisphere T_1-weighted, 2) left hemisphere T_1-weighted, 3) left hemisphere diffusion, 4) right hemisphere diffusion, 5) left hemisphere multimodal, 6) right hemisphere multimodal. We applied nested k-fold cross-validation with 5 outer and 10 inner folds (see Supplement 1 for tuned hyperparameters). We corrected for age-bias and mere age-effects[109, 110] by including age in the regression equations (Eq. 5) when assessing effects of modality, hemisphere, and handedness on brain age, as well as phenotype associations with brain ages (Eq. 9).

Statistical Analyses

All statistical analyses were carried out using Python (v3.7.1) and R (v4.2.0).

Hemispheric differences and age sensitivity

To give an overview of the extent of brain asymmetry, we assessed the significance of T_1-weighted and dMRI features' asymmetry using two-sided t-tests. The lateralisation or asymmetry of the brain features was estimated as the following: we applied the LI[32] to both regional features and features averaged over each hemisphere (see also[33]).

$$LI = \frac{L - R}{L + R},\qquad(1)$$

where L and R belongs to any left and right scalar metric, respectively. Furthermore, when associating LI with age, we used absolute LI values ($|LI|$) allowing to estimate age-effects on asymmetry irrespective of the direction of the asymmetry (leftwards or rightwards).

We then used linear regression models correcting for sex and scanning site to predict age from all regular and LI features:

$$\hat{Age} = F + Sex + Site,\qquad(2)$$

where F is a scalar metric such as, for example, hippocampus volume (derived from T_1-weighted image) or tapetum fractional anisotropy (derived from DTI). The same model setup was used applying generalised additive models (GAM) to model non-linear relationships between F and Age using a smooth s of linked quadratic functions with k = 4 knots and restricted maximum likelihood :

$$\hat{Age} = s(F) + Sex + Site.\qquad(3)$$

Likelihood ratio tests (LRTs)[111] were used to assess the age sensitivity of all T_1-weighted and dMRI features and their asymmetry/LI features by comparing the above models with baseline models not including the respective feature:

$$\hat{Age} = Sex + Site.\qquad(4)$$

We used the same procedure for region-averaged and hemispheric average metrics for regular and LI features. Hemispheric averages of regular features were then visualised

14

by age, including surface area, volume, thickness for T$_1$-weighted data, and intra- and extra-axonal water diffusivities as well as for DTI and DKI metrics.

To compare the model fit of non-linear and linear models we used the Akaike information criterion (AIC)[112] and Bayesian information criterion (BIC)[113].

Brain age assessment

We estimated correlations across HBA and GBA to assess their similarities in addition to the model output provided from the prediction procedure. We also correlated age with the LI (see Eq. (1)) for the three modalities (dMRI, T$_1$-weighted, multimodal MRI), and estimated the age sensitivity of the LI as described in Eqs. (2-4).

As preregistered (https://aspredicted.org/if5yr.pdf), to test the relationships between hemisphere (H), modality (M), and HBA while controlling for age, sex and scanner site, we employed linear mixed effects regression (LMER) models of the following form:

$$H\hat{B}A = H + M + H \times M + Sex + Age + Sex \times Age + (1|Site) + (1|I), \quad (5)$$

where I refers to the random intercept at the level of the individual. Post-hoc group differences were observed for hemisphere, modality and their interaction.

Next, handedness (Ha) was added to the model to observe whether there are model differences between the resulting LMER:

$$H\hat{B}A = Ha + H \times Ha + H + M + H \times M + Sex + Age + Sex \times Age + (1|Site) + (1|I), \quad (6)$$

and the previous model. Models were statistically compared using LRTs[111].

Finally, the LIs (Eq. 1 of left and right brain age predictions for T$_1$-weighted, diffusion and multimodal MRI (LI_{HBA}, i.e. the asymmetry in brain age predictions) were associated with age, controlling for sex and scanner site as random effect:

$$\hat{Age} = LI_{HBA} + Sex + (1|Site). \quad (7)$$

The LI_{HBA}s' age-sensitivity was then assessed (as for brain features, see Eqs. 2-4), using LRTs comparing the above model with a baseline model excluding LI_{HBA} (Eq. 4):

$$\hat{Age} = Sex + (1|Site). \quad (8)$$

Phenotype associations of brain age

In an exploratory analysis step, we assessed association patterns between brain ages and health-and-lifestyle factors which have previously demonstrated an association with brain age[20, 26, 38–41]. This analysis step served to compare phenotype associations across estimated brain ages. The health-and-lifestyle factors included alcohol drinking (binary), height and weight supplementing body mass index (BMI), diabetes diagnosis (binary), diastolic blood pressure, systolic blood pressure, pulse pressure, hypertension

15

(binary), cholesterol level (binary), and smoking (binary describing current smokers). For this last analysis step, LMERs were used with the following structure:

$$\hat{P} = BA + Sex + Age + Sex \times Age + (1|Site), \tag{9}$$

where BA refers brain age incorporating both GBA and HBA, P is the phenotype.

Furthermore, where applicable, we corrected p-values for multiple testing using Bonferroni correction and an α-level of $p < .05$. We used a high-precision approach to calculate exact p-values utilizing the Multiple Precision Floating-Point Reliable R package[114], and report standardized β-values. Sex and site were entered as independent nominal variables in the applicable regression models, with sex being a binary (0 = female, 1 = male) and scanner site a multinominal (0 = Cheadle, 1 = Newcastle, 2 = Reading, 3 = Bristol). Finally, we repeated the presented statistical analyses stratifying for sex.

DATA AVAILABILITY

All raw data are available from the UKB (www.ukbiobank.ac.uk).

CODE AVAILABILITY

Analysis code is available at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age.

16

REFERENCES

[1] Nazlee, N., Waiter, G. D. & Sandu, A.-L. Age-associated sex and asymmetry differentiation in hemispheric and lobar cortical ribbon complexity across adulthood: A uk biobank imaging study. *HBM* **44**, 49–65 (2023).

[2] Karolis, V. R., Corbetta, M. & Thiebaut de Schotten, M. The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nat. Comm.* **10**, 1417 (2019).

[3] Saltoun, K. *et al.* Dissociable brain structural asymmetry patterns reveal unique phenome-wide profiles. *Nat. Hum. Beh.* 1–18 (2022).

[4] Kong, X.-Z. *et al.* Mapping cortical brain asymmetry in 17,141 healthy individuals worldwide via the ENIGMA Consortium. *PNAS* **115**, E5154–E5163 (2018).

[5] Agcaoglu, O., Miller, R., Mayer, A. R., Hugdahl, K. & Calhoun, V. D. Lateralization of resting state networks and relationship to age and gender. *NeuroImage* **104**, 310–325 (2015).

[6] Ocklenburg, S., Friedrich, P., Güntürkün, O. & Genç, E. Intrahemispheric white matter asymmetries: the missing link between brain structure and functional lateralization? *Rev. Neurosci.* **27**, 465–480 (2016).

[7] Barrick, T. R., Lawes, I. N., Mackay, C. E. & Clark, C. A. White Matter Pathway Asymmetry Underlies Functional Lateralization. *Cereb. Cort.* **17**, 591–598 (2006).

[8] Sun, T. & Walsh, C. A. Molecular approaches to brain asymmetry and handedness. *Nat. Rev. Neur.* **7**, 655–662 (2006).

[9] Roe, J. M. *et al.* Asymmetric thinning of the Cereb. Cort. across the adult lifespan is accelerated in Alzheimer's disease. *Nat. Comm.* **12**, 721 (2021).

[10] Thompson, P. M. *et al.* The enigma consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Img. & Beh.* **8**, 153–182 (2014).

[11] Li, P. *et al.* Hemispheric asymmetry in the human brain and in parkinson's disease is linked to divergent epigenetic patterns in neurons. *Gen. Bio.* **21**, 1–23 (2020).

[12] Kong, X.-Z. *et al.* Mapping cortical and subcortical asymmetry in obsessive-compulsive disorder: findings from the enigma consortium. *Bio. Psych.* **87**, 1022–1034 (2020).

[13] Kong, X.-Z. *et al.* Mapping brain asymmetry in health and disease through the enigma consortium. *HBM* **43**, 167–181 (2022).

17

[14] Schijven, D. *et al.* Large-scale analysis of structural brain asymmetries in schizophrenia via the enigma consortium. *PNAS* **120**, e2213880120 (2023).

[15] Comparison of different neurite density metrics with brain asymmetry evaluation. *Zeitschr. Med. Phy.* (2023).

[16] Ardekani, S., Kumar, A., Bartzokis, G. & Sinha, U. Exploratory voxel-based analysis of diffusion indices and hemispheric asymmetry in normal aging. *MRI* **25**, 154–167 (2007).

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

[18] Kaufmann, T. *et al.* Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat. Neur.* **22**, 1617–1623 (2019).

[19] 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 (2021).

[20] Beck, D. *et al.* Adipose tissue distribution from body MRI is associated with cross-sectional and longitudinal brain age in adults. *NeuroImage: Clin.* **33**, 102949 (2022).

[21] Korbmacher, M. *et al.* Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing. *HBM* **44** (2023).

[22] Richard, G. *et al.* Brain age prediction in stroke patients: Highly reliable but limited sensitivity to cognitive performance and response to cognitive training. *NeuroImage: Clin.* **25**, 102159 (2020).

[23] de Lange, A.-M. G. *et al.* The maternal brain: Region-specific patterns of brain aging are traceable decades after childbirth. *HBM* **41**, 4718–4729 (2020).

[24] Voldsbekk, I. *et al.* A history of previous childbirths is linked to women's white matter brain age in midlife and older age. *HBM* **42**, 4372–4386 (2021).

[25] Richard, G. *et al.* Assessing distinct patterns of cognitive aging using tissue-specific brain age prediction based on diffusion tensor imaging and brain morphometry. *PeerJ* **6**, e5908 (2018).

[26] Cole, J. H. Multimodality neuroimaging brain-age in uk biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobio. Aging* **92**, 34–42 (2020).

[27] De Lange, A.-M. G. *et al.* Multimodal brain-age prediction and cardiovascular risk: The Whitehall II MRI sub-study. *NeuroImage* **222**, 117292 (2020).

18

[28] Good, C. D. *et al.* Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains. *NeuroImage* **14**, 685–700 (2001).

[29] Jang, H., Lee, J. Y., Lee, K. I. & Park, K. M. Are there differences in brain morphology according to handedness? *Brain & Beh.* **7**, e00730 (2017).

[30] Ocklenburg, S. *et al.* Polygenic scores for handedness and their association with asymmetries in brain structure. *Brain Struct. & Funct.* 1–13 (2021).

[31] Rentería, M. E. Cerebral asymmetry: a quantitative, multifactorial, and plastic brain phenotype. *Twin Res. & Hum. Gen.* **15**, 401–413 (2012).

[32] Benson, R. *et al.* Language dominance determined by whole brain functional mri in patients with brain lesions. *Neurol.* **52**, 798–798 (1999).

[33] Williams, C. M., Peyre, H., Toro, R. & Ramus, F. Comparing brain asymmetries independently of brain size. *NeuroImage* **254**, 119118 (2022).

[34] Jirsaraie, R. J. *et al.* Benchmarking the generalizability of brain age models: challenges posed by scanner variance and prediction bias. *HBM* **44**, 1118–1128 (2023).

[35] Korbmacher, M. *et al.* Considerations on brain age predictions from repeatedly sampled data across time. *Brain & Beh.* .

[36] Nosek, B. A., Ebersole, C. R., DeHaven, A. C. & Mellor, D. T. The preregistration revolution. *PNAS* **115**, 2600–2606 (2018).

[37] Nosek, B. A. *et al.* Preregistration is hard, and worthwhile. *Trends Cog. Sci* **23**, 815–818 (2019).

[38] Beck, D. *et al.* Cardiometabolic risk factors associated with brain age and accelerate brain ageing. *HBM* **43**, 700–720 (2022).

[39] Korbmacher, M. *et al.* Bio-psycho-social factors' associations with brain age: a large-scale uk biobank diffusion study of 35,749 participants. *Front. Psych.* **14**, 1117732 (2023).

[40] Leonardsen, E. H. *et al.* Deep neural networks learn general and clinically relevant representations of the ageing brain. *NeuroImage* **256**, 119210 (2022).

[41] Smith, S. M. *et al.* Brain aging comprises many modes of structural and functional change with distinct genetic and biophysical associations. *eLife* **9**, e52677 (2020).

[42] Sanford, N. *et al.* Sex differences in predictors and regional patterns of brain age gap estimates. *HBM* **43**, 4689–4698 (2022).

19

[43] Subramaniapillai, S. *et al.* Sex-and age-specific associations between cardiometabolic risk and white matter brain age in the uk biobank cohort. *HBM* **43**, 3759–3774 (2022).

[44] Büchel, C. *et al.* White matter asymmetry in the human brain: a diffusion tensor MRI study. *Cereb. Cort.* **14**, 945–951 (2004).

[45] Mito, R. *et al.* Fibre-specific white matter reductions in Alzheimer's disease and mild cognitive impairment. *Brain* **141**, 888–902 (2018).

[46] Takao, H., Hayashi, N. & Ohtomo, K. White matter asymmetry in healthy individuals: a diffusion tensor imaging study using tract-based spatial statistics. *Neurosci.* **193**, 291–299 (2011).

[47] Song, J. W. *et al.* Asymmetry of White Matter Pathways in Developing Human Brains. *Cereb. Cortex* **25**, 2883–2893 (2014).

[48] Huang, K. *et al.* Asymmetrical alterations of grey matter among psychiatric disorders: a systematic analysis by voxel-based activation likelihood estimation. *Prog. Neuro-Psychopharm. & Bio. Psych.* **110**, 110322 (2021).

[49] Toga, A. W. & Thompson, P. M. Mapping brain asymmetry. *Nat. Rev. Neuro.* **4**, 37–48 (2003).

[50] Chiarello, C., Vazquez, D., Felton, A. & McDowell, A. Structural asymmetry of the human cereb. cort.: Regional and between-subject variability of surface area, cortical thickness, and local gyrification. *Neuropsychologia* **93**, 365–379 (2016).

[51] Di Biase, M. A. *et al.* Mapping human brain charts cross-sectionally and longitudinally. *PNAS* **120**, e2216798120 (2023).

[52] Button, K. S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neuro.* **14**, 365–376 (2013).

[53] Taki, Y. *et al.* Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals. *PloS one* **6**, e22734 (2011).

[54] Bethlehem, R. A. *et al.* Brain charts for the human lifespan. *Nat.* **604**, 525–533 (2022).

[55] Wei, Y., Zhang, H. & Liu, Y. Charting Normative Brain Variability Across the Human Lifespan. *Neuroscience Bulletin* **39**, 362–364 (2023).

[56] Westlye, L. T. *et al.* Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb. Cort.* **20**, 2055–2068 (2010).

[57] Korbmacher, M. *et al.* Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and

20

alzheimer's disease in the uk biobank. *medRxiv* 2023–10 (2023).

[58] Xiao, D., Wang, K., Theriault, L., Charbel, E. & Initiative, A. D. N. White matter integrity and key structures affected in alzheimer's disease characterized by diffusion tensor imaging. *Eur. J. Neurosci.* **56**, 5319–5331 (2022).

[59] Rajmohan, V. & Mohandas, E. The limbic system. *Ind J Psych* **49**, 132 (2007).

[60] Redlich, R. *et al.* The limbic system in youth depression: brain structural and functional alterations in adolescent in-patients with severe depression. *Neuropsychopharm.* **43**, 546–554 (2018).

[61] Zelikowsky, M., Hersman, S., Chawla, M. K., Barnes, C. A. & Fanselow, M. S. Neuronal ensembles in amygdala, hippocampus, and prefrontal cortex track differential components of contextual fear. *J Neurosc* **34**, 8462–8466 (2014).

[62] Bari, A., Niu, T., Langevin, J.-P. & Fried, I. Limbic neuromodulation: implications for addiction, posttraumatic stress disorder, and memory. *Neurosurg. Clin.* **25**, 137–145 (2014).

[63] Coupé, P., Manjón, J. V., Lanuza, E. & Catheline, G. Lifespan changes of the human brain in alzheimer's disease. *Sci. Rep.* **9**, 3998 (2019).

[64] Yamada, S., Ishikawa, M. & Nozaki, K. Exploring mechanisms of ventricular enlargement in idiopathic normal pressure hydrocephalus: a role of cerebrospinal fluid dynamics and motile cilia. *Flu. Barr. CNS* **18**, 1–11 (2021).

[65] Schoenfeld, T. J. & Cameron, H. A. Adult neurogenesis and mental illness. *Neuropsychophar.* **40**, 113–128 (2015).

[66] Peyton, L., Oliveros, A., Choi, D.-S. & Jang, M.-H. Hippocampal regenerative medicine: neurogenic implications for addiction and mental disorders. *Exp. & Mol. Med.* **53**, 358–368 (2021).

[67] Moodley, K. & Chan, D. The hippocampus in neurodegenerative disease. *Hippo. Clin. Neurosci.* **34**, 95–108 (2014).

[68] Van de Pol, L. *et al.* Hippocampal atrophy in alzheimer disease: age matters. *Neurol.* **66**, 236–238 (2006).

[69] Van Hecke, W., Emsell, L., Sunaert, S. *et al. Diffusion tensor imaging: a practical handbook* (Springer, 2016).

[70] Huster, R. J., Westerhausen, R., Kreuder, F., Schweiger, E. & Wittling, W. Morphologic asymmetry of the human anterior cingulate cortex. *NeuroImage* **34**, 888–895 (2007).

21

[71] Kucyi, A., Moayedi, M., Weissman-Fogel, I., Hodaie, M. & Davis, K. D. Hemi-spheric asymmetry in white matter connectivity of the temporoparietal junction with the insula and prefrontal cortex. *PloS one* **7**, e35589 (2012).

[72] Wang, J. *et al.* Asymmetry of the dorsal anterior cingulate cortex: evidences from multiple modalities of mri. *Neuroinf.* **11**, 149–157 (2013).

[73] Takahashi, T. *et al.* Lack of normal structural asymmetry of the anterior cingulate gyrus in female patients with schizophrenia: a volumetric magnetic resonance imaging study. *Schiz. Res.* **55**, 69–81 (2002).

[74] Manoach, D. S. *et al.* Reduced microstructural integrity of the white matter underlying anterior cingulate cortex is associated with increased saccadic latency in schizophrenia. *NeuroImage* **37**, 599–610 (2007).

[75] Joo, S. W. *et al.* Abnormal asymmetry of white matter tracts between ventral posterior cingulate cortex and middle temporal gyrus in recent-onset schizophrenia. *Schiz. Res.* **192**, 159–166 (2018).

[76] Zhao, X. *et al.* Reduced interhemispheric white matter asymmetries in medial temporal lobe epilepsy with hippocampal sclerosis. *Front. Neurol.* **10**, 394 (2019).

[77] Zhang, Y. *et al.* Study of the microstructure of brain white matter in medial temporal lobe epilepsy based on diffusion tensor imaging. *Brain & Beh.* **13**, e2919 (2023).

[78] Cattarinussi, G. *et al.* White matter microstructure associated with the antide-pressant effects of deep brain stimulation in treatment-resistant depression: A review of diffusion tensor imaging studies. *Int. J. Molec. Sci.* **23**, 15379 (2022).

[79] Xu, J. *et al.* Effects of urban living environments on mental health in adults. *Nat. Med.* 1–12 (2023).

[80] Chan, M. Y. *et al.* Long-term prognosis and educational determinants of brain network decline in older adult individuals. *Nature aging* **1**, 1053–1067 (2021).

[81] Schoeler, T. *et al.* Participation bias in the uk biobank distorts genetic associations and downstream analyses. *Nat Hum Beh* 1–12 (2023).

[82] Lyall, D. M. *et al.* Quantifying bias in psychological and physical health in the uk biobank imaging sub-sample. *Brain Comm* **4**, fcac119 (2022).

[83] 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 (2018).

[84] Maximov, I. I. *et al.* Fast qualitY conTrol meThod foR derIved diffUsion Metrics (YTTRIUM) in big data analysis: UK Biobank 18,608 example. *HBM* **42**,

22

3141–3155 (2021).

[85] Rosen, A. F. *et al.* Quantitative assessment of structural image quality. *NeuroImage* **169**, 407–418 (2018).

[86] Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Transact. Img. Proc.* **13**, 600–612 (2004).

[87] Brunet, D., Vrscay, E. R. & Wang, Z. On the mathematical properties of the structural similarity index. *IEEE Transact. Img. Proc.* **21**, 1488–1499 (2011).

[88] Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neur.* **19**, 1523–1536 (2016).

[89] Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* **12**, e1001779 (2015).

[90] Veraart, J., Fieremans, E. & Novikov, D. S. Diffusion MRI noise mapping using random matrix theory. *MR in Med.* **76**, 1582–1593 (2016).

[91] Kellner, E., Dhital, B., Kiselev, V. G. & Reisert, M. Gibbs-ringing artifact removal based on local subvoxel-shifts. *MR in Med.* **76**, 1574–1581 (2016).

[92] Andersson, J. L. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* **125**, 1063–1078 (2016).

[93] Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).

[94] Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).

[95] Basser, P. J., Mattiello, J. & LeBihan, D. Mr diffusion tensor spectroscopy and imaging. *Biophys. J.* **66**, 259–267 (1994).

[96] Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *MR in Med.* **53**, 1432–1440 (2005).

[97] Fieremans, E., Jensen, J. H. & Helpern, J. A. White matter characterization with diffusional kurtosis imaging. *NeuroImage* **58**, 177–188 (2011).

[98] Kaden, E., Kruggel, F. & Alexander, D. C. Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *MR in Med.* **75**, 1752–1763 (2016).

23

[99] Kaden, E., Kelm, N. D., Carson, R. P., Does, M. D. & Alexander, D. C. Multi-compartment microscopic diffusion imaging. *NeuroImage* **139**, 346–359 (2016).

[100] Reisert, M., Kellner, E., Dhital, B., Hennig, J. & Kiselev, V. G. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. *NeuroImage* **147**, 964–975 (2017).

[101] Fischl, B. Freesurfer. *NeuroImage* **62**, 774–781 (2012).

[102] Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).

[103] Smith, S. M. *et al.* Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* **31**, 1487–1505 (2006).

[104] Smith, S. M. Fast robust automated brain extraction. *HBM* **17**, 143–155 (2002).

[105] Mori, S., Wakana, S., Nagae-Poetscher, L. & Van Zijl, P. MRI atlas of human white matter. *Am. J. Neurorad.* **27**, 1384 (2006).

[106] Hua, K. *et al.* Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *NeuroImage* **39**, 336–347 (2008).

[107] Desikan, R. S. *et al.* An automated labeling system for subdividing the human Cereb. Cort. on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).

[108] Chen, T. & Guestrin, C. *Xgboost: A scalable tree boosting system*, 785–794 (2016).

[109] de Lange, A.-M. G. & Cole, J. H. Commentary: Correction procedures in brain-age prediction. *NeuroImage: Clin.* **26** (2020).

[110] de Lange, A.-M. G. *et al.* Mind the gap: Performance metric evaluation in brain-age prediction. *HBM* **43**, 3113–3129 (2022).

[111] Lehmann, E. L. On likelihood ratio tests. *Selected works of E.L. Lehmann* 209–216 (2012).

[112] Akaike, H. in *Information theory and an extension of the maximum likelihood principle* 199–213 (Springer, 1998).

[113] Neath, A. A. & Cavanaugh, J. E. The bayesian information criterion: background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**, 199–203 (2012).

24

[114] Maechler, M., Maechler, M. M., MPFR, S., Suggests, M. & SuggestsNote, M. Package 'rmpfr' (2016).

25

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

M.K..: Study design, Software, Formal analysis, Visualizations, Project administration, Writing—original draft, Writing—review & editing. D.v.d.M.: Software, Writing – review & editing. D.B.: Writing—review & editing. A.M.d.L.: Software, Writing—review & editing. A.L.: Funding acquisition. E.E.: Funding acquisition. O.A.A.: Writing—review & editing, Funding acquisition. L.T.W.: Writing—review & editing, Funding acquisition. I.I.M.: Supervision, Study design, Data preprocessing and quality control, Writing—review & editing, Funding acquisition.

COMPETING INTERESTS

OAA has received a speaker's honorarium from Lundbeck and is a consultant to Coretechs.ai. We declare no other conflicts of interest.

26

Fig. 1 Age curves of standardized and zero-centered mean values of GM and WM features per hemisphere. A cubic smooth function (s) with $k = 4$ knots was applied to plot the relationship between age and brain features correcting for sex and scanner site (F): $a\hat{g}e = s(F) + sex + site$ using restricted maximum likelihood (REML). The grey shaded area indicates the 95% CI. All age-relationships were significant ($p_{adj} < .05$).

27

Fig. 2 T$_1$-weighted and dMRI features linear asymmetry-age-associations. The plot presents the standardized (sex- and site-corrected) regression slopes versus Bonferroni-adjusted -log10 p-values. Modelling was done using Eq. 2: $a\hat{g}e = \beta_0 + \beta_1 \times F + \beta_2 \times Sex + \beta_3 \times Site$, where F is the respective brain feature. Labelling was done separately for T$_1$-weighted and dMRI indicating the 10 most significantly associated features (five for $\beta > 0$ and five for $\beta < 0$). ILF = inferior longitudinal fasciculus, Cereb.Peduncle = cerebral peduncle, Rostro-mid. thicknes = rostro-middle thickness, SLFT = superior longitudinal fasciculus (temporal part), Fornix-Str.Term. = fornix-stria terminalis tract, Caud. ant. cingulate = caudal anterior cingulate. Full tables are available at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age/.

28
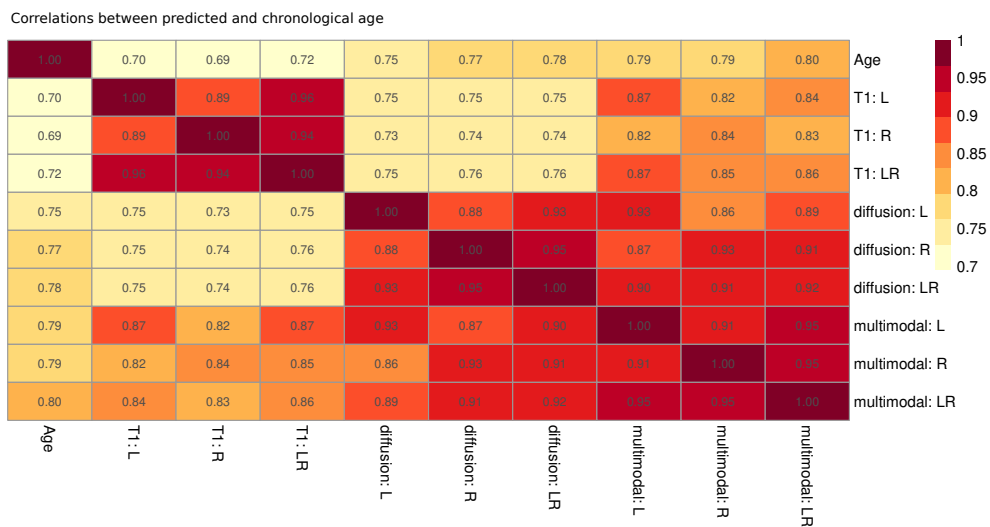
Fig. 3 Pearson correlation coefficients between chronological and predicted ages for T_1-weighted, diffusion, and multimodal MRI for left, right and both hemispheres. All Bonferroni-corrected $p < .001$. L: left hemisphere, R: right hemisphere, LR: both hemispheres.

29
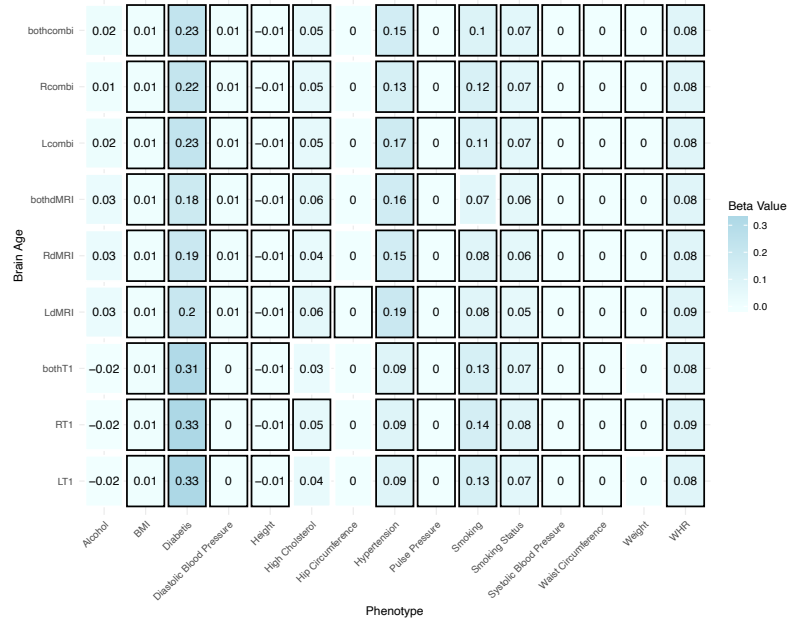
Fig. 4 Linear association between general health-and-lifestyle phenotypes and brain age estimated from different modalities, left, right and both hemispheres. Eq. 9 was used and standardized slopes are presented. For simplicity, standardized slopes with $|\beta| < 0.005$ were rounded down to $\beta = 0$. L: left hemisphere, R: right hemisphere, LR: both hemispheres, BMI: body mass index, WHR: waist-to-hip ratio. Bonferroni-adjusted $p < .05$ is marked by a black frame.

