Longitudinal Characterisation of Healthy Ageing in Rats using Multimodal Magnetic Resonance Imaging

Eilidh MacNicol, MSc

Department of Neuroimaging,

Institute of Psychiatry, Psychology and Neuroscience,

King's College London

This thesis is submitted for examination towards the degree of

Doctor of Philosophy (PhD) in Neuroimaging

December 2020

Word count: 63,900

ABSTRACT

Ageing is a universally heterogeneous process which is associated with a decline in certain cognitive functions and changes to brain structure, but a given individual's ageing phenotype is complex and difficult to predict. Ageing trajectories are non-linear and influenced throughout the lifespan by a multitude of factors, including diet and lifestyle. Non-invasive neuroimaging, such as magnetic resonance imaging, facilitates the *in vivo* characterisation of brain structure and function and is widely used in the endeavour of identifying individual trajectories of the ageing brain.

However, it is unknown how ageing, brain structure, and cognitive function are related, which exaggerates the difficulties in predicting whether an individual is likely to be susceptible to cognitive decline and to what extent. In fact, ageing studies, which are typically cross-sectional, are often not designed to achieve this end; longitudinal studies, which are the most pertinent design to account for the considerable degree of individual variability observed, are less common due to the cost and time restraints. Thus, animal models of healthy ageing are valuable because they develop analogous phenotypes within a shorter lifespan, which reduces the time necessary to complete a longitudinal study. Moreover, there is greater control of environmental and genetic factors in preclinical studies, which can be difficult to account for in human studies.

The Resilient study followed a cohort of 48 male Sprague-Dawley rats who were behaviourally tested and underwent an extensive MRI protocol up to four times, to characterise changes in brain structure and function across the adult lifespan. Half of the rats were also subjected to an environmental enrichment protocol, which involves weekly rotation of toys in the home cage, and a caloric restriction in the form of intermittent fasting. Combined, both interventions are intended to mimic the beneficial lifestyle choices believed to promote healthy ageing. A clinical frailty battery, and behaviour tests, including novel object recognition and attentional set-shifting, were employed to characterise individual ageing trajectories.

Brain structure was assessed with a variety of metrics including a comparison of brain volumes, and mass univariate voxelwise comparisons using voxel-based morphometry, tensor-based morphometry, relaxometry, and diffusion metrics. As structurally similar regions are more likely to be physically connected, this was complemented with the construction of structural connectome graphs, derived from morphometric similarity networks using the multi-modal metrics to characterise 154 brain regions. Changes in brain structure between the groups were most frequently observed at the 5-month time point, which was likely a consequence of the difference in tissue growth.

Brain function was assessed by graph-theoretical analyses of adjacency matrices derived from the regional BOLD signal fluctuations, as measured by multi-echo resting-state fMRI, and regional comparisons of cerebral blood flow maps. Differences were observed between the groups' overall functional connectivity and modularity, particularly at 5 months old. Although the groups did not differ in their overall functional connectivity of the default mode network, the previously observed result of declining connectivity within the default mode network with increasing age was replicated, with the EEDR group deferring their largest decline in DMN covariance later than the decline observed in the control group. Finally, an age-related increase in cerebral blood flow was observed, and variance was observed in the regional trajectories, indicating that the change of perfusion rate is not uniform across the brain.

As there were fewer group differences than hypothesised *a priori*, post-hoc analyses explored the functional and structural changes using physiological and behavioural outcomes instead of experimental grouping. However, the metrics explored did not explain more of the individual variability than the original groups. Consequently, *post-hoc* exploratory analyses used weight, experimental group, and age to predict the observed age-related changes in brain structure and function, revealing substantial residual variance despite controlling for factors controlled by genetics, shared environment, and age. The effect of body weight on brain tissue volume and default

mode network overall connectivity changed with age and is possibly reflective of the changes in metabolic processes that underpin brain growth and remodelling.

Future directions for this work include larger replications, without restrictive timelines, which would better characterise the life outcomes of the subjects. Moreover, future work should measure both cerebral metabolic rate of oxygen and glucose, to measure the relative contributions of metabolic substrates that may underpin the observed differences in brain structure and function in healthy ageing.

ACKNOWLEDGEMENTS

First, thanks to my funders, the King's College London Medical Research Council Doctoral Training Partnership (KCL MRC DTP), for the generosity of their financial support over the last four years. The Resilient study was funded by a grant from the Biotechnology and Biological Sciences Research Council, which was essential for the running of this study.

I am so grateful for my supervisors, Dr Diana Cash and Prof Federico Turkheimer, who adopted me when my initial supervisors left the university, and who provided me with the support to find my way in this project. Thanks to Di for always making time for me, devising plans to send me to incredible places, and defending my work in messy authorship lists. Thanks, FET, for showing me the bigger picture and focusing my analyses on the basics before the hard stuff (which I have found to be applicable in life beyond statistics).

Ultimately, this work could not have been completed without the rats of the Resilient cohort, and the hard work, advice, and patience of the Denmark Hill BSU team, especially Glyn, Jolene, and Anna. Thanks to Karen Randall, without whom I could not have acquired the data. The other half of my scanning and behaviour tag team, who taught me (among many things) the importance of a list. I was always grateful for your patience, but now I am also grateful for your friendship. Dr Eugene Kim, whose tolerance to my inane questions deserves an award, I am indebted for your work on the scan protocol and for giving me a jumping-off point for the image processing. Dr Camilla Simmons, whose SOPs have kept the study running, thank you for your advice and support in all matters of behaviour testing, gently re-acquainting me with a cryostat, and teaching me how to neatly loop cables. Thanks to Dr Michel Bernanos, for being available as an extra pair of hands while scanning and for sharing his analysis expertise. I'm grateful for the wider BRAIN centre and the molecular neuroimaging teams, who have created an atmosphere that I'm proud to share in. Thanks to Dr Paul Wright, who has provided advice to focus my thoughts

since my Master's project. Thanks to Dr Oscar Esteban, Prof Russ Poldrack, and the Poldrack lab, for providing an incredible learning opportunity and lending your expertise, time, and infrastructure to help me to make rodent imaging more reproducible.

To my best friends – the rest of the “Fab Five” – Cheryl, Gillian, Rachael, and Stephanie, who have believed in (and sang about) Dr Eilidh Bailey for a long time. Thanks to Indy, my dog, who begrudgingly stayed up with me until I finished writing for the night, kept my feet warm under my desk, and who generously forewent many scratches to let me focus. To my husband, Andrew, who carried me over the line by making sure that I was fed, watered, and caffeinated, while occasionally insisting on a trip outside the house: you might not be able to retire once I finish this project, but your love, accommodation of my goals, and resolute belief that I can “solve healthy ageing” makes me wish that I could do that for you.

I dedicate this work to my Gran, Isa, whose ageing trajectory was the epitome of resilience and who is forever my inspiration for doing research that can improve quality of life. Finally, this is also dedicated to my parents, Christine and Neil, whose unwavering support and love are exemplified by offers to proofread my work, even though they stopped understanding the content over a decade ago.

TABLE OF CONTENTS

Abstract	2
Acknowledgements	5
Table of Contents	7
1. Heterogeneity in the ageing brain	12
Defining ageing	12
Heterogeneity of cognitive ageing	13
Measuring correlates of ageing with MRI	15
<i>Basic concepts of MRI</i>	16
<i>Age-related changes to brain structure</i>	18
<i>Age-related changes to brain function</i>	21
The importance of longitudinal designs for ageing studies	23
Brain networks	26
<i>The brain as networks</i>	26
<i>Canonical brain networks</i>	29
<i>Heterogeneity in ageing brain networks</i>	31
Summary	37
2. Implementing a longitudinal animal model study	39
Rodent models of healthy ageing	39
<i>Animal models in translational science</i>	39
<i>Animal models of ageing</i>	41
<i>Differences in rodent models</i>	44
The Resilient study protocols	46
<i>Study aims and design</i>	46
<i>Behaviour testing</i>	53
<i>Scanning protocols</i>	60
Generation of analysis resources	64
<i>Age-specific population templates and probability maps</i>	64
<i>Atlas-based brain extraction</i>	67

3. Variance and covariance of brain structure in healthy ageing	72
Rationale	72
Methods	76
<i>Image acquisition and pre-processing</i>	76
<i>Image analysis</i>	80
Results	86
<i>Global volumetric measures</i>	86
<i>Tissue proportions of brain volume</i>	92
<i>Voxelwise volumetric measures</i>	93
<i>Voxelwise relaxometry measures</i>	100
<i>Voxelwise diffusion measures</i>	103
<i>Graph-theoretical analyses of structural networks</i>	108
Summary	117
<i>Increased brain volume with age</i>	117
<i>Regional volume differences</i>	118
<i>Regional microstructural markers</i>	120
<i>Coherence of volumetric and microstructural findings</i>	123
<i>Morphometric similarity networks</i>	124
<i>Multi-modal imaging: benefits and costs</i>	126
<i>Heterogeneity in structural metrics</i>	127
4. Variance and Covariance of brain function in healthy ageing	128
Rationale	128
Methods	135
<i>Functional atlas parcellation</i>	135
<i>Image acquisition and pre-processing</i>	136
<i>Image analysis</i>	138
Results	143
<i>Global metrics of functional connectivity</i>	143
<i>Overall functional connectivity of DMN</i>	148
<i>Cerebral blood flow</i>	151
Summary	155
<i>Group differences in select global functional metrics</i>	155
<i>DMN covariance declined with age</i>	157
<i>Cerebral blood flow increased with age</i>	158
<i>High variance in group-level measures</i>	159

5. Exploring heterogeneity in the Resilient cohort	160
Rationale	160
Methods	163
<i>Lifespan outcomes</i>	163
<i>Behavioural and physical outcomes</i>	163
<i>Statistics</i>	165
Results	167
<i>Description of non-scanning metrics</i>	167
<i>Association of non-scanning metrics</i>	183
<i>How do brain network metrics differ with life outcome?</i>	185
<i>The relative contributions of genetics and shared environment on age-related outcomes</i>	192
Summary	200
<i>Sources of noise in behavioural tests</i>	200
<i>EEDR improves frailty and survival outcomes relative to standard environments</i>	202
<i>Alternative classification did not improve on experimental groups</i>	203
<i>Body weight contributes to brain metrics with different degrees in ageing</i>	206
<i>Residual variance persists after controlling for environment, weight, and age</i>	209
6. Conclusions of the Resilient study	210
Age-related changes in brain structure and function	210
<i>Structural changes</i>	210
<i>Functional changes</i>	212
<i>Other changes</i>	212
Changes associated with EEDR implementation	213
<i>Structural changes</i>	214
<i>Functional changes</i>	215
<i>Other changes</i>	216
The role of diet and enrichment in ageing	217
Limitations and challenges	219
<i>Study design and methods</i>	219
<i>Analysis</i>	222
Future directions	224
<i>Possible alternative biomarkers</i>	224
<i>Large longitudinal studies are necessary to account for individual variance</i>	226
References	228
Appendices	270

List of Appendices

Appendix 2-1: Environmental enrichment sets	270
Appendix 2-2: Frailty index score sheet	271
Appendix 3-1: Example morphometric similarity adjacency matrix	272
Appendix 3-2: Supplemental figures for univariate voxelwise comparisons	273
Appendix 3-3: Supplemental figures for thresholding structural network graphs	275
Appendix 4-1: Mean AUC of functional connectivity metrics	276
Appendix 4-2: Baseline group differences in overall functional connectivity and global efficiency	277
Appendix 4-3: Extended estimates for model predicting cerebral blood flow from age, region, and group	278
Appendix 4-3: Estimated marginal means of the interaction between group and time across regions for cerebral blood flow	281
Appendix 5-1: Physical and behavioural metrics between aged impaired and unimpaired groups	282
Appendix 5-2: Default mode connectivity by outcome	283