TABLES

Table 1 Hemispheric brain age prediction outcomes.

Model	Features	R^2	MAE	RMSE	Correlation*
Left T_1w	117	0.504 (0.010)	4.389 (0.054)	5.472 (0.061)	0.708 [0.703, 0.712]
Right T_1w	117	0.492 (0.008)	4.439 (0.049)	5.529 (0.051)	0.705 [0.700, 0.709]
T_1w	234	0.526 (0.011)	4.294 (0.050)	5.356 (0.062)	0.725 [0.721, 0.730]
Left dMRI	840	0.568 (0.014)	4.000 (0.047)	4.990 (0.067)	0.757 [0.753, 0.762]
Right dMRI	840	0.582 (0.013)	3.960 (0.052)	4.967 (0.079)	0.766 [0.762, 0.771]
dMRI	1680	0.605 (0.010)	3.867 (0.059)	4.821 (0.094)	0.781 [0.777, 0.785]
Left multimodal	957	0.630 (0.009)	3.757 (0.046)	4.673 (0.047)	0.794 [0.790, 0.797]
Right multimodal	957	0.634 (0.014)	3.723 (0.073)	4.673 (0.092)	0.794 [0.791, 0.798]
Multimodal	1914	0.628 (0.017)	3.663 (0.055)	4.563 (0.077)	0.793 [0.789, 0.797]

R^2 = Variance explained, MAE = Mean Absolute Error, RMSE = Root Mean Squared Error, Corr. = Correlation, Values in round parentheses () refer to standard deviations and square brackets [] to 95% confidence interval around correlations (Pearson's r) of uncorrected brain age estimates and chronological age.
* The correlation between raw brain age and chronological age.

31

SUPPLEMENTARY INFORMATION

Supplementary information to the article "Brain asymmetries from mid- to late-life and hemispheric brain age", Korbmacher et al., 2023

SUPPLEMENTARY TABLES

1   Tuned Hyperparameters

Overview of the tuned hyperparameters for each of the used brain age models.

Modality	Hemisphere	Learning Rate	Maximum Depth	Number of Trees
Multimodal	Both	0.1	8	140
Multimodal	Left	0.05	7	180
Multimodal	Right	0.1	8	140
dMRI	Both	0.1	6	100
dMRI	Left	0.1	4	180
dMRI	Right	0.1	5	180
T_1w	Both	0.1	5	140
T_1w	Left	0.1	6	140
T_1w	Right	0.1	6	180

32

2 Most age-sensitive regional features using non-linear models

T$_1$ Metric	Deviance	F	dMRI Metric	Deviance	F
superior temporal thickness (lh)	587304.16	4188.91	DKI - AK anterior limb of the internal capsule (rh)	644106.76	5170.95
hippocampus volume (rh)	576250.86	4101.39	DTI - RD fornix striaterminalis (rh)	627313.71	4981.99
thickness (lh)	576355.10	4082.87	DTI - FA anterior corona radiata (lh)	571637.61	4390.91
inferiorparietal thickness (lh)	569468.00	4041.74	DTI - FA inferior fronto-occipital fasciculus (lh)	568799.27	4366.64
hippocampus volume (lh)	565456.80	4006.59	BRIA - microRD anterior thalamic radiation (rh)	561902.12	4295.66
thickness (rh)	562548.97	3965.59	WMTI - radEAD anterior coronaradiata (rh)	433925.45	4281.90
inferior lateral ventricle volume (lh)	544864.71	3836.12	BRIA - microFA fornix striaterminalis (rh)	557084.22	4247.55
inferior lateral ventricle volume (rh)	539066.94	3786.01	DTI - FA fornix striaterminalis (rh)	545272.55	4125.27
superior temporal thickness (rh)	522564.64	3603.62	BRIA - microRD fornix striaterminalis (rh)	539180.21	4070.72
lateral ventricle volume (lh)	513713.08	3567.34	BRIA - microADC anterior thalamic radiation (rh)	536979.62	4050.30

3 Most age-sensitive regional features using linear models

T$_1$ Metric	Sum of Squares	F	dMRI Metric	Sum of Squares	F
superior temporal thickness (lh)	582215.80	12516.42	DTI - RD fornix striaterminalis (rh)	568838.72	13114.24
thickness (lh)	571936.88	12239.14	DTI - FA anterior coronaradiata (lh)	554045.66	12664.20
hippocampus volume (rh)	564806.62	12048.28	DTI - FA inferior fronto-occipital fasciculus (lh)	527205.85	11866.98
inferiorparietal thickness (rh)	559834.17	11915.90	DTI - FA fornix striaterminalis (rh)	526713.03	11852.57
thickness (rh)	557696.94	11859.18	DTI - RD anterior coronaradiata (lh)	504149.35	11201.30
hippocampus volume (lh)	554478.02	11773.95	DTI - RD anterior coronaradiata (rh)	500047.51	11084.67
superior temporal thickness (rh)	519361.12	10859.68	DTI - FA anterior coronaradiata (rh)	481860.37	10573.93
thalamus volume (rh)	470220.77	9626.25	BRIA - microRD anterior thalamic radiation (rh)	480010.76	10522.57
cortex volume (lh)	455643.18	9270.23	DTI - RD inferior fronto-occipital fasciculus (lh)	471710.65	10293.35
amygdala (lh)	454268.29	9236.88	DTI - FA inferior fronto-occipital fasciculus (rh)	470227.52	10252.62

853 # 4 Global metrics' age sensitivity using linear models

854 LRTs outcomes testing global metrics' age sensitivity using linear models (Eqs. 2 & 4), with p-values being Bonferroni-corrected for multiple comparison. Acronyms lh and rh refer to mean left and right hemisphere, respectively.

Metric	Sum of Squares	F	p	Metric	Sum of Squares	F	p
BRIA vintra (lh)	-14143.51	256.78	<.001	DTI MD (rh)	-294821.39	6256.88	<.001
BRIA vintra (rh)	-13492.91	244.88	<.001	DTI FA (lh)	-294054.08	6237.71	<.001
BRIA vextra (lh)	-8868.68	160.58	<.001	DTI FA (rh)	-290846.08	6157.77	<.001
BRIA vextra (rh)	-8247.91	149.29	<.001	SMT FA (lh)	-96237.02	1824.32	<.001
BRIA vcsf (lh)	-12339.56	223.82	<.001	SMT FA (rh)	-88924.97	1679.10	<.001
BRIA vcsf (rh)	-11691.20	211.99	<.001	SMT MD (lh)	-145717.99	2837.80	<.001
BRIA micrord (lh)	-110749.44	2115.93	<.001	SMT MD (rh)	-138236.90	2681.03	<.001
BRIA micrord (rh)	-112757.19	2156.64	<.001	SMT trans (lh)	-236947.06	4859.34	<.001
BRIA microfa (lh)	-7389.49	133.69	<.001	SMT trans (rh)	-230976.33	4720.50	<.001
BRIA microfa (rh)	-7660.08	138.61	<.001	SMT long (lh)	-233251.22	4773.28	<.001
BRIA microax (lh)	-20330.95	370.29	<.001	SMT long (rh)	-221802.60	4509.03	<.001
BRIA microax (rh)	-19217.81	349.82	<.001	SMTmc d (lh)	-12811.82	232.44	<.001
BRIA microadc (lh)	-244852.70	5044.67	<.001	SMTmc d (rh)	-15325.44	278.40	<.001
BRIA microadc (rh)	-242965.40	5000.27	<.001	SMTmc extramd (lh)	-234164.26	4794.51	<.001
BRIA dradextra (lh)	-0.87	0.02	1.00	SMTmc extramd (rh)	-221755.78	4507.96	<.001
BRIA dradextra (rh)	-0.56	0.01	1.00	SMTmc extratrans (lh)	-269921.51	5643.84	<.001
BRIA daxintra (lh)	-45776.98	844.85	<.001	SMTmc extratrans (rh)	-251971.27	5213.02	<.001
BRIA daxintra (rh)	-32572.59	597.02	<.001	SMTmc intra (lh)	-162286.59	3189.66	<.001
BRIA daxextra (lh)	-33941.70	622.56	<.001	SMTmc intra (rh)	-138122.05	2678.64	<.001
BRIA daxextra (rh)	-29058.51	531.64	<.001	WMTI awf (lh)	-216212.24	4381.26	<.001
DKI AK (lh)	-98394.96	1867.39	<.001	WMTI awf (rh)	-198966.98	3992.24	<.001
DKI AK (rh)	-107687.41	2054.02	<.001	WMTI radead (lh)	-538.93	9.72	0.11
DKI RK (lh)	-134762.36	2608.66	<.001	WMTI radead (rh)	-1786.25	32.22	<.001
DKI RK (rh)	-117109.00	2245.17	<.001	WMTI axead (lh)	-15537.30	282.28	<.001
DKI MK (lh)	-166559.26	3281.46	<.001	WMTI axead (rh)	-140593.59	2730.28	<.001
DKI MK (rh)	-146629.45	2856.99	<.001	T1 (lh) thickness	-361976.02	7460.14	<.001
DTI AD (lh)	-6414.25	115.99	<.001	T1 (rh) thickness	-337720.79	6873.27	<.001
DTI AD (rh)	-32682.00	599.06	<.001	T1 (lh) area	-131984.99	2428.71	<.001
DTI RD (lh)	-103169.79	1963.06	<.001	T1 (rh) area	-115500.16	2109.17	<.001
DTI RD (rh)	-98654.94	1872.59	<.001	T1 (lh) volume	-366138.06	7562.34	<.001
DTI MD (lh)	-296264.43	6292.97	<.001	T1 (rh) volume	-351072.41	7194.50	<.001

36

5 Global metrics' age sensitivity using non-linear models

LRTs outcomes testing global metrics' age sensitivity using generalized additive models (Eqs. (3,4)), with p-values being Bonferroni-corrected for multiple comparison. Acronyms lh and rh refer to mean left and right hemisphere, respectively.

Metric	Deviance	F	p	Metric	Deviance	F	p
BRIA vintra (lh)	298222.11	1980.42	<.001	DTI MD (rh)	420292.82	2975.29	<.001
BRIA vintra (rh)	263814.46	1721.75	<.001	DTI FA (lh)	454980.03	3284.21	<.001
BRIA vextra (lh)	99954.26	601.06	<.001	DTI FA (rh)	437831.65	3130.91	<.001
BRIA vextra (rh)	68415.48	404.41	<.001	SMT FA (lh)	231126.97	1481.17	<.001
BRIA vcsf (lh)	389707.46	2715.85	<.001	SMT FA (rh)	212502.18	1350	<.001
BRIA vcsf (rh)	395906.41	2768.18	<.001	SMT MD (lh)	338666.06	2295.03	<.001
BRIA micrord (lh)	489922	3605.08	<.001	SMT MD (rh)	329913.59	2225.14	<.001
BRIA micrord (rh)	482669.31	3537.29	<.001	SMT trans (lh)	325557.05	2188.8	<.001
BRIA microfa (lh)	468131.01	3399.04	<.001	SMT trans (rh)	309770.56	2066.61	<.001
BRIA microfa (rh)	441798.75	3161.43	<.001	SMT long (lh)	239399.25	1543.22	<.001
BRIA microax (lh)	123284.12	747.18	<.001	SMT long (rh)	220310.81	1406.65	<.001
BRIA microax (rh)	122353.86	741.87	<.001	SMTmc d (lh)	17581.83	100.96	<.001
BRIA microadc (lh)	442217.61	3169.41	<.001	SMTmc d (rh)	18705.17	107	<.001
BRIA microadc (rh)	433573.24	3092.76	<.001	SMTmc extramd (lh)	375805.34	2598.83	<.001
BRIA dradextra (lh)	265199.9	1732.72	<.001	SMTmc extramd (rh)	350591.17	2392.53	<.001
BRIA dradextra (rh)	259410.42	1690.27	1.00	SMTmc extratrans (lh)	381698.57	2646.36	<.001
BRIA daxintra (lh)	227477.58	1459.06	<.001	SMTmc extratrans (rh)	357451.71	2446.62	<.001
BRIA daxintra (rh)	221619.72	1417.52	<.001	SMTmc intra (lh)	230534.22	1477.89	<.001
BRIA daxextra (lh)	269452.37	1764.08	<.001	SMTmc intra (rh)	196608	1238.41	<.001
BRIA daxextra (rh)	265820.53	1737.25	<.001	WMTI awf (lh)	294396.47	1946.39	<.001
DKI AK (lh)	248201.74	1607.19	<.001	WMTI awf (rh)	271308.81	1773.47	<.001
DKI AK (rh)	277452.37	1822.07	<.001	WMTI radead (lh)	356837.69	2444	<.001
DKI RK (lh)	248246.37	1606.74	<.001	WMTI radead (rh)	347896.75	2371.95	<.001
DKI RK (rh)	214591.89	1365.38	<.001	WMTI axead (lh)	22893.57	133.33	<.001
DKI MK (lh)	225899.98	1446.06	<.001	WMTI axead (rh)	30036.73	175.61	<.001
DKI MK (rh)	190685.71	1195.84	<.001	T1 (lh) thickness	363679.29	2447.65	<.001
DTI AD (lh)	91486.87	545.48	<.001	T1 (rh) thickness	339637.31	2256.41	<.001
DTI AD (rh)	63150.51	374.43	<.001	T1 (lh) area	132330.67	818.92	<.001
DTI RD (lh)	492407	3628.18	<.001	T1 (rh) area	115697.02	777.46	<.001
DTI RD (rh)	481438.43	3525.6	<.001	T1 (lh) volume	366575.45	2414.26	<.001
DTI MD (lh)	425442.7	3020.18	<.001	T1 (rh) volume	351519.39	2312.27	<.001

6 Differences of T_1-weighted and dMRI features between hemispheres by sex

The table shows the ten largest regional differences between left and right hemispheres' T_1-weighted and dMRI data indicated by effect size (Cohen's d) indicated by paired samples t-tests (two-sided) and presented separately for males and females. All Bonferroni corrected $p < .05$. SLFT = Superior longitudinal fasciculus (temporal part), ILF = Inferior longitudinal fasciculus. For full tables see the files Hemi_NEW_sex_dMRI_features_diff.csv and Hemi_NEW_sex_T1w_features_diff.csv at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age.

diffusion MRI			
Feature	**Cohen's d_{males}**	**Feature**	**Cohen's $d_{females}$**
DTI - FA ILF	3.44	DTI - FA ILF	3.91
DTI - AD SLFT	2.09	DTI - AD SLFT	2.40
WMTI - axEAD SLFT	2.01	SMTmc - diff SLFT	2.06
DTI - FA cingulate gyrus	1.93	SMT - long SLFT	2.04
DKI - RK cingulate gyrus	1.90	DTI - FA cingulate gyrus	1.98
WMTI - AWF cingulate gyrus	1.83	SMTmc - extratrans cerebral peduncle	1.96
DTI - AD ILF	1.81	SMTmc - extraMD SLFT	1.93
DTI - FA superior frontooccipital fasciculus	1.77	BRIA - microAX SLFT	1.92
DKI - RK SLFT	1.75	DKI - RK SLFT	1.91
SMTmc - extratrans cerebral peduncle	1.74	SMTmc - intra cingulate gyrus	1.89
T_1-weighted MRI			
Feature	**Cohen's d_{males}**	**Feature**	**Cohen's $d_{females}$**
frontal pole area	1.82	transverse temporal area	1.89
pars orbitalis area	1.78	frontal pole area	1.73
transverse temporal area	1.77	pars orbitalis area	1.72
inferior parietal area	1.71	inferior parietal area	1.72
inferior parietal volume	1.62	inferior parietal volume	1.64
frontal pole volume	1.58	frontal pole volume	1.54
thalamus volume	1.40	middle temporal area	1.42
middle temporal area	1.31	transverse temporal volume	1.38
transverse temporal volume	1.29	thalamus volume	1.34
pars orbitalis volume	1.27	pars orbitalis volume	1.29

7 Most age-sensitive regional T$_1$- and diffusion-weighted features using *non*-linear models *by sex*

The table shows the ten largest regional differences between left and right hemispheres' T$_1$-weighted and dMRI data indicated by F from LRTs comparing a baseline model (Eq. 4) to the GAM (Eq. 3) presented separately for males and females. All Bonferroni corrected $p < .05$. ATR = Anterior thalamic radiation, IFOF = inferior fronto-occipital fasciculus. For full tables see the files Hemi_NEW_REGIONAL_dMRI_non_linear_hemi_effects_MALES.csv, Hemi_NEW_REGIONAL_dMRI_non_linear_hemi_effects_FEMALES.csv, Hemi_NEW_REGIONAL_T1_non_linear_hemi_effects_MALES.csv, and Hemi_NEW_REGIONAL_T1_non_linear_hemi_effects_FEMALES.csv at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age.

Males					
T$_1$ Metric	**Deviance**	**F**	**dMRI Metric**	**Deviance**	**F**
Hippocampus volume (rh)	325396.71	2327.14	DTI - RD fornix striaterminalis (rh)	321483.12	2474.87
Inferior lateral ventricle volume (lh)	315630.20	2242.52	DKI - AK Anteriorlimbofinternalcapsule (rh)	315123.70	2406.94
Hippocampus volume (lh)	314178.29	2222.80	DTI - FA fornix striaterminalis (rh)	287920.97	2127.34
Lateral ventricle volume (rh)	294791.55	2055.25	DTI - FA IFOF (lh)	286229.79	2114.11
Superior temporal thickness (lh)	288883.19	1973.34	BRIA - micro Rd ATR (rh)	285675.67	2109.47
Thickness (lh)	286653.65	1944.76	DTI - FA Anteriorcoronaradiata (lh)	285098.32	2099.07
Thickness (rh)	285659.63	1943.02	BRIA - micro FA Fornix Striaterminalis (rh)	280978.76	2062.43
Lateral ventricle volume (lh)	280759.99	1932.22	BRIA - micro Rd ATR (lh)	268901.80	1946.45
Bankssts thickness (lh)	111534.08	1927.07	DTI - RD ATR (rh)	268857.43	1944.99
Rostral middle frontal volume (rh)	112544.65	1887.40	DTI - RD ATR (lh)	268408.62	1942.57
Females					
T$_1$ Metric	**Deviance**	**F**	**dMRI Metric**	**Deviance**	**F**
Superior temporal thickness (lh)	298436.46	2186.28	DKI - AK Anteriorlimbofinternalcapsule (rh)	328299.92	2756.78
Inferior parietal thickness (rh)	294083.30	2157.56	DTI - RD Fornix Striaterminalis (rh)	309568.49	2539.12
Thickness (lh)	289859.17	2098.94	DTI - FA Anteriorcoronaradiata (lh)	287115.30	2288.94
Thickness (rh)	277328.73	1988.89	DTI - FA IFOF (rh)	282230.37	2243.21
Superiortemporal thickness (rh)	268345.92	1902.38	BRIA - micro FA Fornix Striaterminalis (rh)	279454.04	2213.63
Hippocampus volume (lh)	256888.62	1827.15	BRIA - micro Rd Fornix Striaterminalis (rh)	279158.45	2213.20
Hippocampus volume (rh)	256386.19	1820.25	BRIA - micro Rd ATR (rh)	279221.52	2209.87
Lateral ventricle volume (rh)	247973.59	1755.77	DTI - RD ATR (lh)	278213.99	2202.86
Lateral ventricle volume (lh)	237509.49	1666.32	DTI - RD Anteriorcoronaradiata (lh)	277873.07	2197.90
Supramarginal thickness (rh)	235411.09	1632.94	BRIA - micro Rd ATR (lh)	274318.30	2160.44

8 Most age-sensitive regional T$_1$- and diffusion-weighted features using linear models *by sex*

The table shows the ten largest regional differences between left and right hemispheres' T$_1$-weighted and dMRI data indicated by F from LRTs comparing a baseline model (Eq. 4) to the linear model (Eq. 2) presented separately for males and females. All Bonferroni corrected $p <$ .05. ATR = Anterior thalamic radiation, SLFT = superior longitudinal fasciculus (temporal part), IFOF = inferior fronto-occipital fasciculus. For full tables see the files Hemi_NEW_REGIONAL_dMRI_linear_hemi_effects_MALES.csv, Hemi_NEW_REGIONAL_dMRI_linear_hemi_effects_FEMALES.csv, Hemi_NEW_REGIONAL_T1_linear_hemi_effects_MALES.csv, and Hemi_NEW_REGIONAL_T1_linear_hemi_effects_FEMALES.csv at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age.

	Males				
T$_1$ Metric	**SS**	**F**	**dMRI Metric**	**SS**	**F**
Hippocampus volume (rh)	316544.24	6766.75	DTI - RD fornix striaterminalis (rh)	286594.17	6364.49
Hippocampus volume (lh)	306351.25	6487.54	DTI - FA fornix striaterminalis (rh)	276467.17	6067.57
Superior temporal thickness (lh)	286384.87	5955.49	DTI - FA anterior corona radiata (lh)	274955.01	6023.83
Thickness (lh)	284626.62	5909.55	DTI - FA IFOF (lh)	263814.28	5706.24
Thickness (rh)	283428.75	5878.34	DTI - RD anterior corona radiata (lh)	246467.64	5227.51
Inferior parietal thickness (rh)	271677.80	5575.68	DTI - RD anterior corona radiata (rh)	242227.42	5113.30
Inferior lateral ventricle volume (lh)	254969.32	5156.08	DTI - FA anterior corona radiata (rh)	238605.78	5016.60
Superior temporal thickness (rh)	252923.59	5105.55	DTI - FA IFOF (rh)	233533.32	4882.47
Thalamus volume (rh)	247813.43	4980.11	BRIA - microRD ATR (rh)	232131.22	4845.66
Amygdala volume (lh)	243131.64	4866.16	DTI - RD IFOF (lh)	230505.57	4803.12
T$_1$ Metric	**SS**	**F**	**dMRI Metric**	**SS**	**F**
Superior temporal thickness (lh)	295682.20	6561.30	DTI - RD fornix striaterminalis (rh)	282234.33	6743.14
Inferior parietal thickness (rh)	288649.89	6365.48	DTI - FA anterior corona radiata (lh)	279013.72	6641.48
Thickness (lh)	287657.68	6338.05	DTI - FA IFOF (lh)	263549.29	6163.65
Thickness (rh)	274795.27	5986.72	DTI - RD anterior corona radiata (rh)	258369.12	6007.30
Superior temporal thickness (rh)	266558.45	5765.84	DTI - RD anterior corona radiata (lh)	257953.77	5994.85
Hippocampus volume (rh)	248481.56	5291.98	DTI - FA fornix striaterminalis (rh)	251128.24	5791.79
Hippocampus volume (lh)	248438.59	5290.87	BRIA - microRD ATR (rh)	249535.85	5744.86
Supramarginal thickness (rh)	232265.05	4879.14	BRIA - microRD ATR (lh)	243660.64	5573.15
Supramarginal thickness (lh)	225512.01	4710.53	DTI - FA anterior corona radiata (rh)	243324.99	5563.40
Precuneus thickness (rh)	223530.22	4661.41	DTI - RD IFOF (lh)	241625.80	5514.19

9 Description of white matter features by diffusion approaches.

Diffusion Approach	Metrics
Bayesian Rotationally Invariant Approach (BRIA) [100]	intra-axonal axial diffusivity (DAX intra)
	extra-axonal radial diffusivity (DRAD extra)
	microscopic fractional anisotropy (micro FA)
	extra-axonal axial diffusivity (DAX extra)
	intra-axonal water fraction (V intra)
	extra-axonal water fraction (V extra)
	cerebrospinal fluid fraction (vCSF)
	microscopical axial diffusivity (micro AX)
	microscopic radial diffusivity (micro RD)
	microscopical apparent diffusion coefficient (micro ADC)
Diffusion Kurtosis Imaging (DKI) [96, 97]	mean kurtosis (MK)
	radial kurtosis (RK)
	axial kurtosis (AK)
Diffusion Tensor Imaging (DTI) [95]	fractional anisotropy (FA)
	axial diffusivity (AD)
	mean diffusivity (MD)
	radial diffusivity (RD)
Spherical Mean Technique (SMT) [98]	fractional anisotropy (SMT FA)
	mean diffusivity (SMT md)
	transverse diffusion coefficient (SMT trans)
	longitudinal diffusion coefficient (SMT long)
Multi-compartment Spherical Mean Technique (SMTmc) [99]	extra-neurite microscopic mean diffusivity (SMTmc extra md)
	extra-neurite transverse microscopic diffusivity (SMTmc extra trans)
	mc SMTdiffusion coefficient (SMT mcd)
	intra-neurite volume fraction (SMTmc intra)
White Matter Tract Integrity (WMTI) [97]	axonal water fraction (AWF)
	radial extra-axonal diffusivity (radEAD)
	axial extra-axonal diffusivity (axEAD)

10 Differences of T₁w and dMRI features between hemispheres

The table shows the ten largest regional differences between left and right hemispheres' T_1w and dMRI data indicated by effect size (Cohen's d) indicated by paired samples t-tests (two-sided). SLFT = Superior longitudinal fasciculus (temporal part), ILF = Inferior longitudinal fasciculus. Bonferroni-adjusted p-values were $p < 2 \times 10^{-308}$. For full tables see the files Hemi_dMRI_features_diff.csv and Hemi_T1w_features_diff.csv at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age.

T_1-weighted MRI			diffusion MRI		
Feature	**T-value**	**Cohen's d**	**Feature**	**T-value**	**Cohen's d**
Transverse temporal area	397.45	1.81	DTI - FA ILF	725.48	3.64
Frontal pole area	-386.34	1.76	DTI - AD SLFT	-444.89	2.23
Pars orbitalis area	-380.71	1.74	DTI - FA cingulate gyrus	388.09	1.95
Inferior parietal area	-368.85	1.68	DKI - RK cingulate gyrus	375.36	1.89
Inferior parietal volume	-352.95	1.61	SMTmc - diff SLFT	-369.19	1.85
Frontal pole volume	-340.08	1.55	SMTmc - extratrans cerebral peduncle	-367.31	1.84
Middletemporal area	-297.79	1.36	DKI - RK SLFT	-364.52	1.83
Thalamus Proper	296.93	1.35	WMTI - AWF cingulate gyrus	364.46	1.83
Transverse temporal volume	292.04	1.33	SMT - long SLFT	-359.08	1.80
Pars orbitalis volume	-280.74	1.28	DTI - AD ILF	353.43	1.78

SUPPLEMENTARY FIGURES

11 Uncorrected mean values' age curves

Uncorrected standardized and zero-centered age curves and lines for mean values of grey and white and grey matter features by age per hemisphere. For line fitting, first, a cubic smooth function (s) with $k = 4$ knots was applied to plot the relationship between age and brain features (F): $a\hat{g}e = s(F)$. Second, a linear model was applied of the following form: $a\hat{g}e = \beta_0 + \beta_1 \times F$. Models used restricted maximum likelihood (REML). Extreme outliers defined by Mean$\pm$9SD were removed for visualisation purposes.

44
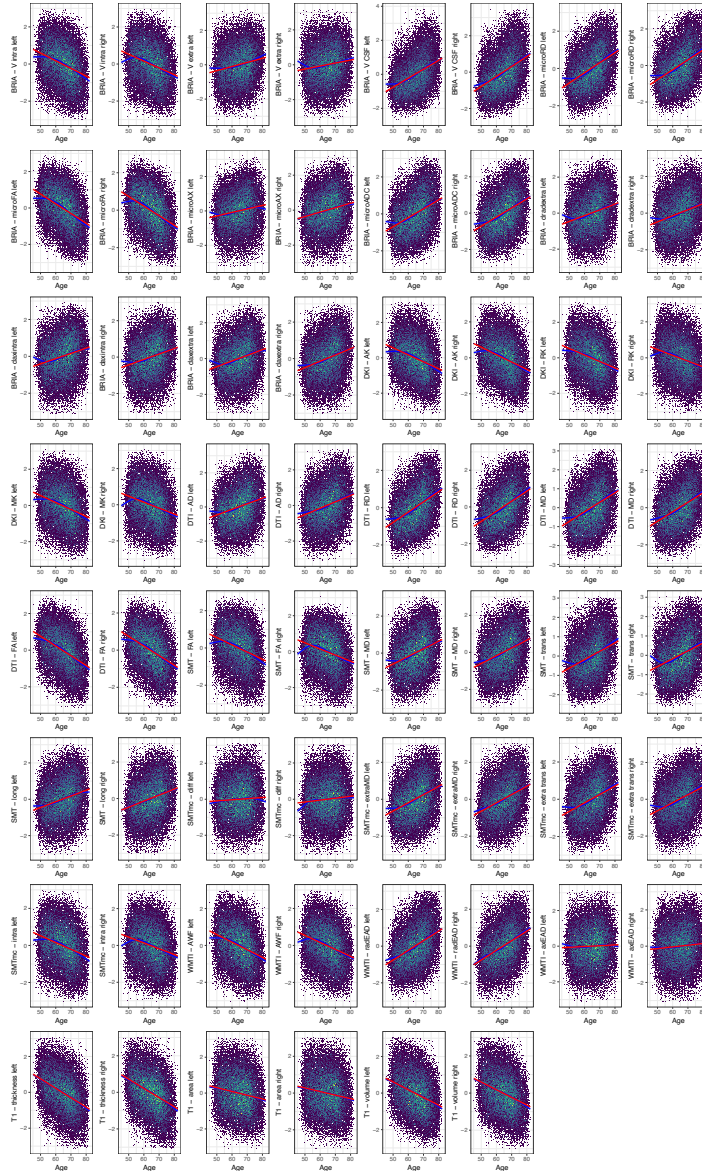
12 Distribution of the *significant and non-significant* slopes of age-related laterality indexed grey and white matter features

We estimated the absolute laterality index ($|LI|$) for each regional feature to assess the overall directionality of asymmetry-age associations. The distrubutions of age-relationship of $|LI|$ are displayed with the six panels showing the distributions for the modality-specific features (T_1-weighted and diffusion-weighted) for both sexes, males and females.