List of Figures

Figure 1-1 Fabricated scores (arbitrary units) against chronological age.....	14
Figure 2-1: Novel object recognition results between rat strains.	48
Figure 2-2: Diagram of longitudinal design.....	53
Figure 2-3: Representations of the open field test.	54
Figure 2-4: Attentional set-shifting diagram.	56
Figure 2-5: Diagrammatic representation of the scanning protocol	62
Figure 2-6: Example raw scans from protocol.....	63
Figure 2-7: Example resources generated for the Resilient study.....	65
Figure 2-8: Illustration of the default artsBrainExtraction algorithm	70
Figure 3-1: Example derived structural images	73
Figure 3-2: Median global tissue volume measures.....	91
Figure 3-3: Proportion of tissue volumes relative to total intracranial volume.....	93
Figure 3-4: Volumetric voxelwise differences between groups	98
Figure 3-5: Voxelwise differences for tensor-based morphometry between groups	99
Figure 3-6: Voxelwise differences for relaxometry between groups.....	102
Figure 3-7: Voxelwise differences for diffusion tensor metrics between groups.....	106
Figure 3-8: Voxelwise differences for NODDI model metrics between groups.....	107
Figure 3-9: Group differences in network-based statistics.	109
Figure 3-10: Group differences in network-based statistics.	110
Figure 3-11: Structural connectome cost-efficiency and small-world propensity for a range of densities	111
Figure 3-12: Global graph theory measures for structural networks	115
Figure 4-1: Change in contrast with increasing echo time in a raw image..	131

Figure 4-2: Example component from group ICA showing DMN-like component from resting-state ...	132
Figure 4-3: Coronal slices of functional atlas used for analysis	135
Figure 4-4: Three coronal slices from a representative cerebral blood flow map.....	138
Figure 4-5: Regions of the functional atlas corresponding to the default mode network.....	141
Figure 4-6: Difference in the area under the curve for functional connectivity metrics at 5 months old .	144
Figure 4-7: Difference in the area under the curve for functional connectivity metrics at 11 months old	146
Figure 4-8: Difference in the area under the curve for functional connectivity metrics at 17 months old	147
Figure 4-9: Default mode functional connectivity adjacency matrices	148
Figure 4-10: Overall functional connectivity of the default mode network	150
Figure 4-11: Median regional perfusion rates of cerebral blood flow.....	151
Figure 5-1: Representation of overlapping outcomes between the groups	161
Figure 5-2: Study outcomes of the Resilient cohort.	168
Figure 5-3: Rat weights across time per group	170
Figure 5-4: Frailty index between groups.	173
Figure 5-5: Measures of grip strength	174
Figure 5-6: Locomotor activity results	176
Figure 5-7: Novel object recognition task results.....	178
Figure 5-8: Attentional set-shifting task results	182
Figure 5-9: The relationship of life outcome to other non-scanning metrics.....	185
Figure 5-10: Differences in structural network graph theory metrics between life outcome groups	188
Figure 5-11: Difference in default mode network overall connectivity by life outcome groups.....	190
Figure 5-12: Difference in DMN connectivity between life outcome groups	190
Figure 5-13: Correlation between weight at scanning time points and total intracranial volume	193
Figure 5-14: Ratio of tissue volume normalised for total intracranial volume with increasing weight	194
Figure 5-15: The change in the mean trend between weight and tissue volume	197
Figure 5-16: The change in the mean trend between weight and default mode network functional connectivity.....	198
Figure 5-17: The change in the mean trend between weight and median global cerebral blood flow	199

List of Tables

Table 2-1: Example combinations for attentional set-shifting trials..	56
Table 3-1: Coefficients for the brain volume models	90
Table 3-2: Coefficients for structural network graph theory models.....	116
Table 4-1: Estimates of the DMN connectivity model from group and age.....	150
Table 4-2: CBF conditional growth model contrasts of factors and interactions.....	152
Table 4-3: Contrasting estimated marginal means of median cerebral blood flow	154
Table 5-1: Number of animals per group at each scanning time point and associated measures	164
Table 5-2: Attentional set-shifting task results.	181

1. HETEROGENEITY IN THE AGEING BRAIN

Defining ageing

Ageing usually refers to the linear accumulation of time that an organism has existed. This description is undeniably appropriate, but so-called chronological ageing is not the only definition of ageing. Improvements in medicine, hygiene, and lifestyle lengthened the average global life expectancy by approximately a decade in the latter half of the 20th century (Jeune, 2002; Vaupel, 2010), which has elicited benefits, but also detriments. Increased chronological age carries a greater risk of developing and dying from chronic non-communicable diseases, including cardiovascular disease, diabetes, and neurodegenerative diseases (Beard et al., 2016), which develop from disruption of common cellular mechanisms (Kennedy et al., 2014; Franceschi et al., 2018). For example, aggregates of the amyloid-beta ($A\beta$) protein, which is characteristic of Alzheimer's disease pathology (Hardy & Higgins, 1992), are routinely discovered in a subset of healthy elderly (Rodrigue et al., 2009). Additionally, it has long been known that non-haem iron accumulates in the brain, particularly with increasing age (Hallgren & Sourander, 1958). Iron is necessary for many brain processes, and therefore its accumulation is considered a typical feature of ageing, but aggregation beyond a threshold of unknown magnitude is associated with axonal injury and apoptosis (Kennedy & Raz, 2015).

The age-related, gradual accumulation of microscopic changes to neuronal cell bodies, myelin, and dendrites of neurons, particularly in the frontal regions (Burke & Barnes, 2006; Flood & Coleman, 1988), is likely to be pathogenic, leading to reduced physiological function that is associated with macroscopic changes in brain structure and function. As detailed in the review by Raz & Rodrigue (2006), old brains were smaller in volume and lighter in mass relative to young brains undergoing post-mortem examination, with the sulci and ventricle spaces of old brains undergoing visible expansion. Some, but not all, aspects of cognition show a decline in older age (Bassett & Bullmore, 2017; Salthouse, 2010) which is known as age-related cognitive

decline, otherwise known as cognitive ageing. Cognitive ageing is robustly observed as a life-long decline in the ability to encode long-term memories and the gradual slowing of information processing. Semantic and short-term memory also decline, but not until later in life, while autobiographical memory, theory of mind tasks, and crystallised intelligence seem to be stable across the lifespan (Hedden & Gabrieli, 2004; Salthouse, 2010; Harada et al., 2013).

Heterogeneity of cognitive ageing

A large body of research endeavours to effectively delineate the decline associated with disease from that observed in normal ageing. For instance, it would be beneficial in a study of older people to differentiate between the early stages of mild cognitive impairment (Petersen et al., 1999) and the robust decline described above, but this is not trivial. The existence of age-related diseases is evidence of a relationship between chronological ageing and physiological processes, but other factors likely contribute, since age alone does not predict cognitive abilities.

However, classification is inherently problematic: categorising behaviour as normal (or not) requires a well-defined characterisation of normal. Although the literature has developed average trends for various cognitive abilities, the individual variation in test scores prohibits using these scores as a measure of normality. Variability between individuals is easily demonstrated with fictitious data; figure 1-1 shows a reproduction of such a figure from Salthouse (1986). While the average trend is analogous to the robustly observed decline in scores with increasing chronological age, individual variability manifests as two individuals (green and red markers) with the same chronological age but with vastly different task scores. Conversely, two individuals (green and yellow markers) can have similar test scores but be completely different ages. Thus, the yellow individual in this example is considered to be ‘functionally old’, and the green individual would be ‘functionally young’. Rowe & Kahn's seminal Science paper (1987) argued that ageing research routinely ignored this observation in favour of exploring differences between group averages. They

suggested that lifestyle factors drive individual variability to produce at least two characterisations: usual ageing and successful ageing.

Figure 1-1 Fabricated scores (arbitrary units) against chronological age

While normal ageing is exemplified by individuals who present the robust trends addressed previously (Ylikoski et al., 1999), the evident variability between individuals indicates that success in ageing occurs as a continuum that is defined by the degree of cognitive decline (Depp & Jeste, 2006). The idealistic end of this continuum is healthy ageing, which is ageing chronologically in the absence of disease, such that an individual has minimal symptoms of age-related cognitive decline. Between normal ageing and healthy ageing is successful ageing, which is maintaining cognitive function despite the presence of age-related lesions (Katzman et al., 1988). Individuals at the optimal end of the continuum can be considered ‘super-agers’: those that are over 80 years old but perform at or above average normative values for individuals in their 50s and 60s (Rogalski et al., 2013). For example, particular participants in the formative Nun Study (Snowdon, 2003), which was a longitudinal evaluation of nearly 700 elderly nuns from the same congregation aged from 75 to 102 at the study launch, are illustrative of points on this continuum. Sister Marcella reached old age with little age-related cellular disruption and had

better cognitive outcomes than other members of her congregation, which is characteristic of healthy ageing. Sister Agnes regrettably epitomised the adverse end of the continuum, as she displayed dramatic cognitive and physical decline as a result of age-related burdens. Sisters Matthia and Bernadette displayed preserved cognition into late life, but their post-mortem examinations found evidence of age-related brain changes, which suggests they were successful agers.

Age-related changes, such as protein plaque accumulation, may be indicative of pre-symptomatic cognitive decline (Hedden et al., 2013). However, if this was the case for the previously described nuns who demonstrated successful ageing, it is not clear how long it would be before the symptomatic cognitive decline would be evident. Moreover, due to the correlational nature of many of these findings between brain structure and cognition, it is ambiguous whether brain changes drive cognitive changes or vice versa. While it is undoubtedly intuitive to believe the former, as many of the studies are cross-sectional, it is arguably unfounded (Salthouse, 2011). Given the demonstrable difficulty in classifying ageing, it is even more difficult to predict individual trajectories, such as when a person will experience cognitive decline and to what extent. Due to its gradual nature, the decline onset is difficult to detect; it appears to be a lifelong, dynamic process which begins before middle age rather than starting after passing an age threshold (Salthouse, 2009). Moreover, although chronological ageing is linear, cognitive ageing is not; there is evidence of accelerating decline with increasing age (Singh-Manoux et al., 2012; Salthouse, 2018), and cognitive domains decline at different rates (Nyberg et al., 2012; Salthouse, 2019).

Measuring correlates of ageing with MRI

Historically, combining cognitive measures for an individual at a given point in their life and their physical brain changes has been difficult because accessing the brain's activity and structure required post-mortem examinations or invasive procedures. This problem is exacerbated for ageing research, which relies on correlation due to the inescapable fact that it is impossible to experimentally manipulate chronological ageing. Thus, as neuroimaging technologies such as magnetic resonance imaging

(MRI) became available, an abundance of research used these tools to detect potential biomarkers that can better clarify the underpinning mechanisms of ageing.

Basic concepts of MRI

MRI is a non-invasive imaging modality that allows us to represent the brain and other tissues by measuring the relaxation of protons in a magnetic field (Lauterbur, 1973). Hydrogen nuclei, which are abundant in cellular components, have one proton that imparts a vector-like magnetic dipole moment from the positive charge. Placing the brain in a strong magnetic field (B_0) causes hydrogen nuclei to align their positive charge either in parallel to B_0 (i.e., spin up) or in the opposite direction (i.e., anti-parallel or spin down). Parallel and anti-parallel moments are equal in magnitude and cancel each other out but, since it is the lower energy state, there are slightly more hydrogen nuclei aligned with B_0 at equilibrium. Thus, the vector sum of all the hydrogen nuclei's magnetic moments, known as the net magnetisation (M_0), is aligned with B_0 . As the dipoles align, the nuclei precess around B_0 at the Larmor frequency. Applying a radiofrequency pulse (B_1) that is oscillating at the Larmor frequency and in a perpendicular direction to B_0 will resonate with the precessing nuclei and transfer energy to the system. With the increased energy, the magnetic moments of nuclei in the lower energy state are no longer aligned with B_0 , reducing the longitudinal component of the net magnetisation (M_z). However, the application of B_1 results in unification of the precessing nuclei's phase, which maximises the transverse component of the net magnetization (M_{xy}).

In the absence of further radiofrequency pulses, the nuclei relax back to equilibrium (Bloch, 1946). As the magnetic moments return to the longitudinal plane from the transverse plane, M_0 grows exponentially to its initial magnitude. The longitudinal relaxation (T_1) is the exponential time constant that represents the time for M_z to reach approximately 63 % of the original magnitude of M_0 , such that $M_z(T_1) = M_0(1 - e^{-1})$. Meanwhile, differences in field strength experienced by individual nuclei causes the magnitude of M_{xy} , which was at its peak after application of B_1 produced phase coherence, decays exponentially to zero. Thus, the transverse

relaxation (T_2) is the exponential time constant that represents the time when M_{xy} is approximately 37 % of the maximum value, such that $M_{xy}(T_2) = M_{xy}(0)e^{-1}$. While this relaxation is driven by neighbouring nuclei imparting relative differences of field strength, it is also influenced by inhomogeneity across the wider magnetic field and as such is not truly exponential. Consequently, T_2^* is the time constant representing the time when M_{xy} is approximately 37 % of its maximum value when accounting for field inhomogeneities and the exponential signal decay of M_{xy} .