13 Distribution of the *significant* slopes of age-related laterality indexed grey and white matter features

We estimated the absolute laterality index ($|LI|$) for each regional feature to assess the overall directionality of asymmetry-age associations. The distrubutions of age-relationship of $|LI|$ are displayed with the six panels showing the distributions for the modality-specific features (T_1-weighted and diffusion-weighted) for both sexes, males and females.

14 Linear, adjusted hemispheric mean values' age associations

Corrected standardized and zero-centered linear age relationships for mean hemispheric values of grey and white matter features by age per hemisphere. Modelling was done using Eq. 2:
$a\hat{g}e = \beta_0 + \beta_1 \times F + \beta_2 \times Sex + \beta_3 \times Site$, where F is the respective brain feature.

15 Adjusted mean values' hemisphere-specific age associations by sex

Age curves of standardized and zero-centered mean values of GM and WM features per hemisphere and by sex. A cubic smooth function (s) with $k = 4$ knots was applied to plot the relationship between age and brain features correcting for sex and scanner site (F): $a\hat{g}e = s(F) + sex + site$ using restricted maximum likelihood (REML). The grey shaded area indicates the 95% CI. All age-relationships were significant ($p_{adj} < .05$).

48

16 Adjusted mean values' sex-specific age associations by hemisphere

Age curves of standardized and zero-centered mean values of GM and WM features per hemisphere and by sex. A cubic smooth function (s) with $k = 4$ knots was applied to plot the relationship between age and brain features correcting for sex and scanner site (F): $a\hat{g}e = s(F) + sex + site$ using restricted maximum likelihood (REML). The grey shaded area indicates the 95% CI. All age-relationships were significant ($p_{adj} < .05$).

49

17 Association between general health-and-lifestyle phenotypes and brain age estimated from different modalities, left, right and both hemispheres by sex

Eq. 9 was used (yet stratifying by sex) and standardized slopes are presented. For simplicity, standardized slopes with $|\beta| < 0.005$ were rounded down to $\beta = 0$. Panel a) males, panel b) females. L: left hemisphere, R: right hemisphere, LR: both hemispheres, BMI: body mass index, WHR: waist-to-hip ratio. Bonferroni-adjusted $p < .05$ is marked by a black frame.

Beta Value 0.0 0.1 0.2 0.3

a — Brain Age (males)

	Alcohol	BMI	Diabetes	Diastolic Blood Pressure	Height	High Cholesterol	Hip Circumference	Hypertension	Pulse Pressure	Smoking	Smoking Status	Systolic Blood Pressure	Waist Circumference	Weight	WHR
bothcombi	0.01	0.02	0.24	0.01	-0.01	0.05	0	0.15	0	0.12	0.08	0	0.01	0	0.1
Rcombi	0	0.02	0.22	0.01	-0.01	0.05	0	0.14	0	0.14	0.09	0	0.01	0	0.11
Lcombi	0.01	0.02	0.26	0.01	-0.01	0.07	0	0.18	0	0.13	0.08	0	0.01	0	0.11
bothdMRI	0.03	0.01	0.21	0.01	-0.01	0.07	0	0.17	0	0.07	0.07	0	0.01	0	0.1
RdMRI	0.02	0.02	0.21	0.01	-0.01	0.05	0	0.15	0	0.07	0.07	0	0.01	0	0.1
LdMRI	0.03	0.02	0.22	0.01	-0.01	0.08	0	0.19	0	0.07	0.06	0	0.01	0	0.11
bothT1	-0.03	0.02	0.32	0	-0.01	0.04	0	0.1	0	0.17	0.09	0	0.01	0	0.1
RT1	-0.03	0.02	0.34	0	-0.01	0.06	0	0.1	0	0.2	0.1	0	0.01	0	0.11
LT1	-0.03	0.02	0.35	0	-0.01	0.05	0	0.1	0	0.17	0.09	0	0.01	0	0.11

Phenotype

b — Brain Age (females)

	Alcohol	BMI	Diabetes	Diastolic Blood Pressure	Height	High Cholesterol	Hip Circumference	Hypertension	Pulse Pressure	Smoking	Smoking Status	Systolic Blood Pressure	Waist Circumference	Weight	WHR
bothcombi	0.02	0	0.21	0.01	-0.01	0.05	0	0.14	0	0.07	0.05	0	0	0	0.06
Rcombi	0.02	0	0.21	0.01	-0.01	0.04	0	0.12	0	0.1	0.06	0	0	0	0.07
Lcombi	0.02	0	0.2	0.01	-0.01	0.04	0	0.15	0	0.09	0.06	0	0	0	0.06
bothdMRI	0.03	0	0.12	0.01	-0.01	0.04	0	0.14	0	0.07	0.06	0	0	0	0.06
RdMRI	0.03	0.01	0.15	0.01	0	0.03	0	0.15	0	0.09	0.06	0	0	0	0.07
LdMRI	0.03	0.01	0.15	0.01	-0.01	0.04	0	0.19	0	0.09	0.05	0.01	0	0	0.07
bothT1	-0.01	0	0.29	0	-0.01	0.02	0	0.08	0	0.08	0.05	0	0	0	0.07
RT1	-0.01	0	0.3	0	-0.01	0.03	0	0.08	0	0.07	0.05	0	0	0	0.07
LT1	-0.01	0	0.3	0	-0.01	0.02	0	0.07	0	0.09	0.05	0	0	0	0.07

Phenotype

50

18 Males' T$_1$-weighted and dMRI features asymmetry-age-associations

T$_1$-weighted and dMRI features linear asymmetry-age-associations. The plot presents the standardized, site-corrected regression slopes versus Bonferroni-adjusted -log10 p-values for males. Modelling was done using a sex-stratified version of Eq. 2: $a\hat{g}e = \beta_0 + \beta_1 \times F + \beta_2 \times Site$, where F is the respective brain feature. Labelling was done separately for T$_1$-weighted and dMRI indicating the 10 most significantly associated features (five for $\beta > 0$ and five for $\beta < 0$). Cereb.Peduncle = cerebral peduncle, Rostro-mid. thicknes = rostro-middle thickness, SLFT = superior longitudinal fasciculus (temporal part), Fornix-Str.Term. = fornix-stria terminalis tract, Rost. ant. cingulate = rostral anterior cingulate. Full tables are available at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age/.

19 Females' T_1-weighted and dMRI features asymmetry-age-associations

960
961
962
963
964
965
966
967

T_1-weighted and dMRI features linear asymmetry-age-associations. The plot presents the standardized, site-corrected regression slopes versus Bonferroni-adjusted -log10 p-values for females. Modelling was done using a sex-stratified version of Eq. 2: $a\hat{g}e = \beta_0 + \beta_1 \times F + \beta_2 \times Site$, where F is the respective brain feature. Labelling was done separately for T_1-weighted and dMRI indicating the 10 most significantly associated features (five for $\beta > 0$ and five for $\beta < 0$). Cereb.Peduncle = cerebral peduncle, Rostro-mid. thicknes = rostro-middle thickness, SLFL = superior longitudinal fasciculus, Sup.front.occ.Fasc. = superior fronto-occipital fasciculus. Full tables are available at https://github.com/MaxKorbmacher/Hemispheric_Brain_Age/.
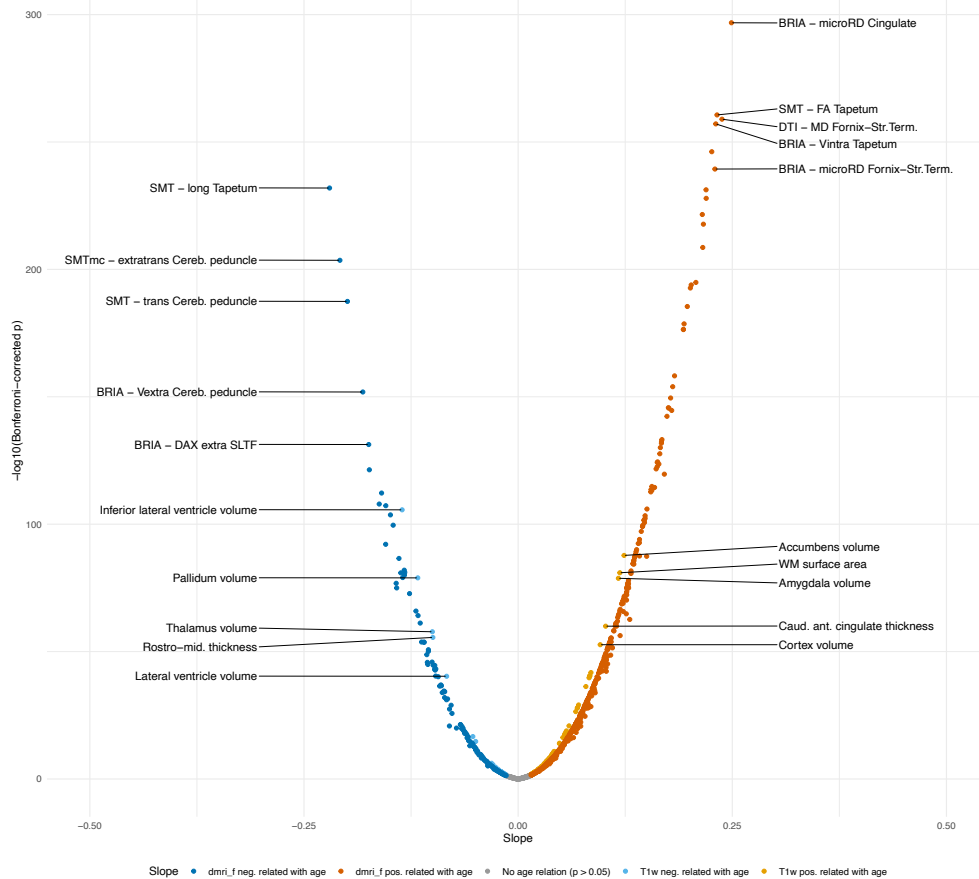
DISTINCT LONGITUDINAL BRAIN WHITE MATTER MICROSTRUCTURE CHANGES AND ASSOCIATED POLYGENIC RISK OF COMMON PSYCHIATRIC DISORDERS AND ALZHEIMER'S DISEASE IN THE UK BIOBANK

M. Korbmacher, D. van der Meer, D. Beck, Daniel E. Askeland-Gjerde, E. Eikefjord, A. Lundervold, O. A. Andreassen, L. T. Westlye, I. I. Maximov

Distinct longitudinal brain white matter microstructure changes and associated polygenic risk of common psychiatric disorders and Alzheimer's disease in the UK Biobank

Max Korbmacher[1,2,3,✉], **Dennis van der Meer**[2,4], **Dani Beck**[2,5,6], **Daniel E. Askeland-Gjerde**[2], **Eli Eikefjord**[1,3], **Arvid Lundervold**[1,3,7,8], **Ole A. Andreassen**[2,9], **Lars T. Westlye**[2,6,9], and **Ivan I. Maximov**[1,2,✉]

[1]Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway
[2]NORMENT Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo and Oslo University Hospital, Oslo, Norway
[3]Mohn Medical Imaging and Visualization Centre (MMIV), Bergen, Norway
[4]Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands
[5]Department of Psychiatric Research, Diakonhjemmet Hospital, Oslo, Norway
[6]Department of Psychology, University of Oslo, Oslo, Norway
[7]Department of Radiology, Haukeland University Hospital, Bergen, Norway
[8]Department of Biomedicine, University of Bergen, Bergen, Norway
[9]KG Jebsen Centre for Neurodevelopmental Disorders, University of Oslo, Oslo, Norway

During the course of adulthood and ageing, white matter (WM) structure and organisation are characterised by slow degradation processes such as demyelination and shrinkage. An acceleration of such ageing process has been linked to the development of a range of diseases. Thus, an accurate description of healthy brain maturation, in particular, in terms of WM features, provides a cornerstone in the understanding of ageing. We use longitudinal diffusion magnetic resonance imaging to provide an overview of WM changes at different spatial and temporal scales in the UK Biobank (UKB) (N=2,678; age_{scan1}=62.38±7.23 years; age_{scan2}=64.81±7.1 years). To examine the genetic overlap between WM structure and common clinical conditions, we tested the associations between WM structure and polygenic risk scores (PGRS) for the most common neurodegenerative disorder, Alzheimer's disease, and common psychiatric disorders (uni- and bipolar depression, anxiety, obsessive-compulsive, autism, schizophrenia, attention-deficit-hyperactivity) in longitudinal (N=2,329) and cross-sectional UKB validation data (N=31,056). Global and regional single and multi-compartment fractional anisotropy, intra-axonal water fraction, and kurtosis metrics decreased ($\bar{\beta}$=-0.04),whereas diffusivity metrics, and free water increased with age (β=0.05), with the annual rate of WM change (ARoC) accelerating at higher ages for both global ($\bar{\beta}$=0.01) and regional WM metrics ($\bar{\beta}$=0.01). Voxel-level trends indicated decreasing anisotropy, and variable spatial patterns for other diffusion metrics, suggesting differential changes in frontal compared to other brain regions. Although effect sizes were small ($|\bar{\beta}_{all}| = 0.01$), ARoC in middle cerebral peduncle WM had the strongest association with PGRS, especially for Alzheimer's: $|\bar{\beta}| = 0.01$. PGRS were more strongly related to ARoC than cross-sectional measures (d_{scan1}=0.03, d_{scan2}=0.03, $d_{validation}$=0.03). Our findings indicate spatially distributed WM changes across the brain, as well as distributed associations of PGRS with WM. Importantly, brain longitudinal changes reflected the genetic risk for disorder development better than the utilised cross-sectional measures, with regional differences giving more specific insights into gene-brain change associations than global averages.

Ageing | White Matter | Microstructure | Brain Ageing | Polygenic Risk | Magnetic Resonance Imaging | Diffusion MRI

Correspondence: *max.korbmacher@hvl.no, ivan.maximov@hvl.no*

Introduction

White matter microstructure (WMM) changes significantly throughout the lifespan. Recent large-scale studies suggest a strong association between WMM and age during both healthy and diseased ageing (1–7). Previous findings demonstrated general trends of tissue anisotropy increases and water diffusivity decreases throughout childhood (8–10), and, as reversal dynamics of these trends, throughout adulthood (3, 11). Importantly, abnormal white matter (WM) development has been associated with the development of neurocognitive skills and mental health symptoms in childhood and adolescence (12) and brain disorders later in life (13, 14). A person's genetically determined propensity to develop a certain disorder can also be summarised by polygenic risk scores (PGRS) (15–23). Combining brain imaging and genetics opens the opportunity to associate the elevated genetic risk for disorders with specific brain features (24) in terms of scalar imaging metrics (25, 26). This allows to identify which brain regions might be more prone to disease developments based on the genetic makeup, and provide additional biological detail to the observed WM changes. An advantage of using WMM for such associations is the level of detail provided by biophysical models, such as intra- and extra axonal diffusion processes (27–31). WM-specific changes are associated with various common psychiatric disorders (13, 14), and precede, for example, the symptom onset in Alzheimer's Disease (AD) (32). This outlines WMM as an important aspect for further investigation in disease formation and outcomes.

As the temporal aspect matters when investigating tissue changes, longitudinal designs are required (6). Hence, longitudinal changes in WM are informative when examining PGRS-tissue associations as they allow to connect PGRS with actual WM change and not only cross-sectional "snapshots" of brain structure at a given moment of an individual's life.

In order to estimate ageing effects on WM alterations, studies commonly focus on diffusion tensor imaging (DTI) (33). While DTI is useful for mesostructural characterisations, DTI is also limited in addressing certain common phenomena in WM such as crossing fibre bundles, non-Gaussian diffu-

sion, and differences between intra- and extra-axonal water compartments (34). Recent achievements in advanced diffusion magnetic resonance imaging (dMRI) techniques offer a spectrum of biophysical models (27–31) addressing these issues, for example, by differentiating between intra and extra-axonal space (27, 29, 31), or by capturing non-Gaussian diffusion (27, 28). In turn, there are few longitudinal studies observing WMM changes (10, 35, 36), and even fewer also utilising advanced diffusion approaches beyond DTI (6, 10). In order to fill this gap, we assess the metrics of a series of dMRI approaches in a large longitudinal mid-to-late life adult sample provided by the UKB (37) and identify the spatio-temporal patterns of ageing-related WMM changes on a global, regional, and voxel-level scale. To further investigate potential genetic underpinnings of these WMM changes, we estimate PGRS informed by previous genome-wide association studies for AD, the most common neurodegenerative disorder, and common psychiatric disorders, including Major Depressive Disorder (MDD), Bipolar Disorder (BIP), Anxiety Disorder (ANX), Autism Spectrum Disorder (ASD), Schizophrenia (SCZ), Attention Deficit Hyperactivity Disorder (ADHD), and Obsessive-Compulsive Disorder (OCD). PGRS capture an individual's genetic propensity for a trait by aggregating the estimated effects of risk variants across the entire genome. Together with brain structure and brain structural changes, PGRS may be informative for the development of disease. For example, these gene-brain associations can help identify concrete spatial patterns for different diseases (38, 39). For further inference on the generalisability of the associations between PGRS and WM in the longitudinal sample with available PGRS (after exclusions: N = 2,329), we estimated the same associations for independent participants from the cross-sectional portion of the UKB (after exclusions: N = 31,056) (37). Based on previous findings (3, 40), we expected near-linear age associations of diffusion metrics, generally outlining lower fractional anisotropy, intra-axonal water fraction, and kurtosis at higher ages, but higher diffusivity, extra-axonal free water fraction. Moreover, we expected to observe PGRS-WMM associations for AD globally and in most age-sensitive regions.

Methods

A. Sample characteristics. We obtained UKB data (37), including the longitudinal dMRI data of N = 4,871 participants at two time points. Participant data were excluded when consent had been withdrawn, and when dMRI data were not meeting quality control (QC) standards using the YTTRIUM method (41) (see also Supplement 1). Additionally, we excluded participants which were diagnosed with any mental and behavioural disorder (ICD-10 category F), disease of the nervous system (ICD-10 category G), and disease of the circulatory system (ICD-10 category I). Remaining data sets after exclusions were applied were N = 2,678 participants (52.99% females). At baseline, participants were on average 62.26±7.19 years old (range: 46.12-80.30 years) and at time point two, mean age was 64.70±7.07 years (range: 49.33-82.59 years), indicating an average age difference as

Δage =2.44±0.73 years (range: 1.12-6.90 years). The data were collected at three sites: (1) in Cheadle (57.36%), (2) Newcastle (37.04%), and (3) Reading (5.60%). PGRS data were available for N = 2,329 of these longitudinal data sets and for N = 31,056 cross-sectional validation data sets (after exclusions).

B. MRI acquisition and post-processing. UKB MRI data acquisition procedures are described elsewhere (37, 42, 43).

After obtaining access to the raw dMRI data, we pre-processed it using an optimised pipeline (41). The pipeline includes corrections for noise (44), Gibbs ringing (45), susceptibility-induced and motion distortions, and eddy currents artefacts (46). Isotropic $1\,mm^3$ Gaussian smoothing was carried out using FSL's (47, 48) *fslmaths*. Employing the multi-shell data, Diffusion Tensor Imaging (DTI), Diffusion Kurtosis Imaging (DKI) (28) and White Matter Tract Integrity (WMTI) (27) metrics were estimated using Matlab 2017b code (https://github.com/NYU-DiffusionMRI/DESIGNER). Spherical mean technique SMT (30), and multi-compartment spherical mean technique (mcSMT) (29) metrics were estimated using original code (https://github.com/ekaden/smt) (29, 30). Estimates from the Bayesian Rotational Invariant Approach (BRIA) were evaluated by the original Matlab code (https://bitbucket.org/reisert/baydiff/src/master/) (31).

In total, we obtained 26 WM metrics from six diffusion approaches (DTI, DKI, WMTI, SMT, mcSMT, BRIA; see for overview in Supplement 2). In order to normalise all metrics, we used Tract-based Spatial Statistics (TBSS) (49), as part of FSL (47, 48). In brief, initially all brain-extracted (50) fractional anisotropy (FA) images were aligned to MNI space using non-linear transformation (FNIRT) (48). Following, the mean FA image and related mean FA skeleton were derived. Each diffusion scalar map was projected onto the mean FA skeleton using TBSS. To provide a quantitative description of diffusion metrics at a region level, we used the John Hopkins University (JHU) atlas (51), and obtained 30 hemisphere-specific white matter (WM) regions of interest (ROIs) based on a probabilistic WM atlas (JHU) (52) for each of the 26 metrics. Altogether, 1,794 diffusion features were derived per individual [26 metrics × (48 ROIs + 20 tracts + 1 global mean value)].

C. Polygenic Risk Scores. We estimated PGRS for each participant with available genomic data, using PRSice2 (53) with default settings. As input for the PGRS, we used summary statistics from recent genome-wide association studies of ASD (15), MDD (19), SCZ (22), ADHD (16), BIP (21), OCD (17), ANX (18), and AD (20). We used a minor allele frequency of 0.05, as most commonly used threshold across PRS studies of psychiatric disorders.

While psychiatric disorders were p-values thresholded at α = 0.05 (15–23), recommendations for AD ($\alpha = 1.07^{-4}$ (54)) lead to the application of a lower threshold of $\alpha = 0.0001$, with the goal of optimising signal to noise in comparison to

previously used $\alpha = 0.001$ (55). The goal with the estimation of the PGRS was to relate cross-sectional WMM metrics and WMM changes to disease-related genetic profiles to examine to which degree these genetic risk profiles can explain WMM changes in midlife to senescence.

Statistical Analyses. All statistical analyses were carried out using R version 4.2.0 (www.r-project.org), and FSL version 6.0.1 (48).

First, we assessed unadjusted time point differences by using paired samples t-tests on a set of cognitive measures, and each of the global and regional scalar diffusion metrics F (such as FA from DTI) and present Cohen's d indicating the effect size:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}}, \tag{1}$$

describing the difference between means ($\bar{x}_1$, $\bar{x}_2$) over an estimate of the sample standard deviation of the data ($\hat{\sigma}$).

We then used linear mixed effects regression models (LMER) adjusting F for age, sex, the sex-age interaction (sex×age), time point (TP) and scanner site ($Site$). ID was treated as random intercept (RRI), and y the y-intercept:

$$\hat{F} = y + \beta_0 \times Age + \beta_1 \times Sex + \\ \beta_2 \times Age \times Sex + \beta_3 \times TP + \beta_4 \times Site + RI(ID) \tag{2}$$

For comparison of age-relationships between time points, we utilised simple linear and generalised additive models (GAMs) on each single time point (treating the data as cross-sectional), with the linear models taking the following form:

$$\hat{F} = y + \beta_0 \times Age + \beta_1 \times Sex + \\ \beta_2 \times Age * Sex + \beta_3 \times Site, \tag{3}$$

and the GAMs this form (including a spline function of age $S(Age)$ to model non-linear associations):

$$\hat{F} = y + S(Age) + \beta_1 \times Sex + \\ \beta_2 \times Age * Sex + \beta_3 \times Site. \tag{4}$$

We also estimated the annual rate of change ($ARoC$) of each global feature by taking the difference in WMM features F between time points (scan1&2) indicated by the $\Delta age = age_{scan2} - age_{scan1}$:

$$ARoC = \frac{F_{scan2} - F_{scan1}}{\Delta age}. \tag{5}$$

Global $ARoC$ was corrected for sex, sex×age, and scanner site (as RI):

$$A\hat{R}oC_{global} = y + \beta_0 \times Age * Sex + \beta_1 \times Sex + RI(Site), \tag{6}$$

for which then age-correlations were estimated. For regional features, we used simple linear models to support model convergence:

$$A\hat{R}oC_{regional} = y + \beta_0 \times Age * Sex + \beta_1 \times Sex + \\ \beta_2 \times Age + \beta_3 \times Site. \tag{7}$$

For the voxel-level analysis, we estimated one-sample t-tests on the contrast between each time point's maps within the FA skeleton accounting for age, Δage, scanner site and sex using FSL randomise with 10,000 permutations (H$_0$: Difference = 0).

Finally, we assess the associations between $ARoC$ and $PGRS$ for global and regional WMM metrics adjusting for age, age×sex, and site using simple linear models:

$$A\hat{R}oC = y + \beta_0 \times Age * Sex + \beta_1 \times Age + \\ \beta_2 \times Sex + \beta_3 \times Site + \beta_4 \times PGRS. \tag{8}$$

For voxel-level analyses, we used TBSS randomise with permutation-based statistics (running 10,000 permutations) (**?**). Mean maps and between-time-point contrast maps served for the computation of one-sample t-tests for each of the observed metric, while accounting for age, sex and site (i.e., random intercept models):

$$\hat{F} = y + \beta_0 \times Age + \beta_1 \times Sex + \\ \beta_2 \times Site + \beta_3 \times TP + RI(ID). \tag{9}$$

P-values were adjusted for multiple comparison using Bonferroni-correction(56) for global and region-averaged metrics, and family-wise error (FWE)-corrections were used for voxelwise inferential statistics (using Threshold-Free Cluster Enhancement(57)). We report conditional variance explained in the main text (R_c^2), which refers to variance explained by fixed factors.

Results

Cognitive changes. Our analysis suggested no significant time point differences in cognitive measures for the inter-scan interval of Δ =2.44±0.73 (Supplement 16).

Global WMM changes. The globally averaged WMM metrics differed between time points (Cohen's $|\bar{d}| = 0.073 \pm 0.055, d_{min} = -0.248, d_{max} = 0.159$), with the exception of BRIA - micro ADC, all SMT metrics, and SMTmc - diffusion coefficient (Supplement 3-4).

Congruently, LMER (Eq.   2) outline the effect of time point, in addition to age and sex (Supplement 5a-b).   While the effects of time point ($|\bar{\beta}| = 0.078 \pm 0.024, \beta_{min} = -0.107, \beta_{max} = 0.113$), age ($|\bar{\beta}| = 0.047 \pm 0.014, \beta_{min} = -0.063, \beta_{max} = 0.067$), and sex×age ($|\bar{\beta}| = 0.26 \pm 0.182, \beta_{min} = -0.619, \beta_{max} = 0.197$) were significant ($p < .001$; Supplement 6-7), and modelling diffusion metrics well $R^2_c = 88.40\% \pm 7.33\%$ (Supplement 8), only two diffusion metrics showed significant sensitivity to sex (BRIA micro FA: $\beta = -0.642, p = 0.034$, DKI radial kurtosis: $\beta = -0.785, p = 0.022$, both lower in males)), and presented large variability (Supplement 5c-d, Supplement 9-10). Moreover, WMM metrics' age-associations were better modelled as non-linear ($\bar{R}^2 = 15.81\% \pm 6.88\%$) than linear associations ($\bar{R}^2 = 14.69\% \pm 6.84\%$), considering variance explained (Fig.   1).   Yet, this difference in model fit was non-significant ($p = 0.557$).   We show decreases in FA,

and kurtosis ($|\bar{\beta}| = -0.043 \pm 0.009, \beta_{min} = -0.063, \beta_{max} = -0.035$). At the same time, water diffusivity and extra-axonal water fraction, including free water (mainly cerebro-spinal fluid) increase ($|\bar{\beta}| = 0.052 \pm 0.013, \beta_{min} = 0.025, \beta_{max} = 0.067$). Notably, the intra-axonal water fraction, estimated by different diffusion approaches, exhibit the same behaviour independent of the diffusion approach (lower at higher ages). Finally, we present accelerations of the annual rate of change (ARoC) at higher ages for most global WMM metrics (Fig. 2). Large age-group differences were observed at Cohen's $|\bar{d}_{sign}| = 0.802 \pm 0.453, |d_{min}| = 0.015, |d_{max}| = 1.920$ (when including also non-significant findings: $|\bar{d}_{all}| = 0.732 \pm 0.477, |d_{min}| = 0.010, |d_{max}| = 1.92$ (see Supplement 18 for details on test statistics, including effect size estimates). Decelerating ARoC were observed for BRIA-DAX intra and extra, and DKI-AK, and relatively stable ARoC for SMT - longitudinal coefficient. See also Supplement 19 for ARoC-age-trends indicating accelerated ageing at $\bar{\beta}_{sig} = 0.012$, Supplement 20 for uncorrected age-stratification, Supplement 21 for corrected sex-stratification, respectively.