A delicate balance is required to acquire informative images: both the available signal, which diminishes over time and the differentiation between tissues, which takes time to emerge, need to be exploited. Maximising the signal magnitude reduces the impact of noise that is integral to all measurements; thus, an image with a low signal-to-noise ratio (SNR) is not faithful to the underlying biology. However, relaxation rates differ between tissues based on their composition; grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) contain fluid, fat, and nervous tissue in varying proportions, and these structural differences underpin their distinct relaxation properties. For example, at a field strength of 3 Tesla (3 T), GM has a T_1 relaxation time of ~ 1.3 seconds, but WM matter has a T_1 of ~ 0.8 seconds, and T_2 relaxation times for GM and WM are approximately 110 and 80 milliseconds respectively (Wansapura et al., 1999). Consequently, protocols must acquire images at various timescales in order to maximise these differences, depending on their goal.

The resulting image consists of volume elements (voxels), the size of which, known as the image resolution, also depends on the length of the acquisition protocol. These elements are assigned a value that represents the signal acquired from that volume and reflects the inherent tissue properties of all nuclei sources contained within. These values are on a gradient, and generally, larger signal intensities are displayed as brighter voxels. Small voxel sizes are better at resolving brain structures, as the likelihood that a voxel contains signals from multiple tissue types (known as a partial volume effect) is reduced, but images with high resolution take longer to acquire as they have more voxels. Achieving the highest resolution possible is advantageous for

structural MRI, the goal of which is to resolve tissue structures in order to represent the underlying anatomy as accurately as possible. However, the long scan times can be undesirable for *in vivo* subjects, as this increases the likelihood of the subject experiencing discomfort and moving. Images with high spatial resolution are also wholly unsuitable for recording brain activity at much shorter timeframes. To represent the changes in magnetisation in the brain across a shorter timeframe, as is necessary for functional MRI (fMRI), it is generally necessary to achieve an optimal compromise between resolution and speed of acquisition. Consequently, fMRI has a lower spatial resolution, but its faster temporal resolution can accurately capture brain magnetisation changes on the order of seconds or less (Feinberg et al., 2010).

Unsurprisingly, due to its non-invasive application, sensitivity to the underlying biology, and the flexibility of the paradigms to measure both structural and functional changes, MRI has become a prevalent tool to interface between the brain and behaviour to investigate mechanisms of ageing.

Age-related changes to brain structure

Structural correlates of ageing have been investigated with MRI studies using various methods of quantification. By hand-drawing and measuring the volume from regions of interest (ROIs) it was shown that prefrontal cortical grey matter is the most affected neocortical region, with moderate changes to the temporal cortices, and even fewer age-related effects on the parietal and occipital cortices (Raz, 1997; Gunning-Dixon & Raz, 2000). This finding was recapitulated by a study measuring cortical thickness, where the distance between the grey-white matter boundary and the outer cortical surface suggested the region subject to the most extensive atrophy was the prefrontal cortex (Salat, 2004). Importantly, these findings of grey matter volume reduction have since been replicated using semi-automated and automated image processing pipelines, as they have become available. For instance, widespread age-related differences in grey matter volumes were exposed using voxel-based morphometry (Good et al., 2001) and automatic segmentation (Walhovd et al., 2005). However, beyond the findings that grey matter volume declines with age and that

the prefrontal cortex is commonly affected, the results between studies vary greatly, particularly for regional patterns of age-related atrophy. For instance, Salat and colleagues (2004) found that the temporal cortices were relatively spared but that parietal cortices were affected. In the study by Walhovd et al. (2005), almost all structures segmented showed an age-related decline, while in the study by Good and colleagues (2001), parietal and cingulate cortices were most affected, and the temporal cortex and associated structures were preserved. All of these studies acknowledged large amounts of heterogeneity in brain morphometry at an individual level, suggesting that the variability seen in cognitive ageing is also reflected in brain structure. In addition to cortical regions, a decline in hippocampal volume has also been observed in healthy ageing (Raz et al., 2004a), but not to the extent of the atrophy observed in Alzheimer's disease (Jack et al., 2000). Frontal and hippocampal regions also show decreases in T_2^* with increasing age, which is not observed in the primary visual cortex (Raz et al., 2007). The authors hypothesise that the shortening of the T_2^* signal measured in these regions demonstrates changes in the concentration of deoxygenated blood, rather than non-haem iron accumulation which collects less frequently in the regions with demonstrable changes.

White matter volume changes are also replicated across ageing studies, with the volume either increasing or remaining steady until it begins to decline in middle age (Raz & Rodrigue, 2006; Michielse et al., 2010). This trajectory appears to be modulated regionally, with prefrontal volumes declining non-linearly (Bartzokis et al., 2003; Salat et al., 1999). For instance, prefrontal white matter and the fornix are regularly reported to decrease in volume with age, but the rate of decrease is not constant (Walhovd et al., 2005; Pagani et al., 2008; Michielse et al., 2010). However, the most straightforward change does not require computational analysis to be seen, as it is observed by the human eye. White matter hyperintensities are so-called because they appear bright on T_2 -weighted (T_{2w}) images and are often spotted in periventricular white matter (Brant-Zawadzki et al., 1985). They are likely caused by a reduced blood supply in cerebral small vessels but their histopathological impact is

variable, so they are associated with heterogeneous cognitive outcomes (Prins & Scheltens, 2015). WMH are common in brain scans of individuals in their forties and increase with increasing age (Wen et al., 2009).

While relaxometry can assess macrostructural features of white matter volumes (Moseley, 2002), diffusion-weighted imaging (DWI) and diffusion tensor imaging (DTI) are related methods that measure the directionality of water proton diffusion. These methods are sensitive to the direction of diffusion, particularly along the length of white matter axons, which is inherently anisotropic as a consequence of the myelin envelopment on all sides except the primary direction of diffusion (Moseley, 2002; Sullivan & Pfefferbaum, 2006). Diffusion imaging is well-placed to measure microstructural integrity of white matter; for instance, the average diffusion coefficient (ADC), which represents water proton motility regardless of direction, is higher in the frontal corpus callosum with age than in younger participants, while decreases in fractional anisotropy (FA; a measure of diffusion directionality) and increases in mean diffusivity (MD; a measure of diffusion magnitude) are repeatedly observed in ageing white matter (O'Sullivan et al., 2001; Burzynska et al., 2010; Sullivan et al., 2001).