Regional white matter microstructure changes. Estimating paired-samples tests indicate that 47.57% of white matter features decreased ($|\bar{d}_{sign}| = -0.191 \pm 0.122, |d_{min}| = -0.825, |d_{max}| = -0.030$), 32.30% increased ($|\bar{d}_{sign}| = 0.159 \pm 0.091, |d_{min}| = 0.283, |d_{max}| = 0.516$) and 20.13% did not change between time points. Most extreme p-values among these unadjusted time-point differences were observed in Fornix (DTI-RD: $d_{p_{min}} = 0.262, 95\% CI[0.252, 0.271]$) and the body of the corpus callosum (DTI-FA: $d_{p_{min}} = -0.270, 95\% CI[-0.281, -0.259]$), and largest effect sizes of Cohen's $|d| > 0.5$ found in the middle cerebellar peduncle (BRIA microAX: $d_{max} = -0.825, 95\% CI[-0.875, -0.775]$; Fig. 3a).

LMER (Eq. 2), outlined strongest age-associations in fornix WMM ($|d| > 0.5$; Fig. 3b, see Supplement 13 for distribution of β-values). We also identify various regions sensitive to sex differences and smaller sex×age-interaction effects, with repeated occurrence of significant fornix, cerebral peduncle, corticospinal tract, and medial lemniscus differences (Supplement 11,12). Across regions, time point was a non-significant fixed effect ($\beta < 1.5 \times 10^{-5}, p > .05$). For comparability across regions, we used $|ARoCs|$ in association with age, adjusting for $age \times sex$, sex, and site effects, showing accelerated ageing ($\bar{\beta}_{sig} = 0.013 \pm 0.005$; Supplement 17).

Voxelwise white matter microstructure changes. When investigating voxelwise changes of WMM between time points (adjusted for age, sex, and scanner site), we find two major patterns examining both increases and decreases in WMM: 1) a global decrease in fractional anisotropy metrics (with the exception of SMT-FA showing distributed changes, including frontal increase and posterior decrease), and 2) an overall decrease in radial, axial and mean diffusivity metrics in orbitofrontal, occipital and brain-stem and cerebellum and

increase in superior frontal areas (Table 1). For an overview of the voxelwise WMM maps at corrected $\alpha = 0.05$ see Supplement 25.
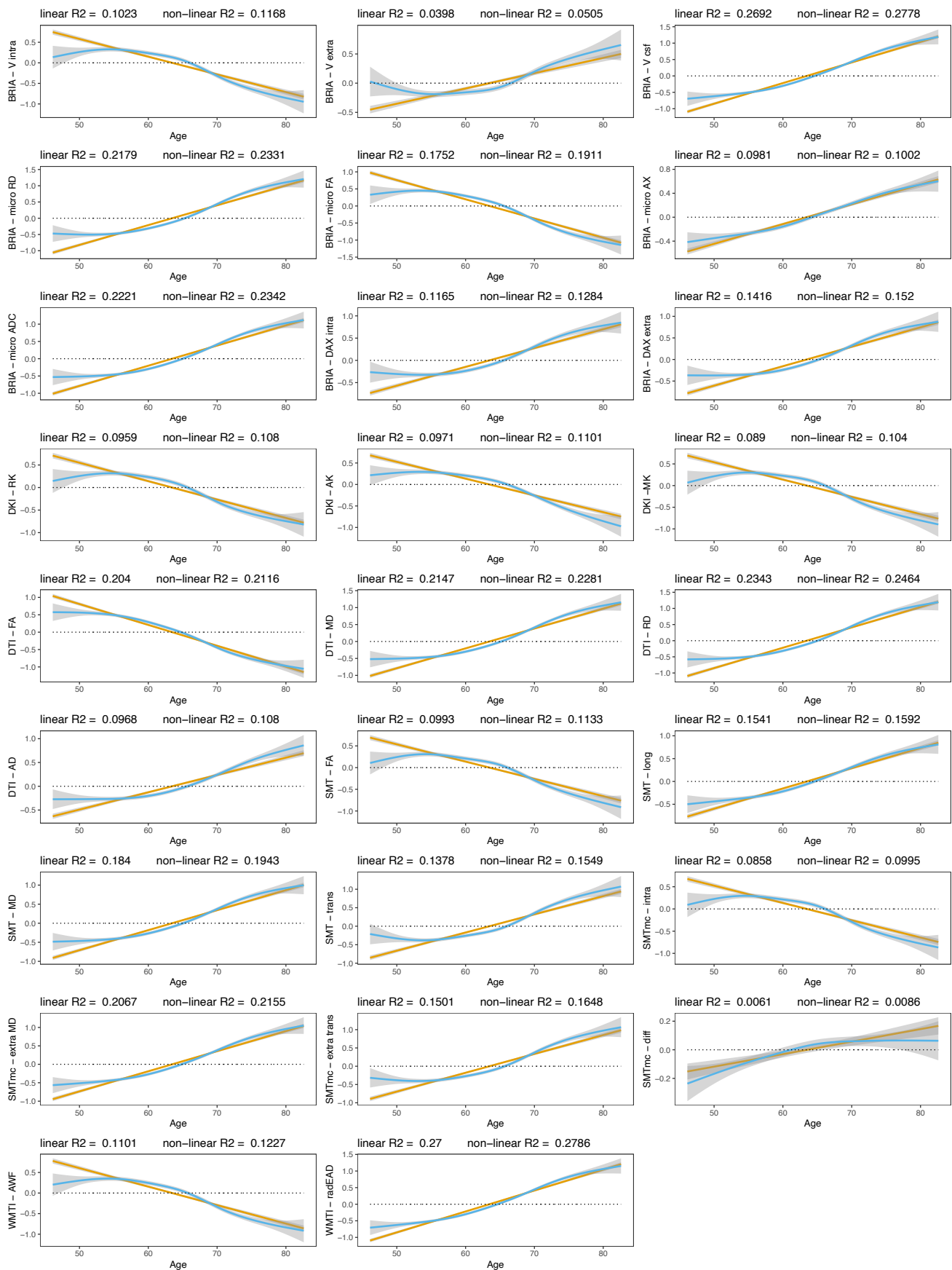
Fig. 1. Global WMM ageing trajectories. WMM values were standardised, mean-centred, and adjusted for covariates of no interest using LMERs. Second, age-WMM relationships were described by linear and non-linear functions.
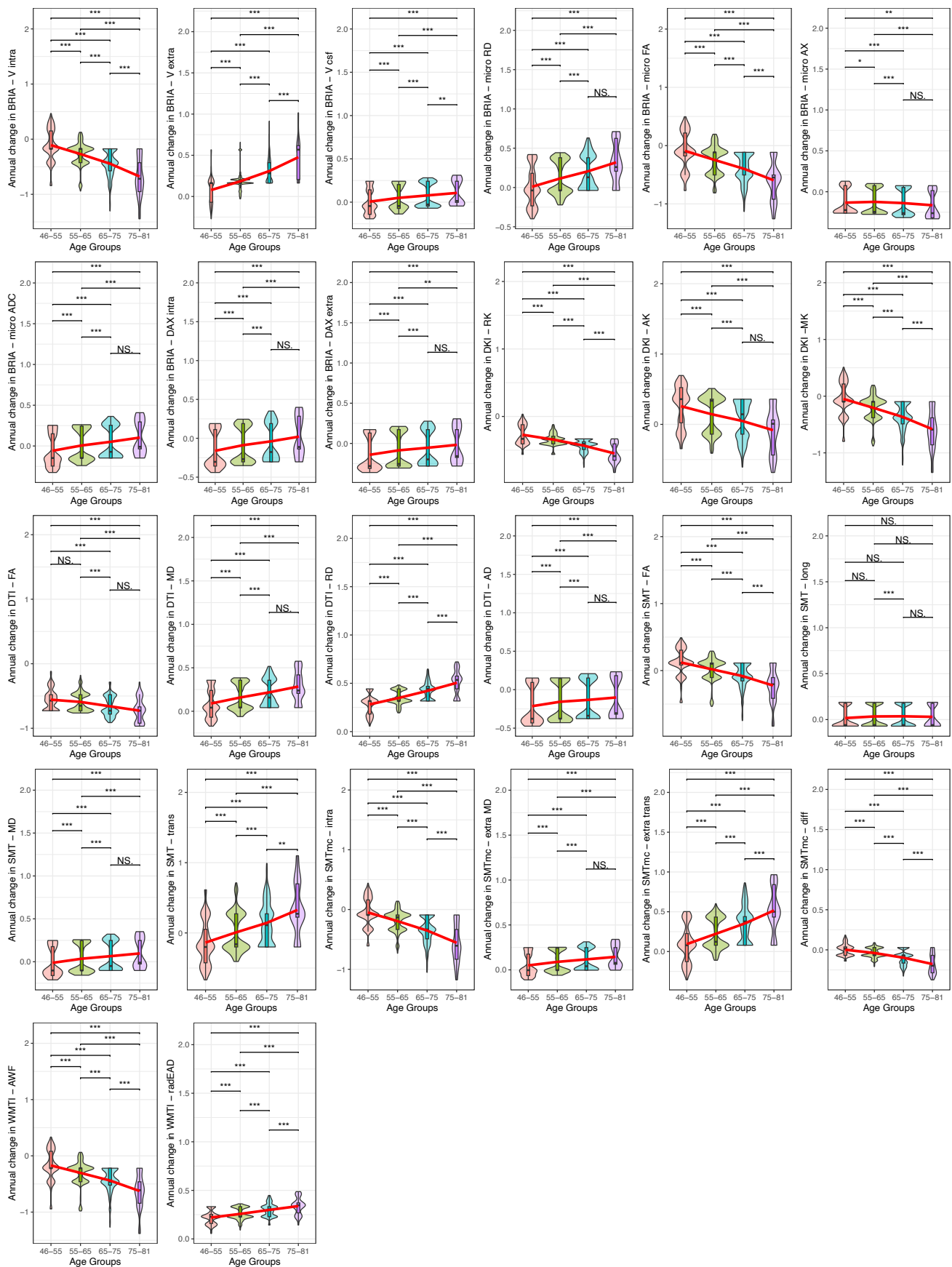
Fig. 2. Age-stratified annual WMM change. WMM were corrected for age, sex, age$\times$sex and site, and standardized for comparability (without mean-centering). We present p-values for Wilcoxon tests, which were significant at the Bonferroni-corrected * $\alpha < 0.05/(26*6) = 3.21 \times 10^{-4}$, ** $\alpha < 0.01/(26*6) = 6.41 \times 10^{-5}$, *** $\alpha < 0.001/(26*6) = 6.41 \times 10^{-6}$, NS = non-significant. The red lines were added as a visual help to identify trends of accelerated or decelerated annual change.
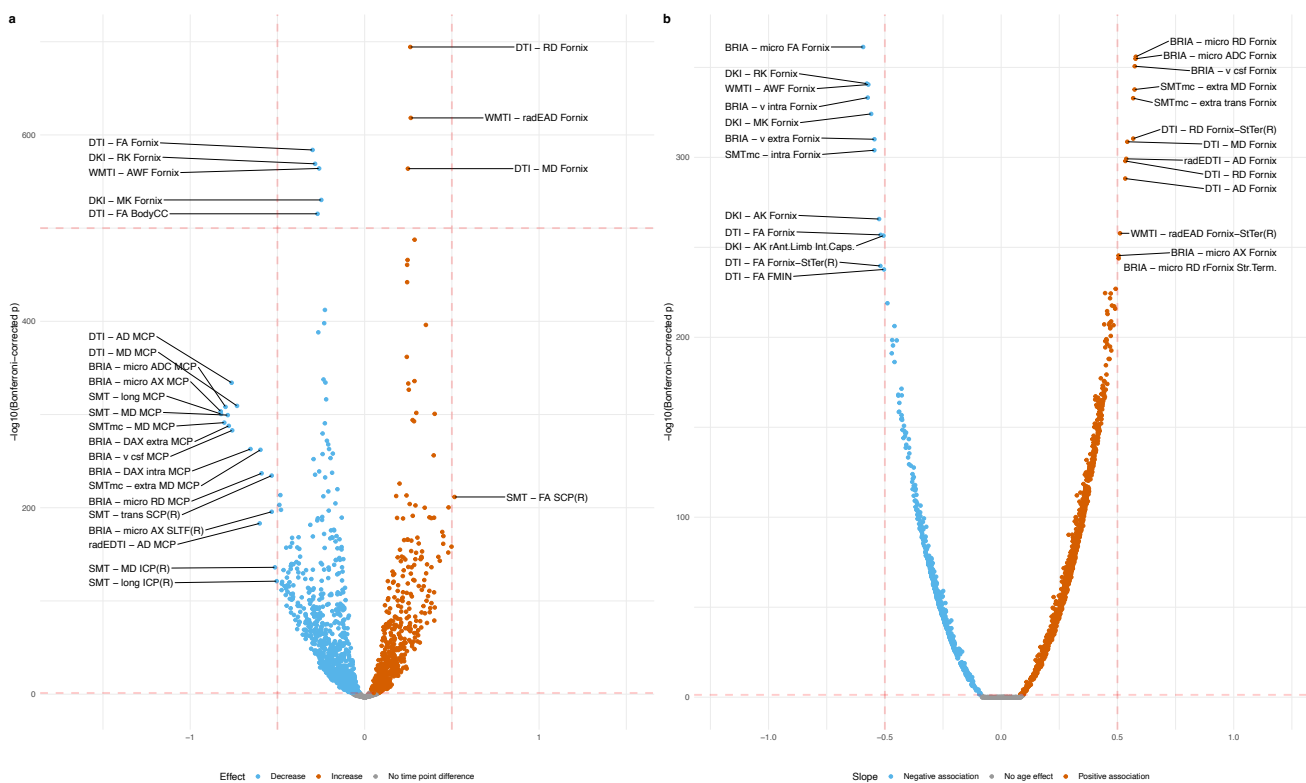
Fig. 3. Regional white matter microstructure changes between time points and age-associations. Panel (a) presents unadjusted effect sizes (Cohen's d) versus Bonferroni-adjusted -log10 p-values. Labelling was done using a medium effect size threshold of Cohen's $|d| > 0.5$ (also marked with vertical lines) as well as extreme Bonferroni-adjusted p-values of $-10log(p) > 500$. Panel (b) presents adjusted WMM associations with age. Age-WMM were adjusted for sex, sex×age, scanner site and time point (Eq. 2). The plot presents standardised slopes (β) versus Bonferroni-adjusted -log10 p-values. Labelling was done using a large association of $|\beta| > 0.5$ (also marked with vertical lines). BodyCC = body of the corpus callosum, SCP(R) = right superior cerebellar peduncle. MCP = middle cerebellar peduncle, ICP = inferior cerebellar peduncle. SLTF(R) = right superior longitudinal temporal fasciculus. StTer(R) = right stria terminalis. Dotted lines were inserted as visual aid: the lower horizontal dotted line represent the significance level of $p = .05$ and the upper horizontal line $-10log(p) = 500$. The vertical lines represented labelling borders based on a medium effect size of Cohen's $|d| > 0.5$ (panel a) and large associations of $|\beta| > 0.5$ (panel b). Tables with test statistics are available at https://github.com/MaxKorbmacher/Long_Diffusion/.

PGRS Associations. Although annual change in the cerebral peduncle showed the strongest associations with PGRS of AD and global WMM with ADHD, both global ($\bar{\beta} = 0.015 \pm 0.012$) and regional ($\bar{\beta} = 0.011 \pm 0.009$) WMM-PGRS associations were non-significant after adjusting the α-level for multiple comparisons (Fig. 4a-b). Nevertheless, a highly brain-region-specific pattern of associations between WMM ARoC and PGRS was observed (Fig. 4b) with the medial cerebral peduncle (Fig. 4c) showing the strongest consistent associations with PGRS and specifically associations with AD ($|\beta_{max}| = 0.053$, $|\bar{\beta}| = 0.014 \pm 0.013$), and MDD ($|\beta_{max}| = 0.051$, $|\bar{\beta}| = 0.014 \pm 0.013$; Fig. 4b-c, see for all peduncle-PGRS associations Supplement 22, indicating strongest ARoC-PGRS ($|\bar{\beta}| = 0.015$) associations but weaker non-replicating associations for cross-sectional assessments ($|\bar{\beta}_{scan1}| = 0.001, |\bar{\beta}_{scan2}| = 0.001, |\bar{\beta}_{validation}| = 0.001$)). Fornix, the most age-sensitive region related strongest to ANX ($|\beta_{max}| = 0.011$, $|\bar{\beta}| = 0.006 \pm 0.003$), and OCD PGRS ($|\beta_{max}| = 0.008$, $|\bar{\beta}| = 0.005 \pm 0.001$; Supplement 23).

Similarly, observing each time point separately, both global and regional WMM were non-significant after adjusting for multiple comparisons. However, WMM-PGRS associations were highly similar at the two time points in the longi-

tudinal sample, yet highlighted BIP and SCZ associations (considering $p_{uncorrected} < 0.05$, Fig. 5a-b), did however not replicate in an independent cross-sectional validation sample of $N = 31,056$ UKB participants (Fig. 5c). Noteworthy, WMM-PGRS associations were strongest for AD ($|\bar{\beta}_{scan1}| = 0.0013, |\bar{\beta}_{scan2}| = 0.0012, |\bar{\beta}_{validation}| = 0.0005$) and ANX ($|\bar{\beta}_{scan1}| = 0.0010, |\bar{\beta}_{scan2}| = 0.0012, |\bar{\beta}_{validation}| = 0.0004$) in the longitudinal sample, but strongest for ADHD ($|\bar{\beta}_{scan1}| = 0.0008, |\bar{\beta}_{scan2}| = 0.0010, |\bar{\beta}_{validation}| = 0.0010$), and ASD in the cross-sectional validation sample ($|\bar{\beta}_{scan1}| = 0.0011, |\bar{\beta}_{scan2}| = 0.0008, |\bar{\beta}_{validation}| = 0.0010$, Fig. 5d-f).

Overall, ARoC-PGRS associations ($\bar{\beta}_{long} = 0.011 \pm 0.009$) were significantly stronger than cross-sectional effects $\bar{\beta}_{scan1} = 0.001 \pm 0.001, d = 0.028, 95\%CI[0.004, 0.051], p < .001$, $\bar{\beta}_{scan2} = 0.001 \pm 0.001, d = 0.026, 95\%CI[0.003, 0.049], p < .001$, $\bar{\beta}_{validation} = 0.0007 \pm 0.001, d = 0.031, 95\%CI[0.008, 0.054], p < .001$ (see Supplement 24 for distribution of effects).

Table 1. Trends in WMM changes.

Model	Metric	Superior Frontal	Orbito-Frontal	Posterior	Cerb.-Brainst.[1]
BRIA	Dax extra	↑	↑↓	↓	↓
	Dax intra	↑	↑↓	↓	↓
	DRAD extra*	↑	↑↓	↓	↓
	micro ADC	↑	↑↓	↓	↓
	micro AX	↑	↑↓	↓	↑↓
	micro FA	↓	↑↓	↑↓	↑↓
	micro RD	↑	↑↓	↓	↑↓
	Vcsf	↑	↑↓	↓	↑↓
	Vextra	↓	↑↓	↑	↑↓
	Vintra	↑↓	↑↓	↓	↑↓
DKI	AK	↓	↑	↑↓	↑↓
	MK	↑↓	↑↓	↓	↑↓
	RK	↑↓	↑↓	↓	↓
DTI	AD	↑	↓	↓	↑↓
	FA	↓	↓	↓	↑↓
	MD	↑	↑	↓	↓
	RD	↑	↑	↑↓	↓
SMT	FA	↑	↑↓	↓	↑↓
	MD	↑	↑↓	↓	↓
	long	↑	↑↓	↓	↓
	trans	↑↓	↑↓	↑↓	↑↓
SMTmc	diffusion	↑	↑↓	↓	↑↓
	Extra MD	↑	↑↓	↓	↑↓
	Extra trans	↑	↑↓	↑↓	↑↓
	intra	↑↓	↑↓	↓	↑↓
WMTI	AWF	↑↓	↑↓	↓	↑↓
	axEAD*	↑	↑↓	↓	↑↓
	radEAD	↑	↑↓	↑↓	↑↓

↑ indicates a general trend of increases, and ↓ of decreases. ↑↓ indicates trends of both increases and decreases.

*Note that Drad extra and axEAD were excluded from the analyses as a significant portion of the produced metrics did not pass our quality control procedure(41).

1 Subcortical refers to the cerebellum and the brain stem.

Fig. 4. Associations of PGRS with the rate of white matter change. Panel (a) presents the global associations between PGRS and WMM change. Colours indicate the strength of association (standardised β-coefficients). Panel (b) presents the regional associations between PGRS and WMM change. The dotted line indicates an uncorrected $\alpha < 0.001$. Labels presenting the respective regions and metrics are supplied above this α-threshold, as well as at $|\beta| > 0.05$. Panel (c) presents the regional associations exclusively for the medial cerebral peduncle, the region where the strongest ARoC-PGRS associations were observed. Boxes in panels (a) and (c) indicate the statistical significance at an uncorrected $\alpha < 0.05$. *Note:* All associations were adjusted for age, sex, the $age \times sex$ interaction, and site. None of the presented associations survived the adjustment of the α-level for multiple comparisons.

Fig. 5. Cross-sectional associations between PGRS and WMM in longitudinal and cross-sectional validation data. Panel (a) shows PGRS-associations of globally averaged WMM for time point one in the longitudinal sample $N = 2,329$, and panel (b) for time point two, respectively. Panel (c) shows the global WMM-PGRS associations for the cross-sectional validation sample $N = 31,056$. Boxes indicate significance at an uncorrected $\alpha < 0.05$. For simplicity, standardised regression coefficients with $|\beta| < 0.005$ were rounded down to $\beta = 0$. Panel (d) shows PGRS-associations of regionally averaged WMM for time point one in the longitudinal sample, and panel (e) for time point two. Panel (f) presents the regional associations for the cross-sectional validation sample. The dotted line in panels (d-f) indicates an uncorrected $\alpha < 0.001$. Labels presenting the respective regions and metrics are supplied above this α threshold, and $|\beta| > 0.001$. *Note:* All associations were adjusted for age, sex, $age \times sex$ as fixed effects, and site as random effect, and none of the associations survived the adjustment of the α-level for multiple comparisons.

Discussion

We investigated WMM changes using longitudinal UKB diffusion data using a series of diffusion approaches and their associations with PGRS. The comparison between two time points, with an average inter-scan interval of $\bar{\Delta}$=2.44 years, was performed at different spatial scales in order to localise the strongest ageing effect and its correlations with other covariates. Unadjusted time point differences in global WMM metrics (Supplement 3) resembled closely the covariate-adjusted time point differences, as well as age effects (Supplement 5a-b). These ageing effects largely confirmed previous findings showing a global decrease of fractional anisotropy and the increase in axial, radial and mean diffusivity for both intra- and extra-axonal space(3, 6, 40, 58–60), whereas kurtosis metrics decrease with age(3, 6). Noteworthy, we provided evidence for the intra-axonal water fraction estimated by different diffusion approaches demonstrating age-related decrease, accompanied by opposite trends for the extra-axonal water fraction and the CSF water fraction (also known as *free water* in other diffusion approaches).

We observed accelerated change in global and regional WMM at higher ages. Observed inconsistencies in age dependencies of the annual change in axonal diffusivity stand in contrast to a previous longitudinal study analysing DTI data on the voxel-level(36). Such difference in the localisation of the increased acceleration and in findings on axial diffusivity might be driven by methodological variations: here, we focused on global and regional averages of annual WMM change. Yet, in contrast to the mentioned study(36), another longitudinal investigation did also not detect accelerated axial diffusivity changes(61). However, our study provides larger statistical power than previous longitudinal investigations on the rate of change in WMM (e.g.,(36, 61)), and is hence bench-marking global and regional increases in the rate of WM change at a higher age.

Region-wise investigations allowed for more differentiated results and presented larger ageing effects compared to global WMM metrics (Fig. 3, Supplement 13). Region-level assessments further underlined the strong effect of age on WMM(3, 6, 40, 60, 61). Differences between unadjusted and adjusted associations gave more detail to dependencies of regional WMM age associations on sex and scanner site. Unadjusted associations outlined the cerebellar peduncle, and the superior longitudinal temporal fasciculus as regions with largest age-association, and corpus callosum and fornix as statistically most significantly age-associated regions. After corrections, the fornix was identified as the most significant and strongest age-associated region, which is congruent with previous cross-sectional findings on limbic WMM age associations(3, 4, 40, 62). These WMM age associations also followed the described global pattern of fractional anisotropy metrics decreasing at higher ages, and axial, radial, and mean diffusivity metrics increasing during ageing.

The observed regional fornix changes were moreover clearly delineated in the voxel-level analysis. Additionally, a more general pattern was observed: radial, axial and mean diffusivity metrics increased in superior frontal areas but decreased in more posterior and inferior areas (Table 1, Supplement 25). While previous findings show consistent DTI AD, RD and MD increases and FA decreases across the brain for global and tract measures throughout ageing(60, 63), or particularly in the brain stem(36), no study has yet outlined such differential, non-homogeneous spatial patterns of WMM changes across the brain. Only one other study(64) revealed a similar pattern for DTI when examining cross-sectional AD, RD, and MD tract-age-associations in the UKB, and recent reviews which highlight the frontal lobes as most susceptible to white matter deterioration(65, 66). Potentially, our findings provide additional evidence for the "last-in-first-out" retrogenesis hypothesis, which states that brain areas that develop slowest (such as the prefrontal cortex) are more vulnerable to negative ageing effects, such as degeneration(67). On the other hand, the observed WMM changes might simply map onto frontal GM areas which are most affected by normal ageing processes instead of areas with higher evolutionary expansion(67). Yet, as the age-dependence of frontal WM seems to be partly explained by cognitive ability(68), and cerebrovascular factors(69), frontal WM might be particularly interesting for further clinical dMRI examinations.

In accordance with our previous findings(3, 40), we found that the fornix is highly sensitive to ageing-related changes. The fornix – a C-shaped bundle of nerve-fibres that acts as a major output tract of the hippocampus – is a brain region implicated in various neurological and psychiatric disorders, such as, mild cognitive decline(70), impairment(71, 72), Parkinson's disease(73), arguably Alzheimer's disease (compare(74–78), and bipolar disorder(79). Moreover, the genetic architecture of fornix WMM is related to various neurological and psychiatric disorders(80). Bridging such genetics and imaging findings indicates that there are genetic underpinnings for the accelerated ageing of the fornix and other regions, which might explain pathology development. While our findings render the fornix as a promising marker of ageing, future studies need to explore this region as potential therapeutic target.

We identified general patterns of WM changes when applying diffusion approaches at voxelwise scale; namely, 1) global fractional anisotropy metric decreases, and 2) superior axial, radial and mean diffusivity metric increases in superior frontal brain regions, but 3) decreases in posterior regions, the brain stem and the cerebellum. Fractional anisotropy (DTI-FA and BRIA-microFA) decreases suggest different potential biological processes such as a seizing myelination or cell death across the brain (with microFA adding information on the fibre orientation coherence(81), Table 1, Supplement 25) or axonal degradation(82, 83) in combination with intra-axonal water fraction metric. The other diffusion metrics contain information on axial, radial and mean diffusivity and coherently suggested WMM degeneration with increasing age in superior frontal lobes and potential cell-swelling in posterior and subcortical regions where diffusivity decreases. This is congruent with multi-compartment metrics from BRIA, SMT, and WMTI (see Supplement 25). Potentially, the decreases in diffusivity metrics in posterior

and subcortical regions depict compensatory mechanisms, accounting for the frontal lobe deterioration. Additional examination of diffusion properties (e.g., using tractography) leveraging both single and multi-shell dMRI might provide further insight into the differential developments of white matter across the adult lifespan.

Finally, this study presents a comprehensive overview of WM association with PGRS of common psychiatric disorders and Alzheimer's Disease. Moreover, we differentiated between global and regional WM metrics at each time point and the annual rate of change, which led to different association patterns. However, the associations of both WM and annual WM-change with PGRS were non-significant when correcting for multiple comparisons. Previous results demonstrate small associations between WM and PGRS in the UKB for MDD(39), SCZ, BIP(84), and AD(38), as well as in the Adolescent Brain and Cognitive Development Study(24). For the further highlighting of key associations, we considered uncorrected p-values, examining the PGRS-WM associations as suggestive for potential associations of WM and genetic risk.

Surprisingly, the global annual rate of WMM change associated only with ADHD PGRS. WMM at each time point provided a more nuanced association pattern including ANX, OCD, and AD, similar to previous findings on AD PGRS-WM associations(38). Additionally, similarities between ANX and OCD PGRS associations with WMM might originate from large symptom overlap between these diagnoses(85). This association pattern was however not replicated in an independent cross-sectional portion of the UKB which instead outlined associations with ASD and BIP. Hence, whether these differential association patterns speak to sample-specific gene-brain relationships, or are simply noise due to a lack of statistical power (as a function of small effect sizes) requires follow-up with larger samples, more time points, and larger inter-scan-intervals.