Complimentary to the observed volume changes in grey and white matter, an increase in the size of CSF compartments with chronological ageing is consistently observed, but there is a discrepancy between whether this relationship is linear or curvilinear (Good et al., 2001; Walhovd et al., 2005; Michielse et al., 2010). Diffusion data also demonstrate increased CSF volume; in particular, histograms generated from the ADC values across all brain voxels showed a lower peak ADC in old brains, and histograms from old brains had a longer tail, indicating a decrease in restricted diffusion and an increase in free diffusion due to a larger relative proportion of the brain being fluid rather than tissue (Klose et al., 2013).

Age-related changes to brain function

Neurons have considerable energy requirements to maintain brain function. Despite accounting for 2 % of the body's overall mass, the brain uses 20 % of the energy. Furthermore, almost all of the brain's energy is derived from oxidative phosphorylation, and the majority of that energy is used for neuronal excitation and signal conduction, thus demanding constant supplies of glucose and oxygen; the former is provided by diet and the latter from cerebral blood flow (CBF; Clarke & Sokoloff, 1999). The brain's proclivity to generate energy from glucose in the presence of oxygen engenders an intricate coupling of cerebral blood flow and neuronal activity, in order to replenish rapacious brain cells with oxygen supplies. Thus, an increase in blood oxygenation is indicative of increased neuronal activity (Roy & Sherrington, 1890). Haemoglobin binds oxygen to its haem component in order to transport it from the lungs to the tissue, and it is known in its bound form as oxyhaemoglobin. At the tissue, the oxygen is released, and the haem is transformed into its deoxygenated state, deoxyhaemoglobin. The binding and release of oxygen alters the chemical structure of the haem group and subsequently adapts its magnetism: oxyhaemoglobin is diamagnetic, while deoxyhaemoglobin is paramagnetic. Consequently, MRI that is sensitive to changes in the magnetic field, such as T_2 and T_2^* relaxation, is employed to measure changes in blood oxygenation as a representation of neuronal activity.

Cerebral perfusion is used as a proxy to measure brain activity, through measurement of the cerebral blood volume and flow. The latter is commonly achieved using a technique called arterial spin labelling (ASL) which uses arterial blood 'labelled' with a separate radiofrequency pulse to invert the net magnetisation of the protons arriving in the brain. The inversion of the net magnetisation reduces the signal available in the subsequently acquired image, providing an endogenous tracer of blood perfusion (Haller et al., 2016).

Justifiably, MRI-derived measures of perfusion have been employed to measure changes in neuronal activity in an ageing population. Perfusion in anterior cerebral

grey matter regions is known to decrease with increasing age, even in cognitively normal individuals (Naritomi et al., 1979; Gur et al., 1987). A negative correlation between age and CBF has also been quantified directly with MRI using ASL protocols (Chen et al., 2011a). Chen and colleagues (2011) found that cortical perfusion reduced with increasing age, while subcortical perfusion was largely preserved, and these regional differences were independent of age-related cortical thinning. Moreover, they found that the observed decrease in cortical blood flow was not linear, and there was a faster rate of decline after middle age. This decline and the independence of grey matter and cerebral blood flow decline trajectories has been replicated, and there was evidence that these trajectories together were better predictors of various cross-domain cognitive scores than age alone (Steffener et al., 2013). Grady and colleagues (Grady et al., 1998) found that younger adults could modulate the right prefrontal cortex blood flow response to the increasing difficulty of a working memory task better than their older counterparts. In contrast, older participants increasingly recruited regions of the left prefrontal cortex at all difficulties.

Another popular MRI measure of brain function uses the Blood Oxygen Level Dependent (BOLD) signal fluctuations as a proxy for the changing oxygenation state of cerebral haemoglobin. There is a relative increase in the T_2^* relaxation-related signal when the concentration of deoxyhaemoglobin decreases (Ogawa et al., 1990), so the BOLD signal has been used to measure neuronal activity both during tasks, such as cognitive batteries, and in the absence of tasks, otherwise known as ‘rest’ (Raichle, 1998; Gusnard & Raichle, 2001).

Changes in the BOLD signal during tasks and at rest have been extensively studied in ageing research. As fMRI provides an opportunity to non-invasively examine neuronal function, albeit indirectly, it is well placed to quantify age-related cognitive decline. When undertaking executive processing tasks in a scanner, older adults demonstrate less neuronal activity in the prefrontal cortex compared to young adults (Stebbins et al., 2002). Furthermore, in memory encoding tasks, for which young

individuals do not recruit prefrontal regions at all, older participants showed increased bilateral prefrontal activity relative to young participants (Logan et al., 2002). The shift in regional recruitment strategy across cognitive domains is also more symmetrical between hemispheres than the pattern observed in young adults, thus inspiring the Hemispheric Asymmetry Reduction in OLDER adults (HAROLD) model (Cabeza, 2002; Hedden & Gabrieli, 2004). Additionally, young adults recruit occipitotemporal regions more than their older counterparts, indicating there is a Posterior-Anterior Shift in Ageing (PASA) that occurs independently of difficulty and task (Davis et al., 2008). There is still debate regarding whether age-related changes to regional recruitment represent a consequence of age-related changes, known as dedifferentiation, or a compensatory mechanism to offset age-related changes. Cabeza and colleagues provided evidence that the compensation hypothesis is a strategy to counteract inefficiency due to age-related changes (Cabeza et al., 2002), which complements the compensation-related utilisation of neural circuits hypothesis (CRUNCH) suggested by Reuter-Lorenz & Mikels (2006), that suggests an older brain requires more neural resources for the same task than a younger brain.

The importance of longitudinal designs for ageing studies

Chronological ageing alone fails to predict the structural and functional proxies for physiological and cognitive ageing addressed above, despite their relationship with age. This disparity is likely a consequence of the heterogeneity observed between individuals in both these proxies and cognitive performance. Salthouse (2011) discusses the abundance of evidence for a complicated ‘ABC’ relationship: there are replicable correlations between age (A) and brain structure (B) changes, brain structure and cognition (C) changes, and age and cognition changes. However, despite plenty of studies, the causality and type of relationship between these factors are still unclear. For instance, although there may be a direct relationship where A influences B which consequently mediates an effect on C, it is currently also likely that the inverse is correct in that C mediates an effect on B as a consequence of A, as suggested by cognitive measures which better predict functional outcomes than

vice versa (Zahodne et al., 2013). Furthermore, Salthouse points out what has already been described here: the relationship between age and brain structure is often weak, and the causality of the relationship between brain and cognition is unverified. It is as likely that there is an unknown mediator, or that A influences B and C independently. It is difficult to understand which is the influencing factor that drives these relationships, particularly when utilising cross-sectional study designs.

In response to Salthouse's discussion, Raz & Lindenberger (2011) argue for the importance of a longitudinal design for ageing on this account. The between individual variability can be addressed with longitudinal studies: these studies scan the same person multiple times to tease apart intraindividual and interindividual differences and the variability associated with each (Schaie & Hofer, 2001). As discussed for the nun study, structural proxies may occur first, which leads to structural changes before cognitive changes. For instance, depositions of beta-amyloid protein are associated with Alzheimer's disease and MCI, but it has been shown that 20-30 % of healthy older adults have a considerable plaque load (Rodrigue et al., 2009). Individuals that are cognitively normal with high plaque loads may be pre-symptomatic at the time of the study, thus potentially biasing a healthy control group (Jack et al., 2008). Furthermore, age-related atrophy occurs with different trajectories and onsets for different regions within the same person, so it is not unreasonable that it occurs at different rates and at different times between people. These differences encumber appropriate matching young and old participants in group comparisons, particularly in cases that include latent generational effects that influence any of the variables (Gerstorf et al., 2015). Likewise, statistically, it is poor practice to categorise continuous variables (Royston et al., 2006), and ageing research is no exception. While an argument can be made that there will be observable and important differences between groups of young and old individuals, it has been demonstrated that ageing is a dynamic and heterogeneous process. Thus, it is crucial to inspect the nuances of individual trajectories rather than

insensitively collecting individuals to devise a mean value, attempting to predict the subtle changes that may be happening at a given time point.

Findings from longitudinal studies of ageing generally confirm those from previous studies but also highlight new information. For instance, a study by Fjell et al. (2009) found evidence of atrophy in otherwise healthy individuals aged over 60, relative to a scan only one year before. The largest atrophies occurred in the prefrontal cortices and the hippocampus, while the smallest changes were in the visual cortex. Longitudinal studies also concur that the rate of decline is ostensibly non-linear; several longitudinal studies found increased CSF volumes and increasing rate of shrinkage for prefrontal white matter with increasing age (Raz et al., 2005, 2004b; Pfefferbaum et al., 2013; Fjell et al., 2009; Nyberg et al., 2010). Furthermore, a meta-analysis of 200 longitudinally-acquired images that aimed to explore the discrepancy between cross-sectional and longitudinal findings regarding hippocampal decline supports the observation that brain volume changes are non-uniform and rates of atrophy are non-linear, especially after the age of 60 (Pfefferbaum & Sullivan, 2015). Longitudinal fMRI studies also support the observation that posterior regions, such as the occipital cortex, are functionally under-recruited with increasing age (Nyberg et al., 2010). Additionally, the BOLD signal in the hippocampus during a facial recognition task compared to rest was measured in a longitudinal study (O'Brien et al., 2010). The change in signal was found to decrease only in individuals who experienced cognitive decline between scans, and not individuals who maintained their status as 'clinically normal'.

However, these studies are not infallible, and there are still discrepancies between studies; some longitudinal studies highlight the decline of the cerebellum and caudate (Raz et al., 2005), but these regions have the smallest degree of decline in another study (Fjell et al., 2009). Moreover, Nyberg et al. (2010) found differences between cross-sectional and longitudinal results with regards to frontal recruitment: while cross-sectionally, successful adults did use frontal regions more than unsuccessful

adults, but when they were followed longitudinally, even these adults showed a reduction in BOLD signal in frontal regions over time.

It is an essential caveat that longitudinal studies have limitations too: they are expensive, financially and in resources. In ageing research, the experimental manipulation is chronological age, so the passage of time is necessary to observe the desired outcome. Participant retention is difficult, and since dropouts are more likely to occur in frailer individuals, this inherently biases the study by increasing the homogeneity of the studied group. It is also common, particularly when measuring changes in cognition, to repeatedly expose a participant to the same test at different time points. This practice engenders undesirable training effects, which may inflate test scores by providing practice and multiple attempts to find an efficient strategy. Finally, longitudinal studies are often constrained to the acquisition technology, i.e., hardware and software, that is available at the beginning of the study, because adapting scanners, even within the same vendors, or updating sequences have been found to alter the estimation of brain volumes (Lee et al., 2019).

Despite the implementation of longitudinal designs, age-related changes in cognition cannot be successfully explained by the over or under activation of a single given region. Fortuitously, the field of neuroscience's shift towards the collection of 'big data' coincided with a boom in interest for characterising the physics of complex networks (Bassett & Sporns, 2017). These fields have given rise to a multidisciplinary field that leverages the high volume of MRI data to generate models of information transfer between brain regions, which has been postulated as a framework for understanding brain structure since the 19th century (Bullmore & Sporns, 2009).

Brain networks

The brain as networks

Cognition is essentially an outcome of the sharing of information between distinct brain regions which synchronously transfer and integrate, assess and produce

information to create appropriate actions. Historically, most higher brain functions were believed to be served by a particular region, which would manifest as alterations in behaviour when they were injured. Lesion studies provide an essential understanding of the relationship between brain and cognition, and Broca and Wernicke's aphasic patients, amnesiac Henry Molaison, and the personality changes of Phineas Gage have justly become textbook examples of cognitive neuropsychology. This view that brain regions are solely responsible for a function has fallen out of favour, due to the evidence that there is not a distinct region for each function, and some functions are distributed across many regions (Fox & Friston, 2012). It is clear that the brain is neither functionally segregated nor integrated, so a middle ground has emerged in the form of modular networks, which each pertain to a particular function, which integrate and interact with each other, otherwise known as a small-world topology (Bassett & Bullmore, 2006). The characterisation as a small-world network has been quantified by mathematical network theory study, which examines the relationships between connected objects using graphs of the relationship structure; graph theory has been lucratively applied in neuroscience to describe and model the connections within and between networks.