Observing relationships of PGRS with both WMM and its annual rate of change in single regions highlight the brain stem, cerebral peduncle, and the limbic system as potential PGRS-association targets. Importantly, PGRS associations were orders of magnitude stronger for WMM rate of change than for cross-sectional metrics, which underlines the importance of examining the genetic underpinnings of WM in longitudinal data. Notably, areas outlined as most age-sensitive (the fornix and cerebral peduncle) were also the strongest related in their annual change to ANX, ADHD, OCD, and SCZ PGRS. Yet, more longitudinal research is needed to validate the presented findings. Genetic overlaps between fornix WMM and the listed disorders(80), as well as the involvement of the cerebellum, which is connected with the cortex via the cerebral peduncles, in various psychiatric disorders, gives additional insight into the role of genetic makeup for WMM development(86). Furthermore, the cerebral peduncles were particularly associated with AD, ADHD, and OCD PGRS, but also with ANX, BIP, MDD, and SCZ PGRS. This spatially specific pattern of PGRS associations emphasises the usefulness of regional investigations, due to highly spatially distributed influence of genetics. While the small effect sizes limit inferences on WMM-PGRS associations, the highlighted associations of PGRS and WMM *change* are worth further investigation.

There are several limitations to be mentioned in the context of this study. First, the age range was limited to individuals older than 40 years, allowing only for generalisations across mid-to-late adulthood. Future studies should consider large samples to cover the whole lifespan, particularly when the objective is to investigate WM ageing or to investigate generalisable associations of WM with genotypes and phenotypes. Second, the inter-scan interval was relatively short ($\bar{\Delta}$=2.44 years). Longer inter-scan intervals might reveal clearer information on accelerated WMM ageing processes at different ages and their genetic underpinnings. Longer intervals are also useful to examine the relationship of cognitive decline and WMM. Third, we used a relatively homogeneous, non-diverse sample including nearly exclusively white UK citizens, which limits the generalisability beyond white Northern Europeans and US Americans in their midlife to older age. Additionally, although the sample size was larger than in previous longitudinal WMM investigations, power was still limited to find PGRS associations (presenting small effects). The volunteer-based sampling of the UKB participants might additionally have introduced bias, reducing generalisability to the UK population(87). Yet, the imaging sample of the UKB shows an additional positive health bias (better physical and mental health) over the rest of the UKB sample(88), rendering this sub-sample as even less representative of the UK population. However, this might not necessarily be a disadvantage considering the objective of this study, which was to map WM change in healthy mid-to-late life adults and associate polygenic risk with the observed WM changes. Fourth, conservative corrections of the α-level using Bonferroni corrections might have led to false negatives, especially for the small observed effects of PGRS on the annual change in regional WMM where many associations were explored. Finally, the fornix, the key region of the presented regional WMM changes, is a region susceptible of free water contamination due to its closeness to the cerebrospinal fluid(2) and has previously been suspected of partial volume effects(62). Hence, findings on this region need to be interpreted carefully.

To conclude, our findings provide insight about short-term WM changes indicating degradation processes indicated by lower FA, kurtosis, intra-axonal water fraction, and higher diffusivity, free water, and extra axonal water fraction. These changes are associated with the de-myelination and structural disintegration across the adult brain as a consequence of ageing without strong or detectable cognitive decline. Demyelination and WM features degradation primarily affect frontal brain regions, whereas posterior regions as well as brain stem and cerebellum show opposing trends. Further investigations should focus on fornix WMM *changes* throughout the lifespan to investigate health and disease outcomes, and the role in such of the genetic architecture of the cerebellar peduncle in such white matter changes and disease devel-

opment.

AUTHOR CONTRIBUTIONS

Max Korbmacher: Study design, Software, Quality control, Formal analysis, Visualizations, Project administration, Writing—original draft, Writing—review & editing. Dennis van der Meer: Software, Writing – review & editing. Eli Eikefjord: Funding acquisition. Dani Beck: Writing—review & editing. Daniel E. Askeland-Gjerde: Writing—review & editing. Ole A. Andreassen: Writing—review & editing, Funding acquisition. Lars T. Westlye: Writing—review & editing, Funding acquisition. Ivan I. Maximov: Supervision, Study design, Software, Data preprocessing and quality control, Writing—review & editing, Funding acquisition.

CONFLICTS OF INTEREST

OAA has received a speaker's honorarium from Lundbeck and is a consultant to Coretechs.ai.

DATA AND CODE AVAILABILITY

All raw data are available from the UKB5 (www.ukbiobank.ac.uk). Analysis code and extended supplementary files will be made available at the time of publication at https://github.com/MaxKorbmacher/Long_Diffusion.

References

1. Sheelakumari Raghavan, Robert I Reid, Scott A Przybelski, Timothy G Lesnick, Jonathan Graff-Radford, Christopher G Schwarz, David S Knopman, Michelle M Mielke, Mary M Machulda, Ronald C Petersen, et al. Diffusion models reveal white matter microstructural changes with ageing, pathology and cognition. *Brain communications*, 3(2):fcab106, 2021.
2. Rafael Neto Henriques, Richard Henson, Marta Morgado Correia, et al. Unique information from common diffusion mri models about white-matter differences across the human adult lifespan. *arXiv preprint arXiv:2306.09942*, 2023.
3. Max Korbmacher, Ann Marie de Lange, Dennis van der Meer, Dani Beck, Eli Eikefjord, Arvid Lundervold, Ole A Andreassen, Lars T Westlye, and Ivan I Maximov. Brain-wide associations between white matter and age highlight the role of fornix microstructure in brain ageing. *Human brain mapping*, 44(10), 2023.
4. Ahmed Salih, Ilaria Boscolo Galazzo, Zahra Raisi-Estabragh, Elisa Rauseo, Polyxeni Gkontra, Steffen E Petersen, Karim Lekadir, André Altmann, Petia Radeva, and Gloria Menegaz. Brain age estimation at tract group level and its association with daily life measures, cardiac risk factors and genetic variants. *Scientific Reports*, 11(1):20563, 2021.
5. Oktay Agcaoglu, R Miller, Andy R Mayer, Kenneth Hugdahl, and Vince D Calhoun. Lateralization of resting state networks and relationship to age and gender. *Neuroimage*, 104: 310–325, 2015.
6. Dani Beck, Ann-Marie G de Lange, Ivan I Maximov, Geneviève Richard, Ole A Andreassen, Jan E Nordvik, and Lars T Westlye. 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, 2021.
7. Dani Beck, Lia Ferschmann, Niamh MacSweeney, Linn B Norbom, Thea Wiker, Eira Aksnes, Valerie Karl, Fanny Dégeilh, Madelene Holm, Kathryn L Mills, et al. Puberty differentially predicts brain maturation in male and female youth: A longitudinal abcd study. *Developmental Cognitive Neuroscience*, page 101261, 2023.
8. Stine K Krogsrud, Anders M Fjell, Christian K Tamnes, Håkon Grydeland, Lia Mork, Paulina Due-Tønnessen, Atle Bjørnerud, Cassandra Sampaio-Baptista, Jesper Andersson, Heidi Johansen-Berg, et al. Changes in white matter microstructure in the developing brain—a longitudinal diffusion tensor imaging study of children from 4 to 11 years of age. *Neuroimage*, 124:473–486, 2016.
9. Christian K Tamnes, David R Roalf, Anne-Lise Goddings, and Catherine Lebel. Diffusion mri of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Developmental cognitive neuroscience*, 33:161–175, 2018.
10. Catherine Lebel, Sarah Treit, and Christian Beaulieu. A review of diffusion mri of typical white matter development from early childhood to young adulthood. *NMR in Biomedicine*, 32(4):e3778, 2019.
11. Lars T Westlye, Kristine B Walhovd, Anders M Dale, Atle Bjørnerud, Paulina Due-Tønnessen, Andreas Engvig, Håkon Grydeland, Christian K Tamnes, Ylva Østby, and Anders M Fjell. Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cerebral cortex*, 20(9):2055–2068, 2010.
12. Rikka Kjelkenes, Thomas Wolfers, Dag Alnæs, Linn B Norbom, Irene Voldsbekk, Madelene Holm, Andreas Dahl, Pierre Berthet, Christian K Tamnes, Andre F Marquand, et al. Deviations from normative brain white and gray matter structure are associated with psychopathology in youth. *Developmental Cognitive Neuroscience*, 58:101173, 2022.

13. Moriah E Thomason and Paul M Thompson. Diffusion imaging, white matter, and psychopathology. *Annual review of clinical psychology*, 7:63–85, 2011.
14. Raza Sagarwala and Henry A Nasrallah. White matter pathology is shared across multiple psychiatric brain disorders: Is abnormal diffusivity a transdiagnostic biomarker for psychopathology? *Biomarkers in Neuropsychiatry*, 2:100010, 2020.
15. Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. Meta-analysis of gwas of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24. 32 and a significant overlap with schizophrenia. *Molecular autism*, 8:1–17, 2017.
16. Ditte Demontis, Raymond K Walters, Joanna Martin, Manuel Mattheisen, Thomas D Als, Esben Agerbo, Gísli Baldursson, Rich Belliveau, Jonas Bybjerg-Grauholm, Marie Bækvad-Hansen, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature genetics*, 51(1):63–75, 2019.
17. Paul D Arnold, Kathleen D Askland, Cristina Barlassina, Laura Bellodi, OJ Bienvenu, Donald Black, Michael Bloch, Helena Brentani, Christie L Burton, Beatriz Camarena, et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular psychiatry*, 23(5):1181–1181, 2018.
18. Takeshi Otowa, Karin Hek, Minyoung Lee, Enda M Byrne, Saira S Mirza, Michel G Nivard, Timothy Bigdeli, Steven H Aggen, Daniel Adkins, Aaron Wolen, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular psychiatry*, 21(10):1391–1399, 2016.
19. Naomi R Wray, Stephan Ripke, Manuel Mattheisen, Maciej Trzaskowski, Enda M Byrne, Abdel Abdellaoui, Mark J Adams, Esben Agerbo, Tracy M Air, Till MF Andlauer, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature genetics*, 50(5):668–681, 2018.
20. Douglas P Wightman, Iris E Jansen, Jeanne E Savage, Alexey A Shadrin, Shahram Bahrami, Dominic Holland, Arvid Rongve, Sigrid Børte, Bendik S Winsvold, Ole Kristian Drange, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for alzheimer's disease. *Nature genetics*, 53(9):1276–1282, 2021.
21. Niamh Mullins, Andreas J Forstner, Kevin S O'Connell, Brandon Coombes, Jonathan RI Coleman, Zhen Qiao, Thomas D Als, Tim B Bigdeli, Sigrid Børte, Julien Bryois, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature genetics*, 53(6):817–829, 2021.
22. Vassily Trubetskoy, Antonio F Pardiñas, Ting Qi, Georgia Panagiotaropoulou, Swapnil Awasthi, Tim B Bigdeli, Julien Bryois, Chia-Yen Chen, Charlotte A Dennison, Lynsey S Hall, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, 604(7906):502–508, 2022.
23. Jean-Charles Lambert, Carla A Ibrahim-Verbaas, Denise Harold, Adam C Naj, Rebecca Sims, Céline Bellenguez, Gyungah Jun, Anita L DeStefano, Joshua C Bis, Gary W Beecham, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for alzheimer's disease. *Nature genetics*, 45(12):1452–1458, 2013.
24. Sara Fernandez-Cabello, Dag Alnæs, Dennis van der Meer, Andreas Dahl, Madelene Holm, Rikka Kjelkenes, Ivan I Maximov, Linn B Norbom, Mads L Pedersen, Irene Voldsbekk, et al. Associations between brain imaging and polygenic scores of mental health and educational attainment in children aged 9–11. *NeuroImage*, 263:119611, 2022.
25. Dennis van der Meer and Tobias Kaufmann. Mapping the genetic architecture of cortical morphology through neuroimaging: progress and perspectives. *Translational Psychiatry*, 12 (1):447, 2022.
26. Dennis Van der Meer, Alexey A Shadrin, Kevin O'Connell, Francesco Bettella, Srdjan Djurovic, Thomas Wolfers, Dag Alnæs, Ingrid Agartz, Olav B Smeland, Ingrid Melle, et al. Boosting schizophrenia genetics by utilizing genetic overlap with brain morphology. *Biological psychiatry*, 92(4):291–298, 2022.
27. Els Fieremans, Jens H Jensen, and Joseph A Helpern. White matter characterization with diffusional kurtosis imaging. *Neuroimage*, 58(1):177–188, 2011.
28. Jens H Jensen, Joseph A Helpern, Anita Ramani, Hanzhang Lu, and Kyle Kaczynski. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 53(6):1432–1440, 2005.
29. Enrico Kaden, Nathaniel D Kelm, Robert P Carson, Mark D Does, and Daniel C Alexander. Multi-compartment microscopic diffusion imaging. *NeuroImage*, 139:346–359, 2016.
30. Enrico Kaden, Frithjof Kruggel, and Daniel C Alexander. Quantitative mapping of the per-axon diffusion coefficients in brain white matter. *Magnetic resonance in medicine*, 75(4): 1752–1763, 2016.
31. Marco Reisert, Elias Kellner, Bibek Dhital, Juergen Hennig, and Valerij G Kiselev. Disentangling micro from mesostructure by diffusion MRI: a Bayesian approach. *Neuroimage*, 147: 964–975, 2017.
32. Miguel Ángel Araque Caballero, Marc Suárez-Calvet, Marco Duering, Nicolai Franzmeier, Tammie Benzinger, Anne M Fagan, Randall J Bateman, Clifford R Jack, Johannes Levin, Martin Dichgans, et al. White matter diffusion alterations precede symptom onset in autosomal dominant alzheimer's disease. *Brain*, 141(10):3065–3080, 2018.
33. Peter J Basser, James Mattiello, and Denis LeBihan. Mr diffusion tensor spectroscopy and imaging. *Biophysical journal*, 66(1):259–267, 1994.
34. Adam R Pines, Matthew Cieslak, Bart Larsen, Graham L Baum, Philip A Cook, Azeez Adebimpe, Diego G Dávila, Mark A Elliott, Robert Jirsaraie, Kristin Murtha, et al. Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Developmental cognitive neuroscience*, 43:100788, 2020.
35. Thomas R Barrick, Rebecca A Charlton, Chris A Clark, and Hugh S Markus. White matter structural decline in normal ageing: a prospective longitudinal study using tract-based spatial statistics. *Neuroimage*, 51(2):565–577, 2010.
36. Claire E Sexton, Kristine B Walhovd, Andreas B Storsve, Christian K Tamnes, Lars T Westlye, Heidi Johansen-Berg, and Anders M Fjell. Accelerated changes in white matter microstructure during aging: a longitudinal diffusion tensor imaging study. *Journal of Neuroscience*, 34(46):15425–15436, 2014.
37. Fidel Alfaro-Almagro, Mark Jenkinson, Neal K Bangerter, Jesper LR Andersson, Ludovica Griffanti, Gwenaëlle Douaud, Stamatios N Sotiropoulos, Saad Jbabdi, Moises Hernandez-Fernandez, Emmanuel Vallee, et al. Image processing and quality control for the first 10,000

brain imaging datasets from uk biobank. *Neuroimage*, 166:400–424, 2018.

38. Xiao-Yu He, Bang-Sheng Wu, Kevin Kuo, Wei Zhang, Qing Ma, Shi-Tong Xiang, Yu-Zhu Li, Zi-yi Wang, Qiang Dong, Jian-Feng Feng, et al. Association between polygenic risk for alzheimer's disease and brain structure in children and adults. *Alzheimer's Research & Therapy*, 15(1):1–12, 2023.

39. Miruna C Barbu, Yanni Zeng, Xueyi Shen, Simon R Cox, Toni-Kim Clarke, Jude Gibson, Mark J Adams, Mandy Johnstone, Chris S Haley, Stephen M Lawrie, et al. Association of whole-genome and netrin1 signaling pathway–derived polygenic risk scores for major depressive disorder and white matter microstructure in the uk biobank. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 4(1):91–100, 2019.

40. Max Korbmacher, Dennis van der Meer, Dani Beck, Ann-Marie G de Lange, Eli Eikefjord, Arvid Lundervold, Ole Andreassen, Lars T Westlye, and Ivan I Maximov. Brain asymmetries from midlife to old adulthood and hemispheric brain age. *bioRxiv*, pages 2023–08, 2023.

41. Ivan I Maximov, Dennis van der Meer, Ann-Marie G de Lange, Tobias Kaufmann, Alexey Shadrin, Oleksandr Frei, Thomas Wolfers, and Lars T Westlye. Fast qualitY conTrol meThod foR derIved diffUsion Metrics (YTTRIUM) in big data analysis: UK Biobank 18,608 example. *Human brain mapping*, 42(10):3141–3155, 2021.

42. Karla L Miller, Fidel Alfaro-Almagro, Neal K Bangerter, David L Thomas, Essa Yacoub, Junqian Xu, Andreas J Bartsch, Saad Jbabdi, Stamatios N Sotiropoulos, Jesper LR Andersson, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature neuroscience*, 19(11):1523–1536, 2016.

43. Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, 2015.

44. Jelle Veraart, Els Fieremans, and Dmitry S Novikov. Diffusion MRI noise mapping using random matrix theory. *Magnetic resonance in medicine*, 76(5):1582–1593, 2016.

45. Elias Kellner, Bibek Dhital, Valerij G Kiselev, and Marco Reisert. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic resonance in medicine*, 76(5):1574–1581, 2016.

46. Jesper LR Andersson and Stamatios N Sotiropoulos. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage*, 125: 1063–1078, 2016.

47. Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy EJ Behrens, Heidi Johansen-Berg, Peter R Bannister, Marilena De Luca, Ivana Drobnjak, David E Flitney, et al. Advances in functional and structural mr image analysis and implementation as fsl. *Neuroimage*, 23:S208–S219, 2004.

48. Mark Jenkinson, Christian F Beckmann, Timothy EJ Behrens, Mark W Woolrich, and Stephen M Smith. FSL. *Neuroimage*, 62(2):782–790, 2012.

49. Stephen M Smith, Mark Jenkinson, Heidi Johansen-Berg, Daniel Rueckert, Thomas E Nichols, Clare E Mackay, Kate E Watkins, Olga Ciccarelli, M Zaheer Cader, Paul M Matthews, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage*, 31(4):1487–1505, 2006.

50. Stephen M Smith. Fast robust automated brain extraction. *Human brain mapping*, 17(3): 143–155, 2002.

51. S Mori, S Wakana, LM Nagae-Poetscher, and PCM Van Zijl. MRI atlas of human white matter. *American Journal of Neuroradiology*, 27(6):1384, 2006.

52. Kegang Hua, Jiangyang Zhang, Setsu Wakana, Hangyi Jiang, Xin Li, Daniel S Reich, Peter A Calabresi, James J Pekar, Peter CM van Zijl, and Susumu Mori. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *Neuroimage*, 39(1):336–347, 2008.

53. Shing Wan Choi and Paul F O'Reilly. Prsice-2: Polygenic risk score software for biobank-scale data. *Gigascience*, 8(7):giz082, 2019.

54. Kaylyn Clark, Yuk Yee Leung, Wan-Ping Lee, Benjamin Voight, and Li-San Wang. Polygenic risk scores in alzheimer's disease genetics: methodology, applications, inclusion, and diversity. *Journal of Alzheimer's Disease*, 89(1):1–12, 2022.

55. Jack Euesden, Cathryn M Lewis, and Paul F O'reilly. Prsice: polygenic risk score software. *Bioinformatics*, 31(9):1466–1468, 2015.

56. Tyler J VanderWeele and Maya B Mathur. Some desirable properties of the bonferroni correction: is the bonferroni correction really so bad? *American journal of epidemiology*, 188(3):617–618, 2019.

57. Stephen M Smith and Thomas E Nichols. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, 44(1):83–98, 2009.

58. Live Eikenes, Eelke Visser, Torgil Vangberg, and Asta K Håberg. Both brain size and biological sex contribute to variation in white matter microstructure in middle-aged healthy adults. *Human Brain Mapping*, 44(2):691–709, 2023.

59. Jenna L Merenstein, María M Corrada, Claudia H Kawas, and Ilana J Bennett. Age affects white matter microstructure and episodic memory across the older adult lifespan. *Neurobiology of aging*, 106:282–291, 2021.

60. Kurt G Schilling, Derek Archer, Fang-Cheng Yeh, Francois Rheault, Leon Y Cai, Colin Hansen, Qi Yang, Karthik Ramdass, Andrea T Shafer, Susan M Resnick, et al. Aging and white matter microstructure and macrostructure: a longitudinal multi-site diffusion mri study of 1218 participants. *Brain Structure and Function*, 227(6):2111–2125, 2022.

61. Andrew R Bender, Manuel C Völkle, and Naftali Raz. Differential aging of cerebral white matter in middle-aged and older adults: a seven-year follow-up. *Neuroimage*, 125:74–83, 2016.

62. Yisu Yang, Kurt Schilling, Niranjana Shashikumar, Varuna Jasodanand, Elizabeth E Moore, Kimberly R Pechman, Murat Bilgel, Lori L Beason-Held, Yang An, Andrea Shafer, et al. White matter microstructural metrics are sensitively associated with clinical staging in alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 15(2):e12425, 2023.

63. Haichao Zhao, Wei Wen, Jian Cheng, Jiyang Jiang, Nicole Kochan, Haijun Niu, Henry Brodaty, Perminder Sachdev, and Tao Liu. An accelerated degeneration of white matter microstructure and networks in the nondemented old–old. *Cerebral Cortex*, 33(8):4688–4698, 2023.

64. Wen-Yih Isaac Tseng, Yung-Chin Hsu, Chang-Le Chen, Yun-Jing Kang, Te-Wei Kao, Pin-Yu Chen, and Gordon D Waiter. Microstructural differences in white matter tracts across middle to late adulthood: a diffusion mri study on 7167 uk biobank participants. *Neurobiology of Aging*, 98:160–172, 2021.

65. Faith M Gunning-Dixon, Adam M Brickman, Janice C Cheng, and George S Alexopoulos. Aging of cerebral white matter: a review of mri findings. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 24(2):109–117, 2009.

66. M Ethan MacDonald and G Bruce Pike. Mri of healthy brain aging: A review. *NMR in Biomedicine*, 34(9):e4564, 2021.

67. Anders M Fjell, Linda McEvoy, Dominic Holland, Anders M Dale, Kristine B Walhovd, Alzheimer's Disease Neuroimaging Initiative, et al. What is normal in normal aging? effects of aging, amyloid and alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in neurobiology*, 117:20–40, 2014.

68. Jan Willem Koten Jr, Karl Koschutnig, and Guilherme Wood. An attempt to model the causal structure behind white matter aging and cognitive decline. *Scientific Reports*, 13(1):10883, 2023.

69. Prashanthi Vemuri, Jonathan Graff-Radford, Timothy G Lesnick, Scott A Przybelski, Robert I Reid, Ashritha L Reddy, Val J Lowe, Michelle M Mielke, Mary M Machulda, Ronald C Petersen, et al. White matter abnormalities are key components of cerebrovascular disease impacting cognitive decline. *Brain Communications*, 3(2):fcab076, 2021.

70. Andrea T Shafer, Owen A Williams, Evian Perez, Yang An, Bennett A Landman, Luigi Ferrucci, and Susan M Resnick. Accelerated decline in white matter microstructure in subsequently impaired older adults and its relationship with cognitive decline. *Brain communications*, 4(2):fcac051, 2022.

71. Yumeng Ju, Wenwen Ou, Jingzhi Su, Christopher L Averill, Jin Liu, Mi Wang, Zhen Wang, Yan Zhang, Bangshan Liu, Lingjiang Li, et al. White matter microstructural alterations in posttraumatic stress disorder: An roi and whole-brain based meta-analysis. *Journal of affective disorders*, 266:655–670, 2020.

72. Junhong Yu, Charlene LM Lam, and Tatia MC Lee. White matter microstructural abnormalities in amnestic mild cognitive impairment: A meta-analysis of whole-brain and roi-based studies. *Neuroscience & Biobehavioral Reviews*, 83:405–416, 2017.

73. Xia Wei, Chunyan Luo, Qian Li, Na Hu, Yuan Xiao, Nian Liu, Su Lui, and Qiyong Gong. White matter abnormalities in patients with parkinson's disease: a meta-analysis of diffusion tensor imaging using tract-based spatial statistics. *Frontiers in aging neuroscience*, 12: 610962, 2021.

74. John M Ringman, Joseph O'Neill, Daniel Geschwind, Luis Medina, Liana G Apostolova, Yaneth Rodriguez, Barbara Schaffer, Arousiak Varpetian, Benjamin Tseng, Freddy Ortiz, et al. Diffusion tensor imaging in preclinical and presymptomatic carriers of familial alzheimer's disease mutations. *Brain*, 130(7):1767–1776, 2007.

75. X Li, E Westman, AK Ståhlbom, S Thordardottir, Ove Almkvist, K Blennow, L-O Wahlund, and C Graff. White matter changes in familial alzheimer's disease. *Journal of internal medicine*, 278(2):211–218, 2015.

76. Nhat Trung Doan, Andreas Engvig, Karin Persson, Dag Alnæs, Tobias Kaufmann, Jaroslav Rokicki, Aldo Córdova-Palomera, Torgeir Moberget, Anne Brækhus, Maria Lage Barca, et al. Dissociable diffusion mri patterns of white matter microstructure and connectivity in alzheimer's disease spectrum. *Scientific reports*, 7(1):45131, 2017.

77. Fang Ji, Ofer Pasternak, Kwun Kei Ng, Joanna Su Xian Chong, Siwei Liu, Liwen Zhang, Hee Youn Shim, Yng Miin Loke, Boon Yeow Tan, Narayanaswamy Venketasubramanian, et al. White matter microstructural abnormalities and default network degeneration are associated with early memory deficit in alzheimer's disease continuum. *Scientific reports*, 9 (1):4749, 2019.

78. Lotte GM Cremers, Frank J Wolters, Marius de Groot, M Kamran Ikram, Aad van Der Lugt, Wiro J Niessen, Meike W Vernooij, and M Arfan Ikram. Structural disconnectivity and the risk of dementia in the general population. *Neurology*, 95(11):e1528–e1537, 2020.

79. Pauline Favre, Melissa Pauling, Jacques Stout, Franz Hozer, Samuel Sarrazin, Christoph Abé, Martin Alda, Clara Alloza, Silvia Alonso-Lana, Ole A Andreassen, et al. Widespread white matter microstructural abnormalities in bipolar disorder: evidence from mega-and meta-analyses across 3033 individuals. *Neuropsychopharmacology*, 44(13):2285–2293, 2019.

80. Ya-Nan Ou, Yi-Jun Ge, Bang-Sheng Wu, Yi Zhang, Yu-Chao Jiang, Kevin Kuo, Liu Yang, Lan Tan, Jian-Feng Feng, Wei Cheng, et al. The genetic architecture of fornix white matter microstructure and their involvement in neuropsychiatric disorders. *Translational Psychiatry*, 13(1):180, 2023.

81. Filip Szczepankiewicz, Samo Lasič, Danielle van Westen, Pia C Sundgren, Elisabet Englund, Carl-Fredrik Westin, Freddy Ståhlberg, Jimmy Lätt, Daniel Topgaard, and Markus Nilsson. Quantification of microscopic diffusion anisotropy disentangles effects of orientation dispersion from microstructure: applications in healthy volunteers and in brain tumors. *Neuroimage*, 104:241–252, 2015.

82. Elan J Grossman, Ivan I Kirov, Oded Gonen, Dmitry S Novikov, Matthew S Davitz, Yvonne W Lui, Robert I Grossman, Matilde Inglese, and Els Fieremans. N-acetyl-aspartate levels correlate with intra-axonal compartment parameters from diffusion mri. *Neuroimage*, 118: 334–343, 2015.

83. Andreana Benitez, Jens H Jensen, Kathryn Thorn, Siddhartha Dhiman, Stephanie Fountain-Zaragoza, William J Rieter, Maria Vittoria Spampinato, Eric D Hamlett, Paul J Nietert, Maria de Fatima Falangola, et al. Greater diffusion restriction in white matter in preclinical alzheimer disease. *Annals of Neurology*, 91(6):864–877, 2022.

84. Lianne M Reus, Xueyi Shen, Jude Gibson, Ella Wigmore, Lannie Ligthart, Mark J Adams, Gail Davies, Simon R Cox, Saskia P Hagenaars, Mark E Bastin, et al. Association of polygenic risk for major psychiatric illness with subcortical volumes and white matter integrity in uk biobank. *Scientific reports*, 7(1):42140, 2017.

85. Miriam K. Forbes, Bryan Neo, Omid Mohamed Nezami, Eiko I. Fried, Katherine Faure, Brier Michelsen, Maddison Twose, and Mark Dras. Elemental psychopathology: distilling constituent symptoms and patterns of repetition in the diagnostic criteria of the dsm-5. *Psych Med*, page 1–9, 2023. doi: 10.1017/S0033291723002544.

86. Joseph R Phillips, Doaa H Hewedi, Abeer M Eissa, and Ahmed A Moustafa. The cerebellum and psychiatric disorders. *Frontiers in public health*, page 66, 2015.

87. Tabea Schoeler, Doug Speed, Eleonora Porcu, Nicola Pirastu, Jean-Baptiste Pingault, and Zoltán Kutalik. Participation bias in the uk biobank distorts genetic associations and downstream analyses. *Nature Human Behaviour*, pages 1–12, 2023.

88. Donald M Lyall, Terry Quinn, Laura M Lyall, Joey Ward, Jana J Anderson, Daniel J Smith, William Stewart, Rona J Strawbridge, Mark ES Bailey, and Breda Cullen. Quantifying bias in psychological and physical health in the uk biobank imaging sub-sample. *Brain communications*, 4(3):fcac119, 2022.

89. Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13 (4):600–612, 2004.

90. Dominique Brunet, Edward R Vrscay, and Zhou Wang. On the mathematical properties of the structural similarity index. *IEEE Transactions on Image Processing*, 21(4):1488–1499, 2011.

Supplementary Note 1: The utilized white matter quality control pipeline.

In brief, YTTRIUM (41) converts dMRI scalar metrics into 2D format, using a structural similarity (89, 90) extension of each scalar map to their mean image in order to create a 2D distribution of image and diffusion parameters. These quality assessments are based on a 2-step clustering algorithm applied to identify subjects located out of the main distribution. Additionally, rows including impossible values, such as diffusion coefficients d: $0 < d < 4$ $\mu m^2 \cdot ms$, kurtosis values K $0 < K < 3$, and FA values $0 < FA < 1$ were excluded. QC of the mean skeleton values rendered $N = 622$ datasets of the BRIA metric extra-axonal radial diffusivity (DRADextra) as such impossible values, and examining regional and tract averages $N = 643$ of the WMTI metric axial extra-axonal diffusivity (axEAD). Due to the relatively large share of these outliers on the total sample, we excluded these metrics from the analyses.

Supplementary Note 2: Description of white matter features by diffusion approaches.

Diffusion Approach	Metrics
Bayesian Rotationally Invariant Approach (BRIA) (31)	intra-axonal axial diffusivity (DAX intra)
	extra-axonal radial diffusivity (DRAD extra)*
	microscopic fractional anisotropy (micro FA)
	extra-axonal axial diffusivity (DAX extra)
	intra-axonal water fraction (V intra)
	extra-axonal water fraction (V extra)
	cerebrospinal fluid fraction (vCSF)
	microscopical axial diffusivity (micro AX)
	microscopic radial diffusivity (micro RD)
	microscopical apparent diffusion coefficient (micro ADC)
Diffusion Kurtosis Imaging (DKI) (27, 28)	mean kurtosis (MK)
	radial kurtosis (RK)
	axial kurtosis (AK)
Diffusion Tensor Imaging (DTI) (33)	fractional anisotropy (FA)
	axial diffusivity (AD)
	mean diffusivity (MD)
	radial diffusivity (RD)
Spherical Mean Technique (SMT) (30)	fractional anisotropy (SMT FA)
	mean diffusivity (SMT md)
	transverse diffusion coefficient (SMT trans)
	longitudinal diffusion coefficient (SMT long)
Multi-compartment Spherical Mean Technique (SMTmc) (29)	extra-neurite microscopic mean diffusivity (SMTmc extra md)
	extra-neurite transverse microscopic diffusivity (SMTmc extra trans)
	mc SMTdiffusion coefficient (SMT mcd)
	intra-neurite volume fraction (SMTmc intra)
White Matter Tract Integrity (WMTI) (27)	axonal water fraction (AWF)
	radial extra-axonal diffusivity (radEAD)
	axial extra-axonal diffusivity (axEAD)*

*Note that Drad extra and axEAD were excluded from the analyses as a significant portion of the produced metrics did not pass our quality control procedure (41).

Supplementary Note 3: Effect sizes for unadjusted microstructure changes between time points
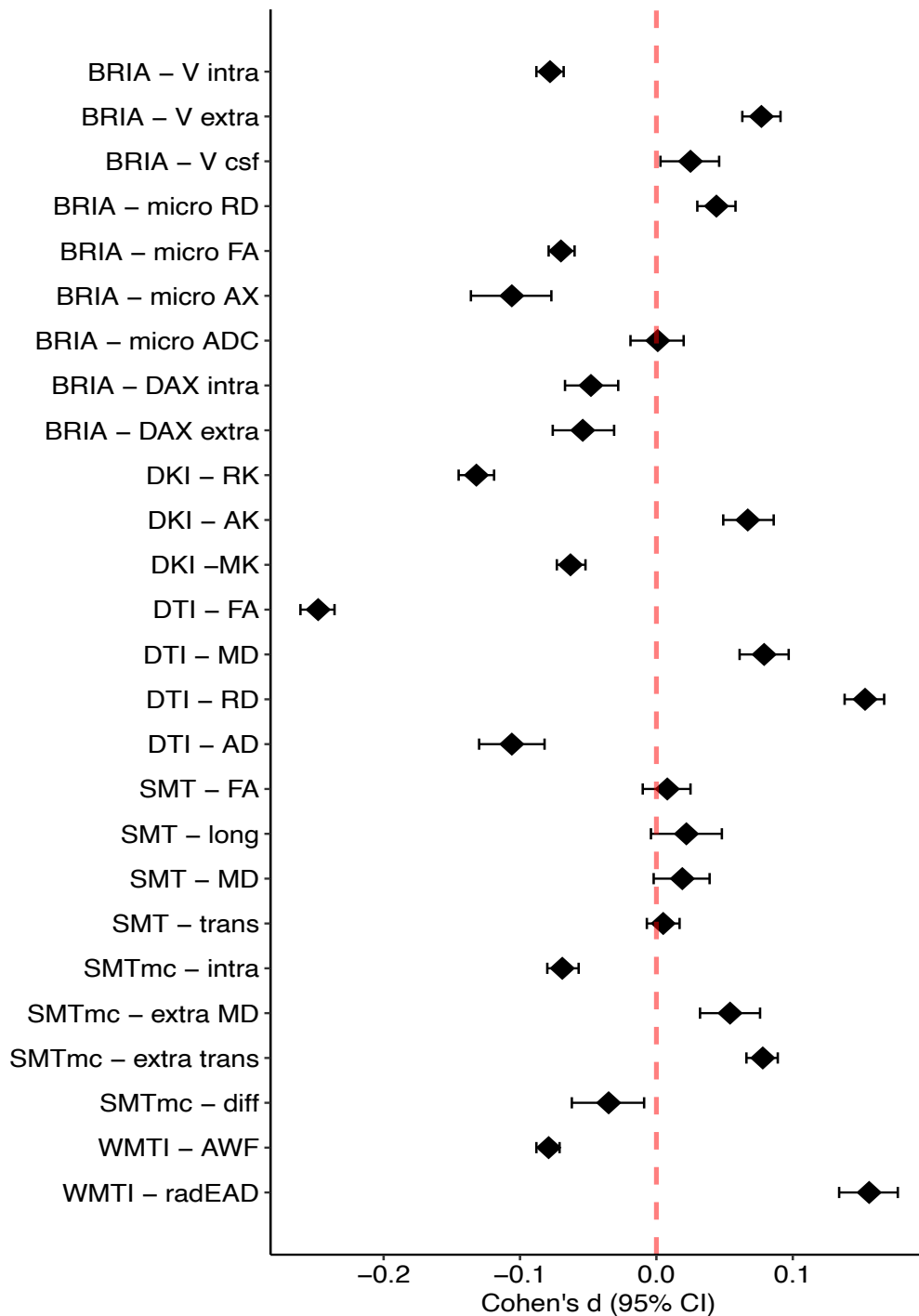
<0.001

Supplementary Note 4: Test statistics for time point differences for global diffusion metrics.

The table presents mean (M), standard deviation (SD) for each time point (TP), T-statistic (T), uncorrected p-value (p),
Bonferroni-corrected p-value p_{Bonf}, and Cohen's d with 95% CI as effect size.

Metric	M_{TP1}	SD_{TP1}	M_{TP2}	SD_{TP2}	T	p	p_{adj}	Cohens's d	L 95% CI	U 95% CI
BRIA - V intra	0.499	0.027	0.497	0.028	16.051	<0.001	<0.001	-0.078	-0.088	-0.068
BRIA - V extra	0.414	0.025	0.416	0.026	-10.937	<0.001	<0.001	0.077	0.063	0.091
BRIA - V csf	0.087	0.008	0.087	0.009	-2.254	0.024	0.631	0.025	0.003	0.046
BRIA - micro RD	0.524	0.038	0.526	0.039	-6.166	<0.001	<0.001	0.044	0.03	0.058
BRIA - micro FA	0.64	0.019	0.638	0.02	13.977	<0.001	<0.001	-0.07	-0.079	-0.06
BRIA - micro AX	1.899	0.032	1.896	0.032	7.132	<0.001	<0.001	-0.106	-0.136	-0.077
BRIA - micro ADC	0.982	0.032	0.982	0.033	-0.082	0.935	1	0.001	-0.019	0.02
BRIA - DAX intra	2.196	0.035	2.194	0.036	4.825	<0.001	<0.001	-0.048	-0.067	-0.028
BRIA - DAX extra	1.199	0.024	1.198	0.025	4.747	<0.001	<0.001	-0.054	-0.076	-0.031
DKI - RK	1.467	0.072	1.458	0.074	19.808	<0.001	<0.001	-0.132	-0.145	-0.119
DKI - AK	0.777	0.021	0.778	0.022	-7.124	<0.001	<0.001	0.067	0.049	0.086
DKI -MK	1.058	0.038	1.056	0.039	11.5	<0.001	<0.001	-0.063	-0.073	-0.052
DTI - FA	0.463	0.018	0.458	0.019	40.372	<0.001	<0.001	-0.248	-0.261	-0.236
DTI - MD	0.886	0.028	0.888	0.029	-8.679	<0.001	<0.001	0.079	0.061	0.097
DTI - RD	0.64	0.031	0.645	0.033	-20.295	<0.001	<0.001	0.153	0.138	0.167
DTI - AD	1.377	0.028	1.374	0.028	8.761	<0.001	<0.001	-0.106	-0.13	-0.082
SMT - FA	0.935	0.007	0.935	0.007	-0.844	0.399	1	0.008	-0.01	0.025
SMT - long	0.003	0	0.003	0	-1.652	0.099	1	0.022	-0.004	0.048
SMT - MD	0.001	0	0.001	0	-1.787	0.074	1	0.019	-0.002	0.039
SMT - trans	0	0	0	0	-0.778	0.436	1	0.005	-0.007	0.017
SMTmc - intra	0.592	0.03	0.59	0.031	11.449	<0.001	<0.001	-0.069	-0.08	-0.057
SMTmc - extra MD	0.001	0	0.001	0	-4.781	<0.001	<0.001	0.054	0.032	0.076
SMTmc - extra trans	0.001	0	0.001	0	-12.994	<0.001	<0.001	0.078	0.066	0.089
SMTmc - diff	0.002	0	0.002	0	2.6	0.009	0.244	-0.035	-0.062	-0.009
WMTI - AWF	0.39	0.014	0.388	0.014	18.187	<0.001	<0.001	-0.079	-0.088	-0.071
WMTI - radEAD	0.973	0.034	0.978	0.036	-14.445	<0.001	<0.001	0.156	0.134	0.177

Supplementary Note 5: Effects of age, time point, sex and the age-sex interaction on global diffusion metrics in LMER.

Standardized β values $\pm$ Standard Error indicating the fixed effects of (a) age, (b) time point, (c) sex, and (d) sex×age on global diffusion metrics. The colour indicates the metric's corresponding diffusion approach.

Supplementary Note 6: Test statistics for the effect of age on global diffusion metrics in LMER.

β_{std} refers to the standardized slopes and SE to the standard error.

Metric	β_{std}	SE	p	$p_{Bonferroni}$
BRIA - V intra	-0.037	0.003	<.001	<.001
BRIA - V extra	0.025	0.003	<.001	<.001
BRIA - V csf	0.066	0.003	<.001	<.001
BRIA - micro RD	0.063	0.003	<.001	<.001
BRIA - micro FA	-0.053	0.003	<.001	<.001
BRIA - micro AX	0.033	0.003	<.001	<.001
BRIA - micro ADC	0.061	0.003	<.001	<.001
BRIA - DAX intra	0.043	0.003	<.001	<.001
BRIA - DAX extra	0.047	0.003	<.001	<.001
DKI - RK	-0.042	0.003	<.001	<.001
DKI - AK	-0.04	0.003	<.001	<.001
DKI -MK	-0.035	0.003	<.001	<.001
DTI - FA	-0.063	0.003	<.001	<.001
DTI - MD	0.062	0.003	<.001	<.001
DTI - RD	0.066	0.003	<.001	<.001
DTI - AD	0.038	0.003	<.001	<.001
SMT - FA	-0.04	0.003	<.001	<.001
SMT - long	0.048	0.003	<.001	<.001
SMT - MD	0.056	0.003	<.001	<.001
SMT - trans	0.045	0.003	<.001	<.001
SMTmc - intra	-0.035	0.003	<.001	<.001
SMTmc - extra MD	0.058	0.003	<.001	<.001
SMTmc - extra trans	0.05	0.003	<.001	<.001
SMTmc - diff	0.008	0.003	0.01	0.269
WMTI - AWF	-0.042	0.003	<.001	<.001
WMTI - radEAD	0.067	0.003	<.001	<.001

Supplementary Note 7: The effect of time point on global diffusion metrics in mixed linear models.

Metric	β_{std}	SE	p	$p_{Bonferroni}$
BRIA - V intra	-0.06	0.005	<.001	<.001
BRIA - V extra	0.039	0.007	<.001	<.001
BRIA - V csf	0.11	0.009	<.001	<.001
BRIA - micro RD	0.102	0.006	<.001	<.001
BRIA - micro FA	-0.084	0.005	<.001	<.001
BRIA - micro AX	0.06	0.011	<.001	<.001
BRIA - micro ADC	0.102	0.008	<.001	<.001
BRIA - DAX intra	0.073	0.008	<.001	<.001
BRIA - DAX extra	0.079	0.009	<.001	<.001
DKI - RK	-0.064	0.006	<.001	<.001
DKI - AK	-0.065	0.008	<.001	<.001
DKI -MK	-0.054	0.006	<.001	<.001
DTI - FA	-0.107	0.006	<.001	<.001
DTI - MD	0.103	0.008	<.001	<.001
DTI - RD	0.111	0.007	<.001	<.001
DTI - AD	0.064	0.009	<.001	<.001
SMT - FA	-0.063	0.008	<.001	<.001
SMT - long	0.08	0.01	<.001	<.001
SMT - MD	0.093	0.008	<.001	<.001
SMT - trans	0.074	0.006	<.001	<.001
SMTmc - intra	-0.058	0.006	<.001	<.001
SMTmc - extra MD	0.097	0.009	<.001	<.001
SMTmc - extra trans	0.084	0.006	<.001	<.001
SMTmc - diff	0.018	0.01	0.082	1
WMTI - AWF	-0.067	0.005	<.001	<.001
WMTI - radEAD	0.113	0.009	<.001	<.001

Supplementary Note 8: Mixed linear model diagnostics

The following table shows the model diagnostics when predicting diffusion metrics from age, sex, age×sex, and time point, while accounting for the participants nested in scanner site as random intercept.

Metric	AIC	BIC	Conditional R2	Marginal R2
BRIA - V intra	7618.31	7677.58	0.97	0.07
BRIA - V extra	9655.94	9715.21	0.93	0.03
BRIA - V csf	11005.42	11064.69	0.84	0.28
BRIA - micro RD	9178.08	9237.35	0.93	0.21
BRIA - micro FA	7516.52	7575.80	0.96	0.14
BRIA - micro AX	13016.03	13075.30	0.70	0.13
BRIA - micro ADC	10798.61	10857.88	0.86	0.22
BRIA - DAX intra	11163.76	11223.03	0.87	0.11
BRIA - DAX extra	11683.87	11743.15	0.83	0.14
DKI - RK	9232.04	9291.32	0.94	0.09
DKI - AK	11036.27	11095.54	0.88	0.08
DKI -MK	8242.74	8302.02	0.96	0.06
DTI - FA	8388.40	8447.67	0.95	0.20
DTI - MD	10352.80	10412.07	0.89	0.22
DTI - RD	9350.35	9409.62	0.92	0.23

Supplementary Note 9: The effect of sex on global diffusion metrics in LMER.

Metric	β_{std}	SE	p	$p_{Bonferroni}$
BRIA - V intra	-0.363	0.2	0.069	1
BRIA - V extra	0.299	0.245	0.222	1
BRIA - V csf	0.47	0.248	0.059	1
BRIA - micro RD	0.564	0.232	0.015	0.386
BRIA - micro FA	-0.642	0.199	0.001	0.034
BRIA - micro AX	-0.072	0.276	0.793	1
BRIA - micro ADC	0.425	0.253	0.093	1
BRIA - DAX intra	0.155	0.267	0.561	1
BRIA - DAX extra	0.135	0.269	0.616	1
DKI - RK	-0.785	0.235	0.001	0.022
DKI - AK	-0.323	0.267	0.227	1
DKI -MK	-0.558	0.214	0.009	0.237
DTI - FA	-0.106	0.218	0.626	1
DTI - MD	0.386	0.249	0.121	1
DTI - RD	0.282	0.234	0.229	1
DTI - AD	0.293	0.268	0.274	1
SMT - FA	-0.611	0.264	0.021	0.538
SMT - long	0.206	0.271	0.446	1
SMT - MD	0.326	0.261	0.211	1
SMT - trans	0.335	0.226	0.138	1
SMTmc - intra	-0.262	0.225	0.245	1
SMTmc - extra MD	0.292	0.262	0.265	1
SMTmc - extra trans	0.202	0.22	0.359	1
SMTmc - diff	-0.267	0.288	0.353	1
WMTI - AWF	-0.511	0.185	0.006	0.149
WMTI - radEAD	0.121	0.255	0.635	1

Supplementary Note 10: The effect of the sex-age interaction on global diffusion metrics in LMER.

Metric	β_{std}	SE	p	$p_{Bonferroni}$
BRIA - V intra	0.014	0.038	0.708	1
BRIA - V extra	0.152	0.039	<0.001	0.003
BRIA - V csf	-0.512	0.033	<0.001	<0.001
BRIA - micro RD	-0.328	0.035	<0.001	<0.001
BRIA - micro FA	0.197	0.037	<0.001	<0.001
BRIA - micro AX	-0.531	0.035	<0.001	<0.001
BRIA - micro ADC	-0.429	0.034	<0.001	<0.001
BRIA - DAX intra	-0.282	0.037	<0.001	<0.001
BRIA - DAX extra	-0.378	0.036	<0.001	<0.001
DKI - RK	0.193	0.038	<0.001	<0.001
DKI - AK	-0.133	0.038	<0.001	0.012
DKI -MK	0.085	0.039	0.028	0.737
DTI - FA	-0.025	0.036	0.491	1
DTI - MD	-0.383	0.035	<0.001	<0.001
DTI - RD	-0.241	0.035	<0.001	<0.001
DTI - AD	-0.619	0.035	<0.001	<0.001
SMT - FA	-0.038	0.038	0.314	1
SMT - long	-0.465	0.035	<0.001	<0.001
SMT - MD	-0.39	0.035	<0.001	<0.001
SMT - trans	-0.062	0.037	0.1	1
SMTmc - intra	-0.038	0.039	0.327	1
SMTmc - extra MD	-0.404	0.034	<0.001	<0.001
SMTmc - extra trans	-0.143	0.037	<0.001	0.003
SMTmc - diff	-0.443	0.037	<0.001	<0.001
WMTI - AWF	0.039	0.038	0.302	1
WMTI - radEAD	-0.278	0.034	<0.001	<0.001

Supplementary Note 11: The effect of the sex on regional WMM metrics.

Increases refer to higher WMM values for males compared to females and decrease to lower WMM values, respectively.
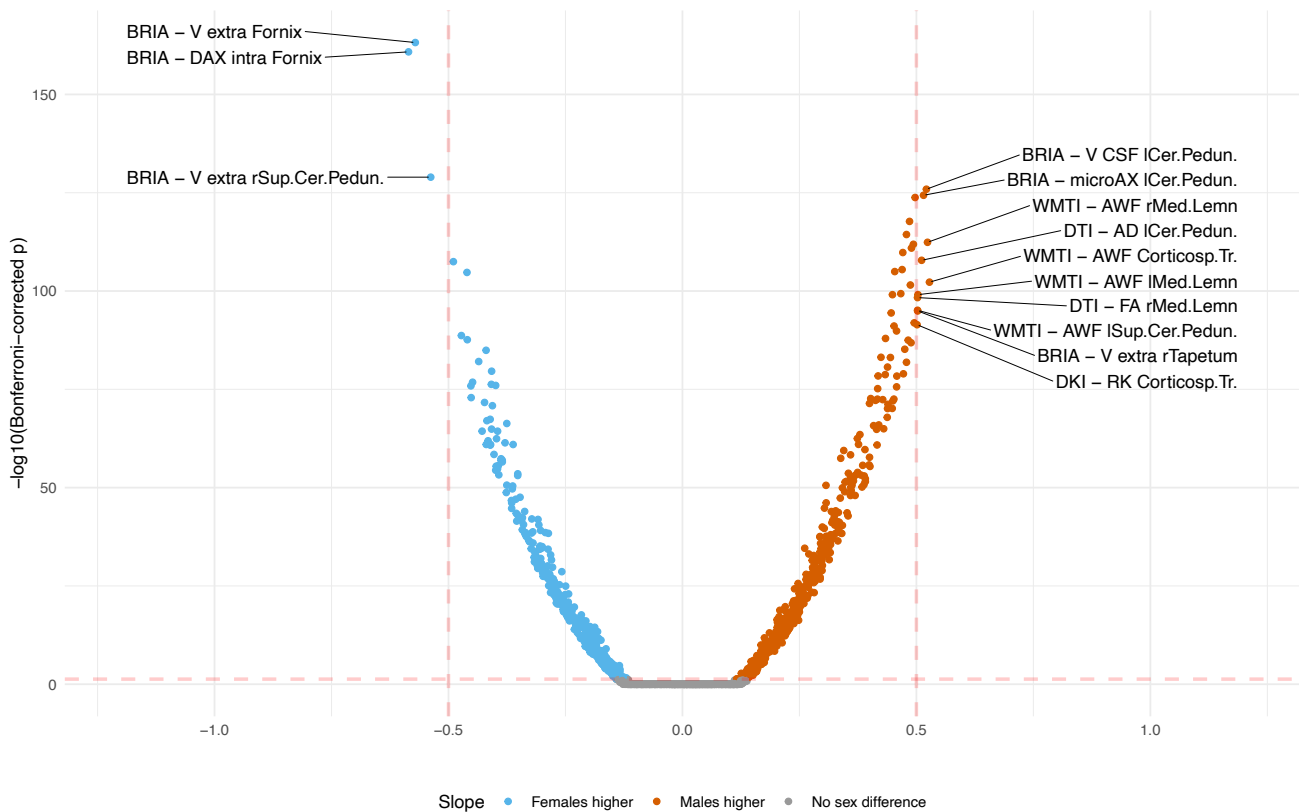
Supplementary Note 12: The effect of the sex-age interaction on regional WMM metrics.

Increases refer for higher WMM values for males at higher ages compared to females and decrease to lower WMM values, respectively.
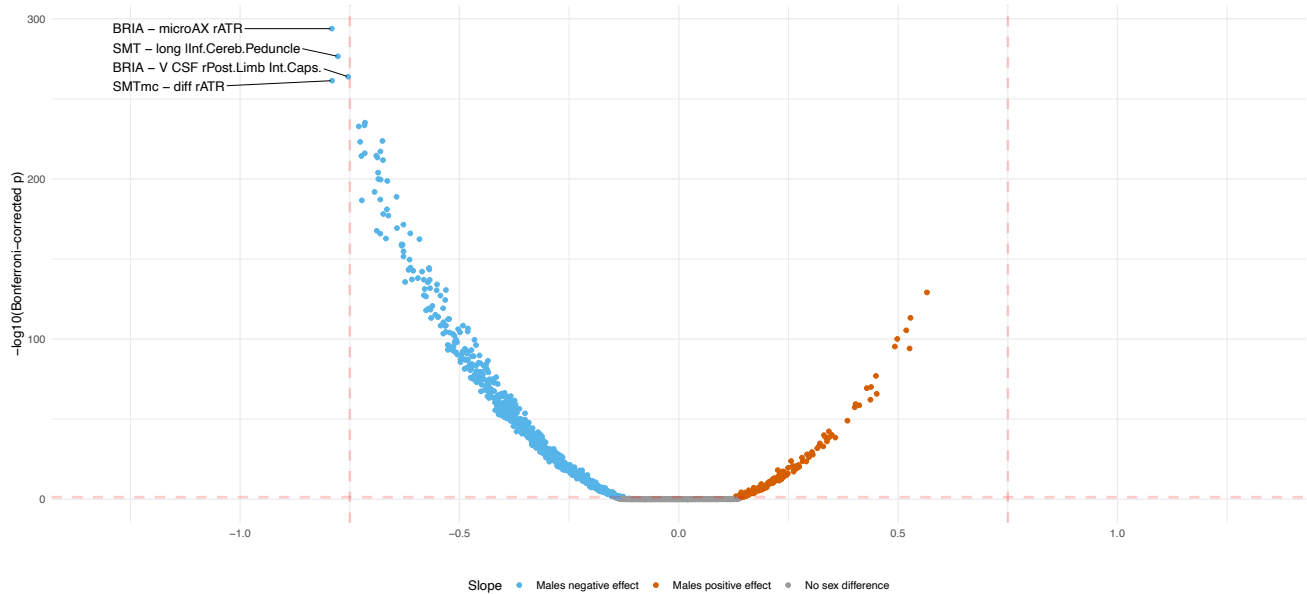
Supplementary Note 13: Distribution of the effect of age on regional WMM metrics.

Estimating average slopes from all β-values results in the following meta-statistics for negative β-values (left side of 0): Mean = -0.21±0.13, Median = -0.23± 0.12. For positive β-values (right side of 0) the following averages were estimated: Mean = 0.25±0.13, Median = 0.27± 0.14.

Estimating average slopes from significant effects only (right panel) results in similar meta-statistics for negative β-values (left side of 0): Mean = -0.21±0.13, Median = -0.24± 0.12, as well as for positive β-values (right side of 0): Mean = 0.25±0.13, Median = 0.27± 0.15.
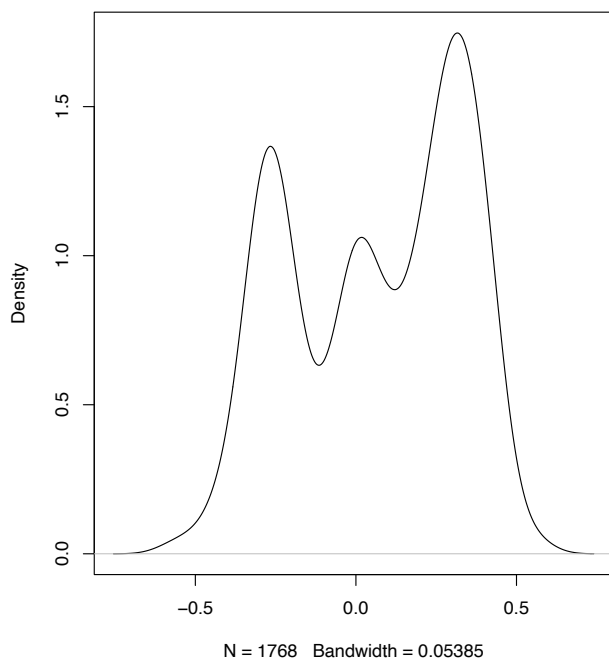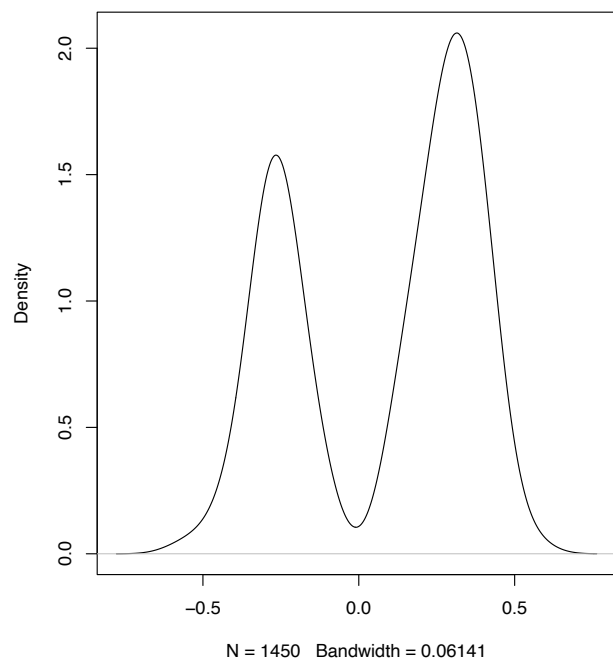
Supplementary Note 14: Adjusted cross-sectional associations between global WMM and age at time point 1.

Age-WMM relationships were adjusted for sex, the sex-age interaction, and site with the fit indicated for each time point.
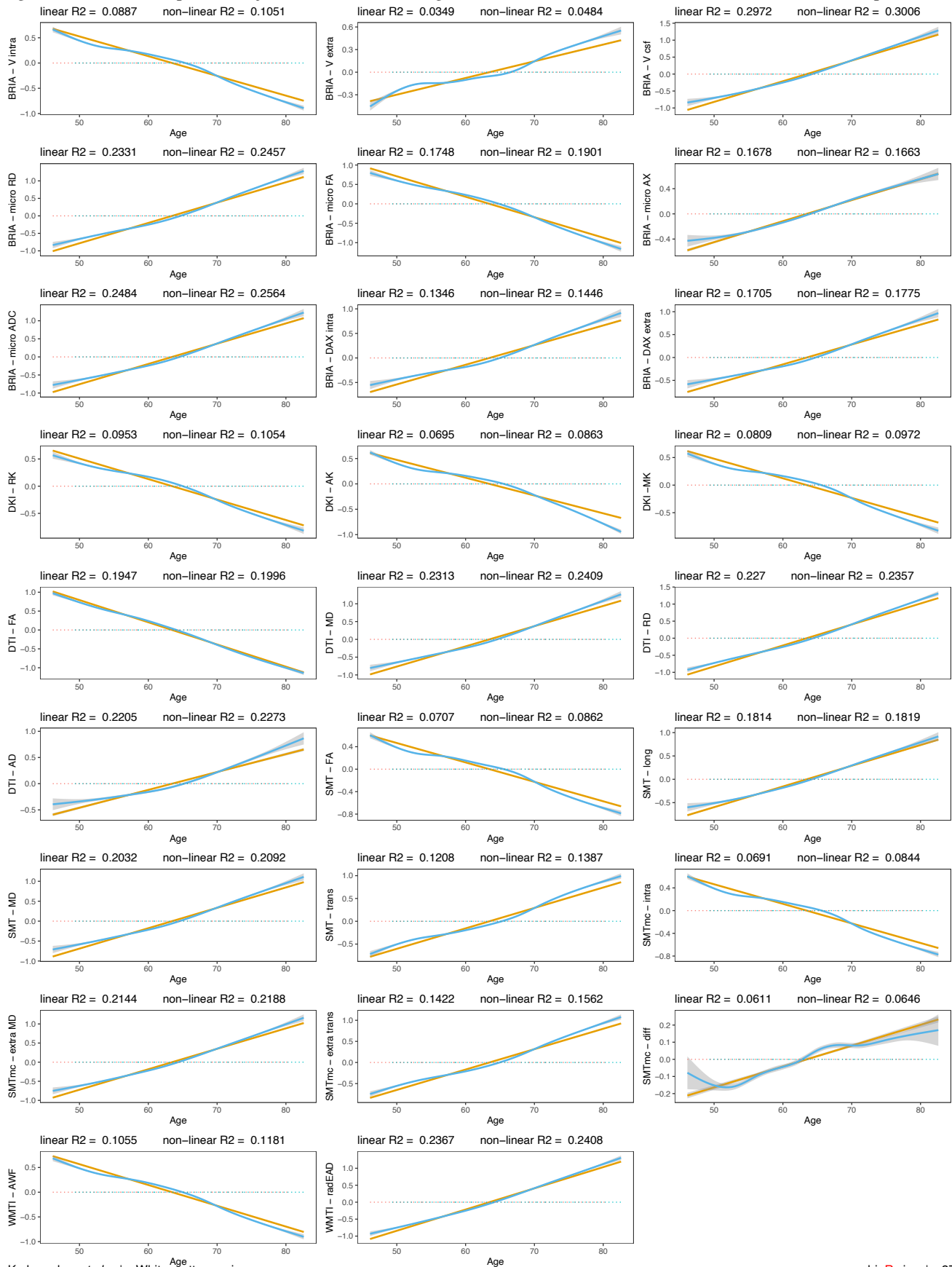
Supplementary Note 15: Adjusted cross-sectional associations between global WMM and age at time point 1.

Age-WMM relationships were adjusted for sex, the sex-age interaction, and site with the fit indicated for each time point.
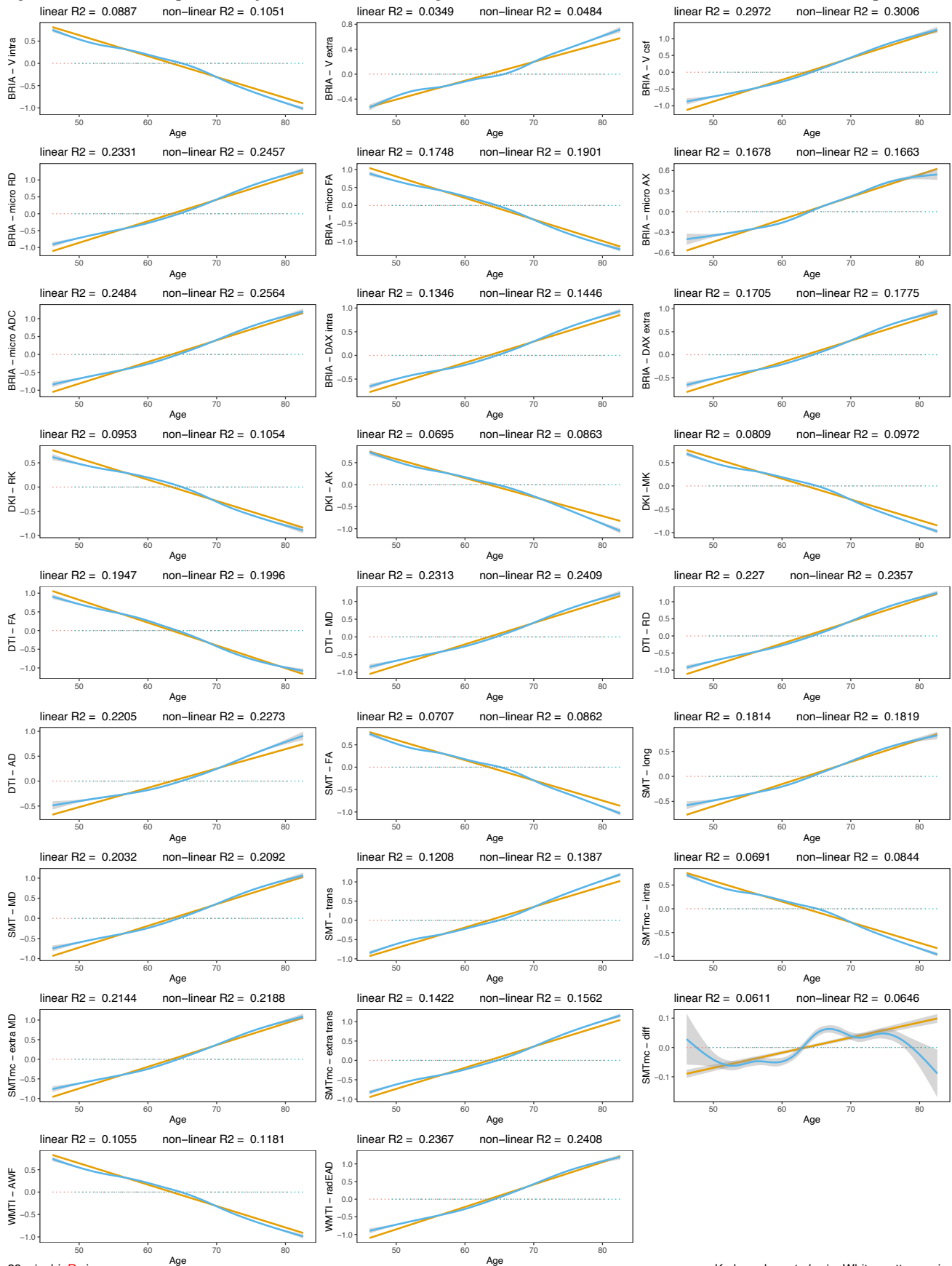
Supplementary Note 16: Cognitive test differences between time points.

Pair matches refers to incorrectly solved pair matches (executed in 3 rounds). Digits refers to the maximum number of digits remembered; intelligence to fluid intelligence; memory to prospective memory; health to a rating of the own overall health; matrix RTs refer to matrix puzzle completion times; matrix correct to the number of correctly solved matrix puzzles; matrix viewed to the number of viewed matrix puzzles; tower correct to the number of correctly solved tower puzzles; and Sym/Dig matches to the number of correct symbol-digit matches.

Metric	M_{TP1}	SD_{TP1}	M_{TP2}	SD_{TP2}	df	p	p_{adj}	d	**L95%CI**	**U95%CI**
Inc. pair matches R1	0.297	0.758	0.289	0.716	2282	0.359	1	-0.026	-0.08	0.029
Inc. pair matches R2	3.344	2.52	3.225	2.586	2282	0.011	0.292	-0.067	-0.119	-0.015
Inc. pair matches R3	3.742	4.077	3.664	3.815	1037	0.049	1	-0.087	-0.174	0
Max. digits rem.	6.844	1.39	6.656	1.79	2184	<0.001	<0.001	-0.103	-0.149	-0.056
Fluid intel.	6.854	2.018	6.796	2.026	2219	0.9	1	0.002	-0.033	0.037
Prosp. mem.	1.045	0.363	1.048	0.354	2282	0.411	1	0.023	-0.031	0.077
Health self-rating	1.896	0.622	1.915	0.654	2654	0.123	1	0.027	-0.007	0.06
Matrix RT1	29.95	9.828	30.627	10.277	1488	0.02	0.526	0.07	0.011	0.129
Matrix RT2	32.601	12.741	33.117	24.712	1488	0.691	1	0.014	-0.053	0.08
Matrix RT3	58.936	26.351	61.716	30.574	1488	0.008	0.201	0.082	0.022	0.142
Matrix RT4	67.765	35.483	68.263	35.013	1488	0.778	1	0.009	-0.053	0.07
Matrix RT5	43.224	20.951	42.608	19.722	1488	0.124	1	-0.044	-0.1	0.012
Matrix RT6	96.575	53.267	95.419	52.752	1488	0.053	1	-0.058	-0.118	0.001
Matrix RT7	136.815	66.333	137.116	67.855	1488	0.987	1	0	-0.058	0.059
Matrix RT8	93.788	43.243	97.961	48.273	1487	0	0.011	0.097	0.043	0.151
Matrix RT9	173.672	102.302	173.778	98.117	1475	0.802	1	0.008	-0.053	0.068
Matrix RT10	92.188	46.633	92.784	48.648	1456	0.42	1	0.025	-0.036	0.086
Matrix RT11	203.247	108.918	192.572	106.136	1329	0.007	0.179	-0.087	-0.15	-0.024
Matrix RT12	189.715	90.756	178.99	87.538	1077	0	0.007	-0.131	-0.201	-0.06
Matrix RT13	127.154	72.968	122.417	72.97	865	0.112	1	-0.066	-0.147	0.015
Matrix RT14	164.675	75.616	155.445	73.219	576	0.204	1	-0.064	-0.164	0.035
Matrix RT15	134.735	70.439	128.53	66.241	350	0.278	1	-0.065	-0.182	0.052
Correct matrix puzzles	8.355	2.015	8.338	2.046	1488	0.618	1	0.012	-0.036	0.061
Viewed matrix puzzles	13.911	1.359	13.967	1.361	1488	0.151	1	0.036	-0.013	0.085
Correct tower puzzles	10.38	3.118	10.522	3.156	1473	0	0.002	0.105	0.053	0.156
Correct sym.-dig. matches	20.099	4.964	19.938	5.256	1472	0.764	1	0.007	-0.038	0.052

Supplementary Note 17: Distribution of the relationship between age and the absolute annual rate of white matter microstructure change

The Figure shows the distribution of adjusted |ARoC|-age associations across brain regions ($\bar{\beta}_{padj<0.05} = 0.013 \pm 0.005$). Mean associations were smaller when including non-significant associations ($\bar{\beta}_{all} = 0.007 \pm 0.005$).
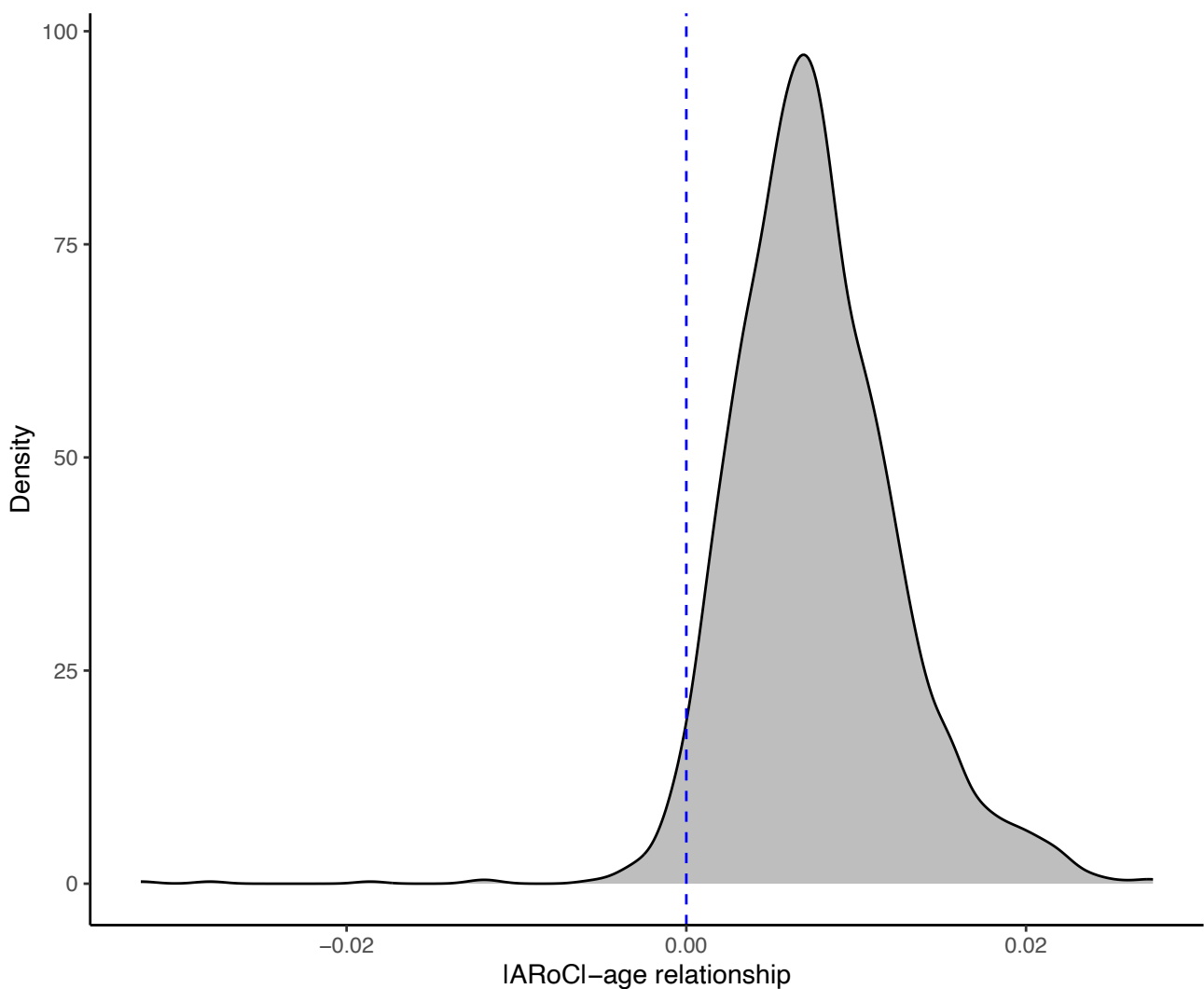
Supplementary Note 18: Age-stratified annual rate of WMM change group-comparisons

Table 2: Age-stratified annual rate of WMM change group-comparisons

Diffusion Metric	Group 2	Group 2	Cohen's d	n_1	n_2	CI_{low}	CI_{high}	Magnitude	p_{adj}
BRIA - v intra	46-55	55-65	0.65	525	1157	0.54	0.78	moderate	<.001
BRIA - v intra	46-55	65-75	1.31	525	905	1.18	1.43	large	<.001
BRIA - v intra	46-55	75-81	1.78	525	91	1.54	2.06	large	<.001
BRIA - v intra	55-65	65-75	0.80	1157	905	0.71	0.9	large	<.001
BRIA - v intra	55-65	75-81	1.41	1157	91	1.14	1.73	large	<.001
BRIA - v intra	65-75	75-81	0.78	905	91	0.52	1.05	moderate	<.001
BRIA – v extra	46-55	55-65	-0.73	525	1157	-0.86	-0.6	moderate	<.001
BRIA – v extra	46-55	65-75	-1.44	525	905	-1.57	-1.31	large	<.001
BRIA – v extra	46-55	75-81	-1.88	525	91	-2.2	-1.62	large	<.001
BRIA – v extra	55-65	65-75	-0.94	1157	905	-1.04	-0.85	large	<.001
BRIA – v extra	55-65	75-81	-1.53	1157	91	-1.88	-1.26	large	<.001
BRIA – v extra	65-75	75-81	-0.83	905	91	-1.12	-0.58	large	<.001
BRIA – v CSF	46-55	55-65	-0.30	525	1157	-0.41	-0.19	small	<.001
BRIA – v CSF	46-55	65-75	-0.51	525	905	-0.62	-0.39	moderate	<.001
BRIA – v CSF	46-55	75-81	-0.71	525	91	-0.92	-0.5	moderate	<.001
BRIA – v CSF	55-65	65-75	-0.20	1157	905	-0.29	-0.12	small	<.001
BRIA – v CSF	55-65	75-81	-0.41	1157	91	-0.61	-0.2	small	0.001
BRIA – v CSF	65-75	75-81	-0.21	905	91	-0.43	0.01	small	0.064
BRIA – micro RD	46-55	55-65	-0.45	525	1157	-0.55	-0.35	small	<.001
BRIA – micro RD	46-55	65-75	-0.85	525	905	-0.98	-0.74	large	<.001
BRIA – micro RD	46-55	75-81	-1.25	525	91	-1.5	-1.02	large	<.001
BRIA – micro RD	55-65	65-75	-0.42	1157	905	-0.51	-0.33	small	<.001
BRIA – micro RD	55-65	75-81	-0.87	1157	91	-1.13	-0.66	large	<.001
BRIA – micro RD	65-75	75-81	-0.48	905	91	-0.72	-0.26	small	<.001
BRIA – micro FA	46-55	55-65	0.56	525	1157	0.45	0.67	moderate	<.001
BRIA – micro FA	46-55	65-75	1.11	525	905	0.99	1.24	large	<.001
BRIA – micro FA	46-55	75-81	1.58	525	91	1.35	1.84	large	<.001
BRIA – micro FA	55-65	65-75	0.63	1157	905	0.55	0.72	moderate	<.001
BRIA – micro FA	55-65	75-81	1.21	1157	91	0.96	1.49	large	<.001
BRIA – micro FA	65-75	75-81	0.68	905	91	0.44	0.92	moderate	<.001
BRIA – micro AX	46-55	55-65	-0.06	525	1157	-0.16	0.04	negligible	0.478
BRIA – micro AX	46-55	65-75	0.04	525	905	-0.07	0.14	negligible	0.519
BRIA – micro AX	46-55	75-81	0.20	525	91	-0.05	0.43	small	0.332
BRIA – micro AX	55-65	65-75	0.09	1157	905	0.01	0.19	negligible	0.170
BRIA – micro AX	55-65	75-81	0.26	1157	91	0.04	0.48	small	0.121
BRIA – micro AX	65-75	75-81	0.16	905	91	-0.04	0.39	negligible	0.426
BRIA – micro ADC	46-55	55-65	-0.32	525	1157	-0.44	-0.22	small	<.001

Continued on next page

Table 2: Age-stratified annual rate of WMM change group-comparisons (Continued)

BRIA – micro ADC	46-55	65-75	-0.55	525	905	-0.66	-0.43	moderate	<.001
BRIA – micro ADC	46-55	75-81	-0.80	525	91	-1.03	-0.58	moderate	<.001
BRIA – micro ADC	55-65	65-75	-0.23	1157	905	-0.32	-0.14	small	<.001
BRIA – micro ADC	55-65	75-81	-0.48	1157	91	-0.7	-0.28	small	<.001
BRIA – micro ADC	65-75	75-81	-0.26	905	91	-0.48	-0.04	small	0.023
BRIA – DAX intra	46-55	55-65	-0.30	525	1157	-0.41	-0.2	small	<.001
BRIA – DAX intra	46-55	65-75	-0.51	525	905	-0.62	-0.4	moderate	<.001
BRIA – DAX intra	46-55	75-81	-0.74	525	91	-0.96	-0.52	moderate	<.001
BRIA – DAX intra	55-65	65-75	-0.20	1157	905	-0.29	-0.11	small	<.001
BRIA – DAX intra	55-65	75-81	-0.44	1157	91	-0.67	-0.23	small	<.001
BRIA – DAX intra	65-75	75-81	-0.24	905	91	-0.47	-0.02	small	0.033
BRIA – DAX extra	46-55	55-65	-0.26	525	1157	-0.36	-0.15	small	<.001
BRIA – DAX extra	46-55	65-75	-0.40	525	905	-0.51	-0.29	small	<.001
BRIA – DAX extra	46-55	75-81	-0.56	525	91	-0.78	-0.33	moderate	<.001
BRIA – DAX extra	55-65	65-75	-0.15	1157	905	-0.24	-0.05	negligible	0.003
BRIA – DAX extra	55-65	75-81	-0.31	1157	91	-0.53	-0.09	small	0.012
BRIA – DAX extra	65-75	75-81	-0.16	905	91	-0.39	0.06	negligible	0.145
DKI – RK	46-55	55-65	0.59	525	1157	0.48	0.7	moderate	<.001
DKI – RK	46-55	65-75	1.34	525	905	1.24	1.45	large	<.001
DKI – RK	46-55	75-81	1.90	525	91	1.69	2.17	large	<.001
DKI – RK	55-65	65-75	1.05	1157	905	0.98	1.14	large	<.001
DKI – RK	55-65	75-81	1.73	1157	91	1.44	2.05	large	<.001
DKI – RK	65-75	75-81	0.98	905	91	0.71	1.29	large	<.001
DKI – AK	46-55	55-65	0.41	525	1157	0.31	0.51	small	<.001
DKI – AK	46-55	65-75	0.75	525	905	0.64	0.87	moderate	<.001
DKI – AK	46-55	75-81	1.12	525	91	0.88	1.36	large	<.001
DKI – AK	55-65	65-75	0.35	1157	905	0.27	0.44	small	<.001
DKI – AK	55-65	75-81	0.76	1157	91	0.54	0.99	moderate	<.001
DKI – AK	65-75	75-81	0.44	905	91	0.23	0.67	small	<.001
DKI – MK	46-55	55-65	0.61	525	1157	0.5	0.72	moderate	<.001
DKI – MK	46-55	65-75	1.22	525	905	1.09	1.34	large	<.001
DKI – MK	46-55	75-81	1.69	525	91	1.45	1.97	large	<.001
DKI – MK	55-65	65-75	0.72	1157	905	0.64	0.81	moderate	<.001
DKI – MK	55-65	75-81	1.32	1157	91	1.07	1.6	large	<.001
DKI – MK	65-75	75-81	0.74	905	91	0.51	0.99	moderate	<.001
DTI – FA	46-55	55-65	0.23	525	1157	0.11	0.34	small	<.001
DTI – FA	46-55	65-75	0.66	525	905	0.56	0.77	moderate	<.001
DTI – FA	46-55	75-81	1.00	525	91	0.74	1.29	large	<.001
DTI – FA	55-65	65-75	0.46	1157	905	0.36	0.55	small	<.001
DTI – FA	55-65	75-81	0.82	1157	91	0.58	1.1	large	<.001
DTI – FA	65-75	75-81	0.40	905	91	0.17	0.66	small	0.001

Table 2: Age-stratified annual rate of WMM change group-comparisons (Continued)

DTI – MD	46-55	55-65	-0.43	525	1157	-0.54	-0.33	small	<.001
DTI – MD	46-55	65-75	-0.79	525	905	-0.91	-0.67	moderate	<.001
DTI – MD	46-55	75-81	-1.15	525	91	-1.41	-0.94	large	<.001
DTI – MD	55-65	65-75	-0.38	1157	905	-0.46	-0.29	small	<.001
DTI – MD	55-65	75-81	-0.77	1157	91	-1	-0.56	moderate	<.001
DTI – MD	65-75	75-81	-0.40	905	91	-0.62	-0.17	small	0.001
DTI – RD	46-55	55-65	-0.74	525	1157	-0.85	-0.63	moderate	<.001
DTI – RD	46-55	65-75	-1.47	525	905	-1.59	-1.36	large	<.001
DTI – RD	46-55	75-81	-1.92	525	91	-2.27	-1.63	large	<.001
DTI – RD	55-65	65-75	-0.93	1157	905	-1.04	-0.85	large	<.001
DTI – RD	55-65	75-81	-1.53	1157	91	-1.88	-1.24	large	<.001
DTI – RD	65-75	75-81	-0.76	905	91	-1.04	-0.52	moderate	<.001
DTI – AD	46-55	55-65	-0.22	525	1157	-0.33	-0.11	small	<.001
DTI – AD	46-55	65-75	-0.32	525	905	-0.43	-0.2	small	<.001
DTI – AD	46-55	75-81	-0.43	525	91	-0.67	-0.2	small	0.001
DTI – AD	55-65	65-75	-0.09	1157	905	-0.19	-0.0066	negligible	0.109
DTI – AD	55-65	75-81	-0.21	1157	91	-0.42	0.007	small	0.114
DTI – AD	65-75	75-81	-0.12	905	91	-0.35	0.08	negligible	0.289
SMT – FA	46-55	55-65	0.53	525	1157	0.43	0.64	moderate	<.001
SMT – FA	46-55	65-75	1.06	525	905	0.94	1.19	large	<.001
SMT – FA	46-55	75-81	1.51	525	91	1.28	1.77	large	<.001
SMT – FA	55-65	65-75	0.59	1157	905	0.5	0.68	moderate	<.001
SMT – FA	55-65	75-81	1.15	1157	91	0.93	1.41	large	<.001
SMT – FA	65-75	75-81	0.66	905	91	0.43	0.91	moderate	<.001
SMT – long	46-55	55-65	-0.15	525	1157	-0.26	-0.04	negligible	0.021
SMT – long	46-55	65-75	-0.16	525	905	-0.27	-0.05	negligible	0.018
SMT – long	46-55	75-81	-0.11	525	91	-0.35	0.1	negligible	1.000
SMT – long	55-65	65-75	-0.01	1157	905	-0.11	0.07	negligible	1.000
SMT – long	55-65	75-81	0.04	1157	91	-0.19	0.25	negligible	1.000
SMT – long	65-75	75-81	0.05	905	91	-0.18	0.3	negligible	1.000
SMT – MD	46-55	55-65	-0.28	525	1157	-0.39	-0.18	small	<.001
SMT – MD	46-55	65-75	-0.46	525	905	-0.57	-0.35	small	<.001
SMT – MD	46-55	75-81	-0.64	525	91	-0.87	-0.42	moderate	<.001
SMT – MD	55-65	65-75	-0.18	1157	905	-0.26	-0.09	negligible	<.001
SMT – MD	55-65	75-81	-0.36	1157	91	-0.58	-0.15	small	0.003
SMT – MD	65-75	75-81	-0.19	905	91	-0.42	0.02	negligible	0.096
SMT – trans	46-55	55-65	-0.49	525	1157	-0.6	-0.38	small	<.001
SMT – trans	46-55	65-75	-0.93	525	905	-1.06	-0.82	large	<.001
SMT – trans	46-55	75-81	-1.36	525	91	-1.61	-1.14	large	<.001
SMT – trans	55-65	65-75	-0.47	1157	905	-0.56	-0.39	small	<.001
SMT – trans	55-65	75-81	-0.97	1157	91	-1.23	-0.74	large	<.001

Table 2: Age-stratified annual rate of WMM change group-comparisons (Continued)

SMT – trans	65-75	75-81	-0.55	905	91	-0.8	-0.32	moderate	<.001
SMTmc – intra	46-55	55-65	0.70	525	1157	0.6	0.81	moderate	<.001
SMTmc – intra	46-55	65-75	1.36	525	905	1.23	1.49	large	<.001
SMTmc – intra	46-55	75-81	1.82	525	91	1.56	2.14	large	<.001
SMTmc – intra	55-65	65-75	0.82	1157	905	0.73	0.9	large	<.001
SMTmc – intra	55-65	75-81	1.42	1157	91	1.15	1.75	large	<.001
SMTmc – intra	65-75	75-81	0.76	905	91	0.5	1.03	moderate	<.001
SMTmc – extra MD	46-55	55-65	-0.32	525	1157	-0.43	-0.21	small	<.001
SMTmc – extra MD	46-55	65-75	-0.55	525	905	-0.67	-0.44	moderate	<.001
SMTmc – extra MD	46-55	75-81	-0.77	525	91	-1	-0.55	moderate	<.001
SMTmc – extra MD	55-65	65-75	-0.23	1157	905	-0.32	-0.14	small	<.001
SMTmc – extra MD	55-65	75-81	-0.46	1157	91	-0.67	-0.24	small	<.001
SMTmc – extra MD	65-75	75-81	-0.23	905	91	-0.45	-0.01	small	0.041
SMTmc – extra trans	46-55	55-65	-0.61	525	1157	-0.72	-0.49	moderate	<.001
SMTmc – extra trans	46-55	65-75	-1.16	525	905	-1.28	-1.05	large	<.001
SMTmc – extra trans	46-55	75-81	-1.63	525	91	-1.92	-1.36	large	<.001
SMTmc – extra trans	55-65	65-75	-0.64	1157	905	-0.73	-0.55	moderate	<.001
SMTmc – extra trans	55-65	75-81	-1.20	1157	91	-1.49	-0.95	large	<.001
SMTmc – extra trans	65-75	75-81	-0.64	905	91	-0.9	-0.4	moderate	<.001
SMTmc – diff	46-55	55-65	0.61	525	1157	0.51	0.72	moderate	<.001
SMTmc – diff	46-55	65-75	1.23	525	905	1.12	1.34	large	<.001
SMTmc – diff	46-55	75-81	1.73	525	91	1.44	2.06	large	<.001
SMTmc – diff	55-65	65-75	0.78	1157	905	0.68	0.87	moderate	<.001
SMTmc – diff	55-65	75-81	1.40	1157	91	1.1	1.79	large	<.001
SMTmc – diff	65-75	75-81	0.73	905	91	0.47	1.01	moderate	<.001
WMTI – AWF	46-55	55-65	0.55	525	1157	0.43	0.66	moderate	<.001
WMTI – AWF	46-55	65-75	1.11	525	905	0.99	1.24	large	<.001
WMTI – AWF	46-55	75-81	1.58	525	91	1.38	1.84	large	<.001
WMTI – AWF	55-65	65-75	0.65	1157	905	0.56	0.74	moderate	<.001
WMTI – AWF	55-65	75-81	1.24	1157	91	1.02	1.49	large	<.001
WMTI – AWF	65-75	75-81	0.72	905	91	0.51	0.97	moderate	<.001
WMTI – radEAD	46-55	55-65	-0.58	525	1157	-0.7	-0.47	moderate	<.001
WMTI – radEAD	46-55	65-75	-1.16	525	905	-1.28	-1.04	large	<.001
WMTI – radEAD	46-55	75-81	-1.52	525	91	-1.79	-1.27	large	<.001
WMTI – radEAD	55-65	65-75	-0.64	1157	905	-0.73	-0.56	moderate	<.001
WMTI – radEAD	55-65	75-81	-1.09	1157	91	-1.38	-0.87	large	<.001
WMTI – radEAD	65-75	75-81	-0.52	905	91	-0.77	-0.28	moderate	<.001

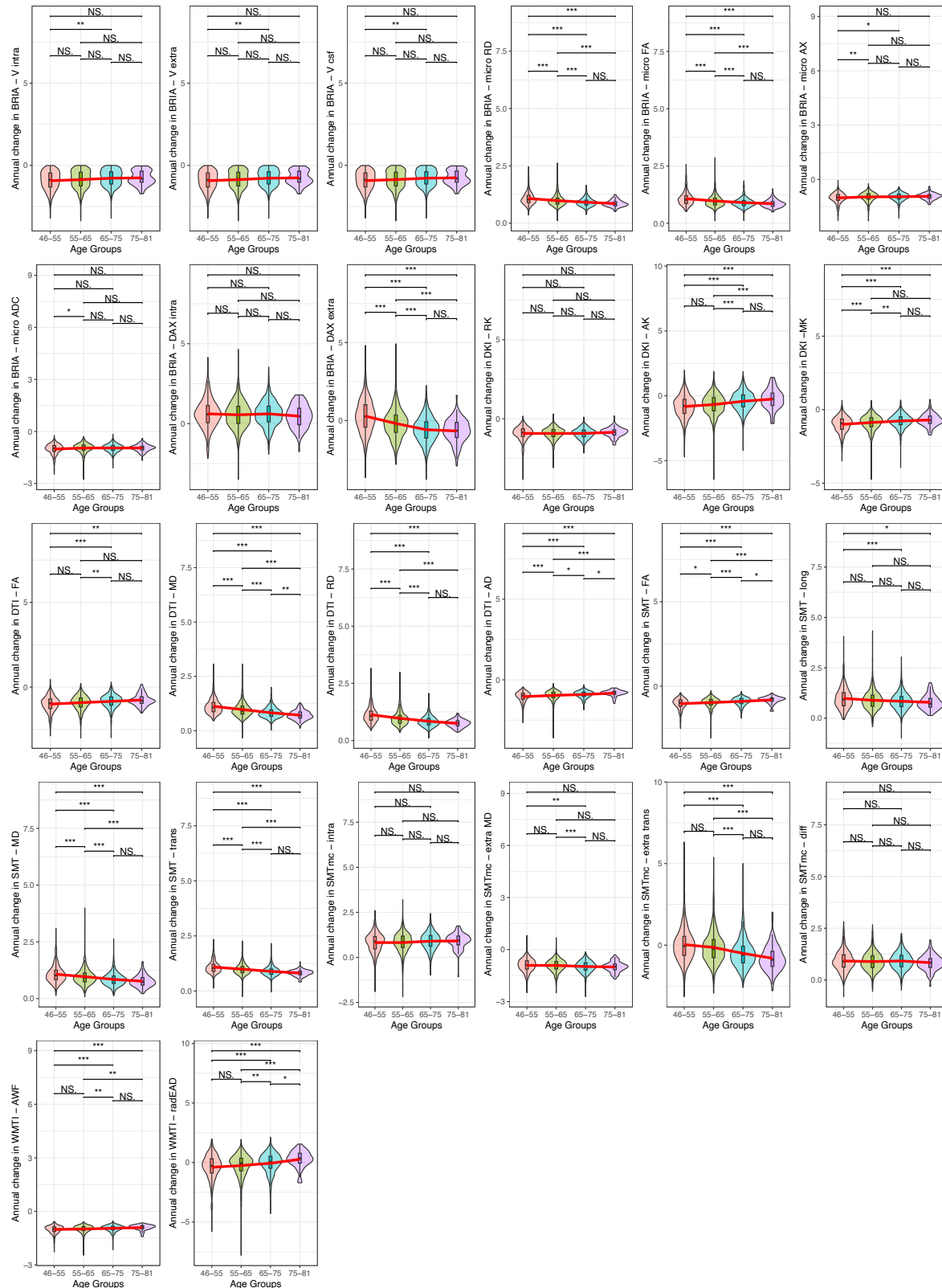## Supplementary Note 19: Adjusted annual rate of WMM change throughout ages

WMM are standardised for comparability without mean centering for a better understanding of onset values. Non-linear curves were fitted, using splines of generalized additive models, and Pearson's correlation coefficients were estimated (top left in each plot). We present uncorrected p-values, which were significant at the Bonferroni-corrected $\alpha < 0.05/26 = 0.0019$. Additionally, we estimated associations between |ARoC| and age (to meaningfully estimate meta-statistics), showing a higher |ARoC| at higher ages ($\bar{\beta}_{sig} = 0.012 \pm 0.003$), with a slightly lower average slopes when considering non-significant age-|ARoC| associations ($\bar{\beta}_{all} = 0.008 \pm 0.004$)

Paper E   **231**
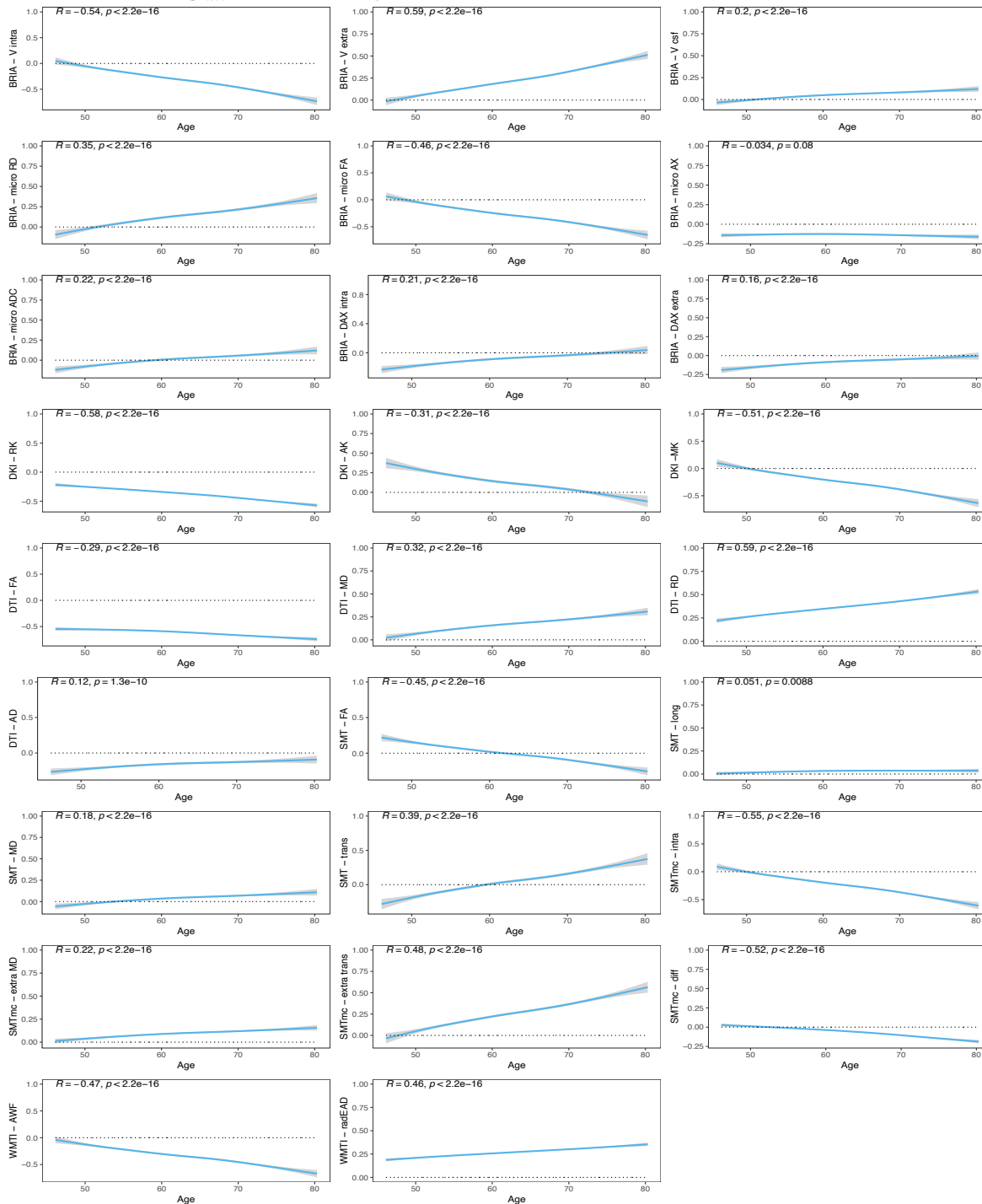
Supplementary Note 20: Unadjusted annual rate of WMM change throughout ages

WMM are standardised for comparability without mean centering for a better understanding of onset values. Red lines were implemented as visual aid.

232 Paper E

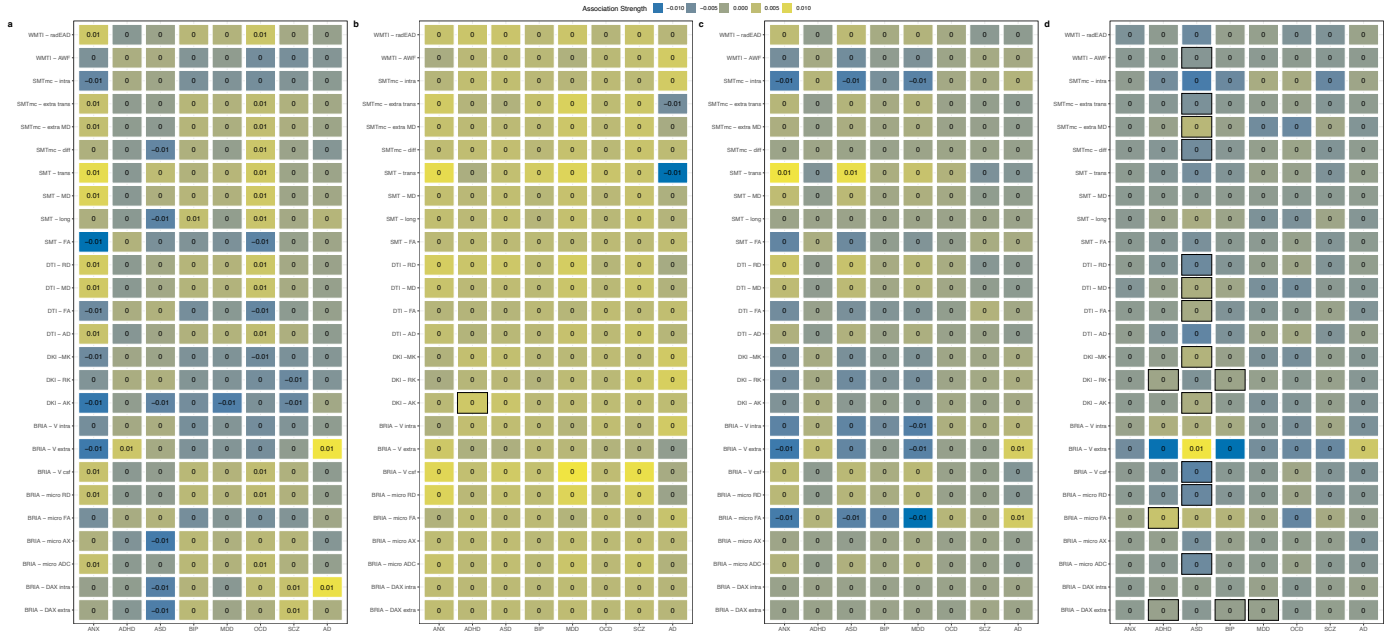## Supplementary Note 21: Corrected sex- and age-stratified annual rate of WMM change

WMM are standardised for comparability (including mean centering). Red colours indicate females and blue colours males.

Supplementary Note 22: Regional associations between WMM change in the Cerebral Peduncle and PGRS

Panel (a) presents the associations between PGRS and WMM change. Panel (b) presents the regional associations between PGRS and cross-sectional regional WMM at time point one. Panel (c) presents the cross-sectional regional WMM associations with PGRS at time point two. Panel (d) presents cross-sectional regional associations between WMM and PGRS for the validation sample. Boxes the statistical significance at an uncorrected $\alpha < 0.05$. Colours indicate the association strength (standardised β-coefficients). *Note:* All associations were adjusted for age, sex, the $age \times sex$ interaction, and site. None of the presented associations survived the adjustment of the α-level for multiple comparisons.

C Polygenic Risk Scores

Supplementary Note 23: Regional associations between WMM change in the Fornix and PGRS

Panel (a) presents the associations between PGRS and WMM change. Panel (b) presents the regional associations between PGRS and cross-sectional regional WMM at time point one. Panel (c) presents the cross-sectional regional WMM associations with PGRS at time point two. Panel (d) presents cross-sectional regional associations between WMM and PGRS for the validation sample. Boxes the statistical significance at an uncorrected $\alpha < 0.05$. Colours indicate the association strength (standardised β-coefficients). *Note:* All associations were adjusted for age, sex, the $age \times sex$ interaction, and site. None of the presented associations survived the adjustment of the α-level for multiple comparisons.

Paper E   **235**

Supplementary Note 24: Distribution of the relationship between PGRS and both the annual rate of WMM change as well as cross-sectional WMM
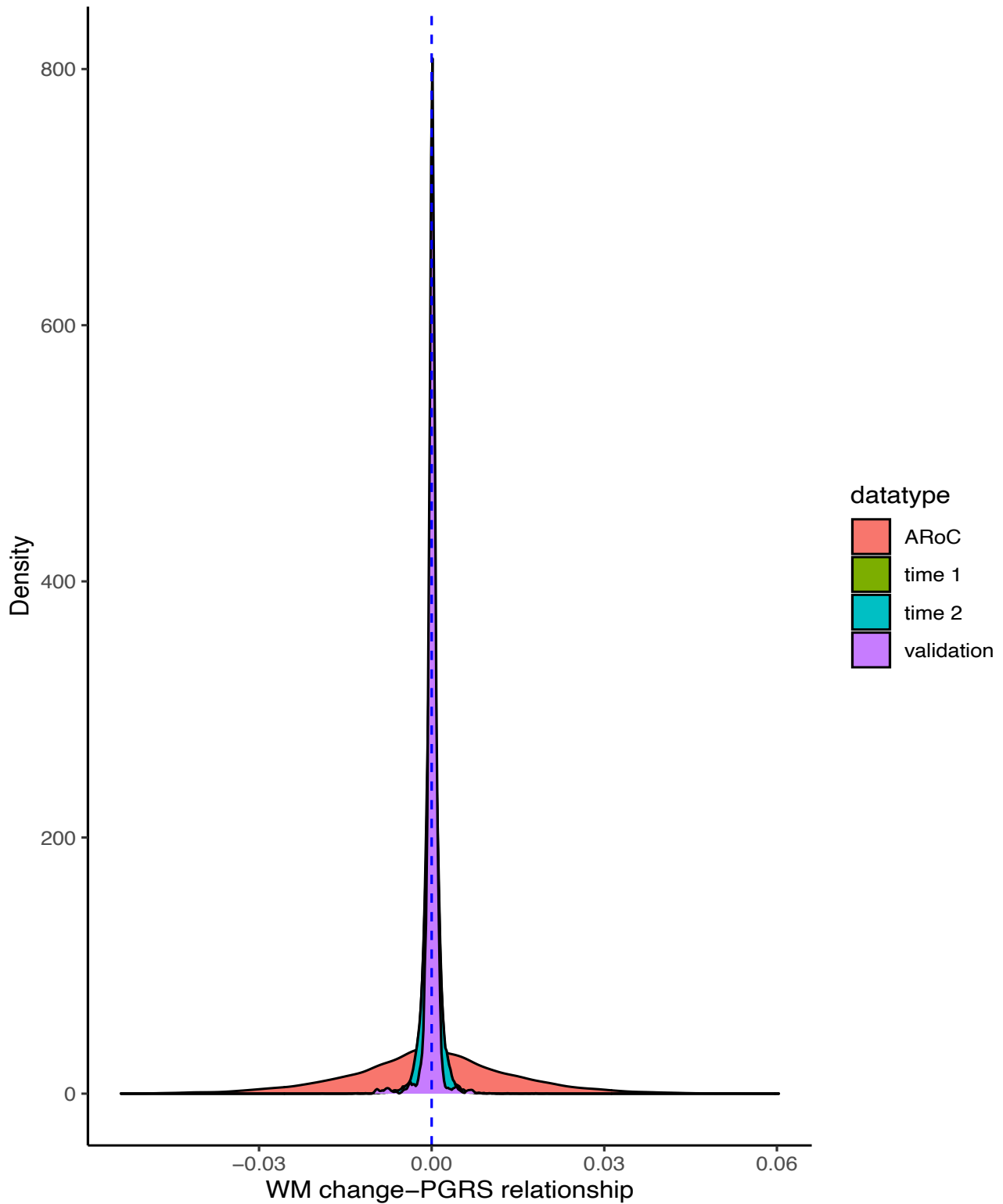
Part III

APPENDIX

APPENDIX

1 MRI pulse sequences

Pulse sequences can be categorised into two principle types: spin echo and gradient echo, which both utilize two RF pulses to create the echo to measure signal intensity. Spin echo sequences can produce all three mentioned contrasts (T_1, T_2, proton density) at a high quality but longer acquisition times. The resulting contrasts depend on the choice of TE and TR (short TR & TE = T_1-weighted, long TR & short TE = proton density-weighted, long TR & TE = T_2-weighted, long TE & short TR do not produce useful contrasts). Also gradient echo sequences can produce all three contrasts at shorter times but is more susceptible to inhomogeneities in the magnetic field (in combination with T_1&T_2 = T_1*&T_2*), and affected by TE and TR. For gradient echo sequences TR is always short compared to spin echo sequences, and the flip angle α plays a role. Lower flip angles ($\alpha < 40°$ short) produce proton density-weighted images when combined with a short TE and T_2-weighted contrasts when combined with long TE. Higher flip angles ($\alpha > 50°$) produce T_1-weighted contrasts when combined with short TE, but no useful contrast when combined with long TE [203].

2 MRI advantages and safety

The possibility of fine-tuning sequences to provide detailed soft tissue contrast, makes MRI extremely useful for detailed imaging of the brain in the clinic and research. Moreover, no injections of contrast agents are strictly required in contrast to other imaging techniques such as positron emission tomography, and no exposure to ionizing radiation (as in CT) is necessary. This makes MRI a safe imaging option independent of the target group (e.g., children, elderly, fragile health states). However, it is possible to utilize contrast agents for MRI, mainly to emphasise or increase certain details in the diagnostic images [133]. These contrast agents are gadolinium-based considering the currently Food and Drug Administration (FDA) approved MRI contrast agents [133]. The only known safety restrictions to routine clinical acquisition procedures are metal objects inside the body, such as implants, pacemakers, tattoos (less than six weeks old), and claustrophobic anxiety [108]. Yet, there are possibilities to also facilitate scanning when there is metal in the body [147] or even electrical devices such as pacemakers or defilibrators [96, 112, 325].
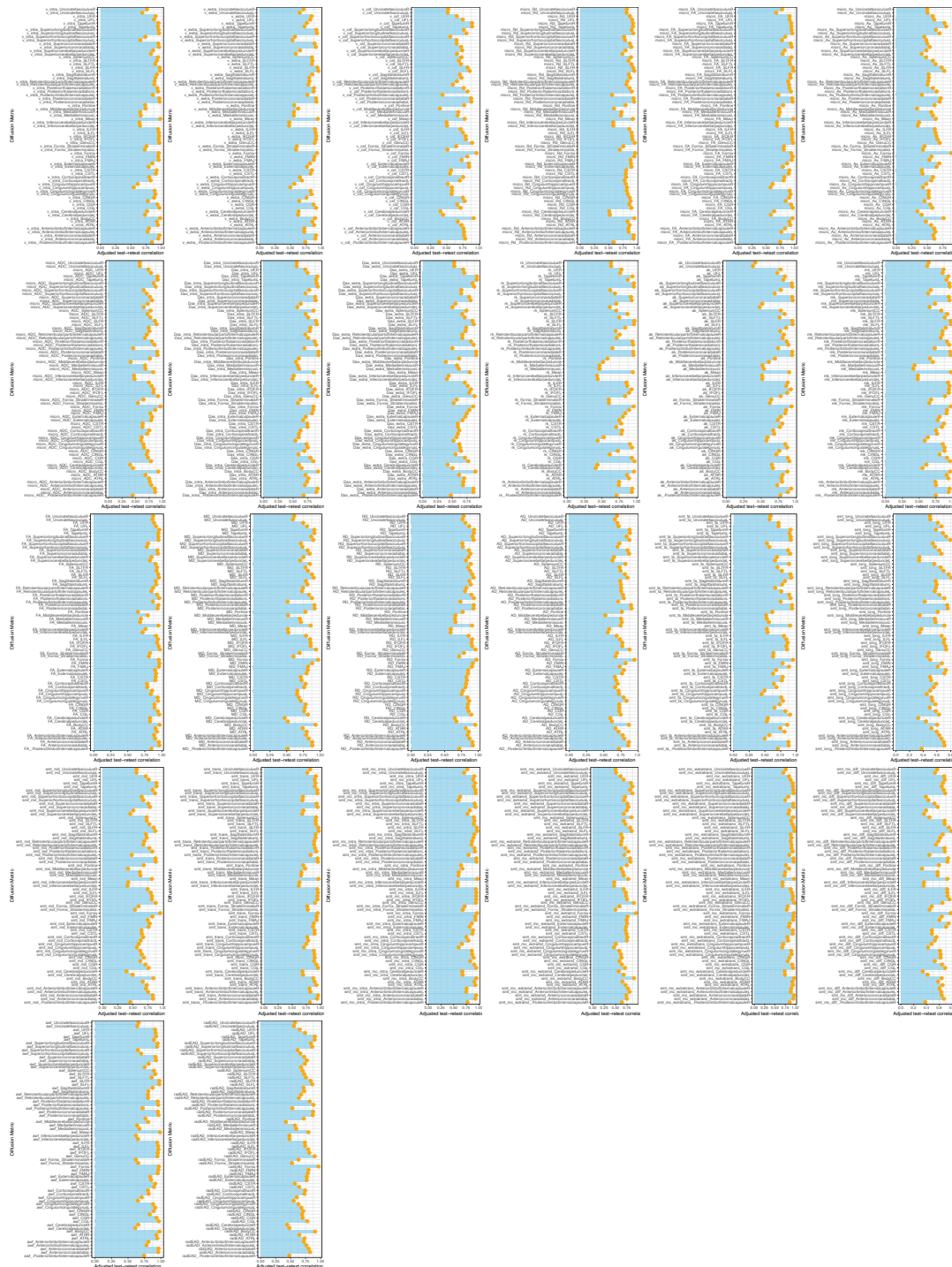
3 Utilized diffusion approaches and their scalar metrics

Diffusion Approach	Metrics
Bayesian Rotationally Invariant Approach (BRIA) [253]	intra-axonal axial diffusivity (DAX intra)
	extra-axonal radial diffusivity (DRAD extra)
	microscopic fractional anisotropy (micro FA)
	extra-axonal axial diffusivity (DAX extra)
	intra-axonal water fraction (V intra)
	extra-axonal water fraction (V extra)
	cerebrospinal fluid fraction (vCSF)
	microscopical axial diffusivity (micro AX)
	microscopic radial diffusivity (micro RD)
	microscopical apparent diffusion coefficient (micro ADC)
Diffusion Kurtosis Imaging (DKI) [142]	mean kurtosis (MK)
	radial kurtosis (RK)
	axial kurtosis (AK)
Diffusion Tensor Imaging (DTI) [27]	fractional anisotropy (FA)
	axial diffusivity (AD)
	mean diffusivity (MD)
	radial diffusivity (RD)
Spherical Mean Technique (SMT) [149]	fractional anisotropy (SMT FA)
	mean diffusivity (SMT md)
	transverse diffusion coefficient (SMT trans)
	longitudinal diffusion coefficient (SMT long)
Multi-compartment Spherical Mean Technique (SMTmc) [148]	extra-neurite microscopic mean diffusivity (SMTmc extra md)
	extra-neurite transverse microscopic diffusivity (SMTmc extra trans)
	mc SMTdiffusion coefficient (SMT mcd)
	intra-neurite volume fraction (SMTmc intra)
White Matter Tract Integrity (WMTI) [97]	axonal water fraction (AWF)
	radial extra-axonal diffusivity (radEAD)
	axial extra-axonal diffusivity (axEAD)

4 Global TBSS test-retest reliability
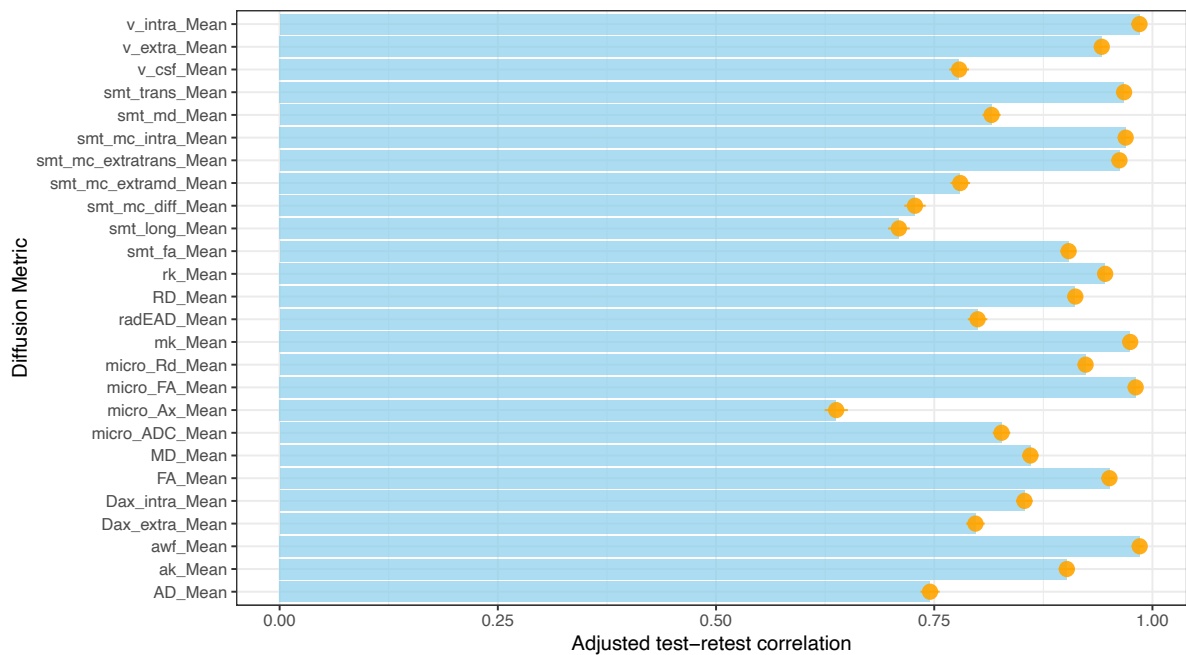
Fig. 1: Global white matter metrics' test-retest correlation with standard error. White matter metrics were derived from different diffusion approaches in N = 2,915 UK Biobank participants which were scanned twice (including patients) with the average inter-scan interval of Δ=2.43 years. Test-retest correlations were corrected for scanner site, age, and sex.

5 Regional TBSS test-retest reliability

Fig. 2: Regional white matter metrics' test-retest correlation with standard error. White matter metrics were derived from different diffusion approaches in N = 2,915 UK Biobank participants which were scanned twice (including patients) with the average inter-scan interval of Δ=2.43 years. Test-retest correlations were corrected for scanner site, age, and sex.