Brain networks can be described as a collection of nodes, where each node represents a brain region. Brain regions can be defined on different scales, such that a region can be derived from predefined histological atlases with various degrees of parcellation, or on a voxelwise basis which is therefore determined by the resolution of the image. Each node is connected, either functionally or structurally, by edges. Edges are assigned a weight which represents the strength of the connection between each node, and these edge weights are frequently derived from the covariance of structural metrics from two nodes, such as diffusion scalars or grey matter volume, or functional metrics, such as the cross-correlation of the BOLD signal time series from two nodes. Structural connectivity relates to the bundles of white matter which transmit information between distinct brain regions since all communication is

underpinned by some physical connection, even if indirectly. On the other hand, functional connectivity measures the synchronicity between BOLD signal fluctuations which are thought to be representative of brain activity. Although functional networks have been shown to reflect the underlying white matter architecture (van den Heuvel et al., 2009), being structurally connected does not inherently qualify for functional connectivity, and not all regions in a functional network are directly connected by a white matter pathway. Therefore, these should be considered separately to understand their relative contributions to cognition.

Networks can be described as being small-world if they have a high local clustering coefficient and short characteristic path length when compared to a randomly-connected network (Bullmore & Sporns, 2009; Watts & Strogatz, 1998). The local clustering coefficient, C , measures how connected a node's neighbours are. For a node, i , which has k_i neighbours, the nodal clustering coefficient, C_i , is given by:

$$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

where k_i denotes the node's degree, i.e., the number of edges connected to i , and e_i is the number of edges between the neighbours. To calculate the clustering coefficient for an undirected graph, $G = (V, E)$, with nodes, V , and edges, E , it is simply the average of the nodal clustering coefficients:

$$\bar{C} = \frac{1}{V} \sum_i^V C_i$$

The distance, d , between two nodes $v_i, v_j \in V$, is the number of edges between them. Thus, the characteristic path length, L , is the average distance between pairs of nodes in a graph with n nodes:

$$\mathcal{L} = \frac{1}{n \cdot (n - 1)} \cdot \sum_{i \neq j} d(v_i, v_j)$$

By comparing local clustering coefficient and characteristic path length to the same measures derived from an equivalent random graph (C_{rand} and L_{rand} , respectively), one can quantify small-worldness, σ :

$$\sigma = \frac{C/C_{rand}}{\mathcal{L}/\mathcal{L}_{rand}}$$

A network has a small-world topology if $\sigma > 1$ because the clustering coefficient should be much greater than in a random network ($C \gg C_{rand}$) while maintaining approximately the same short path length ($L \approx L_{rand}$). In the seminal paper by Watts & Strogatz (1998), they outlined these equations and demonstrated that the neural network of the *Caenorhabditis elegans* worm exhibited small-worldness, and this has since been replicated across species suggesting that this is a conserved property of brain networks (Bullmore & Sporns, 2009).

As detailed by Bassett and Bullmore (2006), there are many other reasons why the brain should be considered as a small-world network. The complexity of the network, which is simultaneously capable of integrating distributed information and utilising modularity for specific functions, is observed over different spatial and time scales. Moreover, brain networks demonstrate a trade-off between efficiency and cost (Bullmore & Sporns, 2012): the most efficient brain network would have all nodes connected directly (thus reducing the characteristic path length), but this would be costly to maintain in terms of energy and compositional material. Conversely, a network with minimal connections would require less energy to run and less material to build, but this would be detrimental to network efficiency; the sparse connections would leave nodes without edges or as a ring lattice, where the average path length scales linearly with the size of the graph (Fornito et al., 2016).

Canonical brain networks

The popularity of brain network analysis has been bolstered with the improvement of methodologies that permit *in vivo* reconstruction of connections that are known from post-mortem inspection and lesion studies. Tractography allows *in vivo* virtual

dissection of white matter connections using the diffusion tensor to determine the structural orientation of tracts (Basser et al., 2000); the resulting tracts derived from virtual dissection replicate known white matter pathways (Wakana et al., 2004). Functional networks are more difficult to reproduce since, before neuroimaging techniques, large-scale brain function could only be measured with invasive techniques. However, measuring correlations and anti-correlations of the synchronous fluctuations in BOLD signal during tasks and at rest, known as functional connectivity (Friston, 1994), can infer continual communication between spatially distinct regions in patterns that replicate various hypothesised cognitive networks. While it is possible to extract these networks from signals derived from the brain undergoing a task (Rissman et al., 2004), networks derived from resting-state fMRI have proved popular, likely due to the capability of acquiring information about multiple networks simultaneously without extra tasks or equipment (O'Connor & Zeffiro, 2019). Resting-state networks were first described as task-independent attenuations in the BOLD signal which lead researchers to consider a baseline resting brain state (Gusnard & Raichle, 2001). Upon further examination, a canonical set of networks could be reproduced from low-frequency (< 1 hertz) amplitude fluctuations in fMRI time courses collected when individuals are in the absence of a task (Damoiseaux et al., 2006).

Of these functional networks, the most studied is the default mode network (DMN) which anatomically comprises the medial prefrontal cortices (MPFC), posterior cingulate (PCC) and retrosplenial (RSC) cortices, the hippocampal formation, inferior parietal cortices, and lateral temporal cortices (Buckner et al., 2008). The DMN characteristically displays decreases in BOLD signal during goal-directed activities and increases during rest. Network analysis of the anatomical regions previously listed shows that it has a modular structure: the midline core, which consists of the PCC and anterior MPFC, integrates information between two subsystems that consist of the temporal and remaining prefrontal components respectively (Andrews-Hanna et al., 2010). This network is functionally ambiguous,

at least compared to other networks such as visual processing and motor networks, but has been associated with ‘spontaneous cognition’ or processing of self-relevant information; it is shown to be active during self-reflection, meditation, and mind-wandering (Greicius et al., 2003; Andrews-Hanna et al., 2010; Brewer et al., 2011; Weissman et al., 2006).

Heterogeneity in ageing brain networks

Task-based networks

While the components of brain networks are still robustly reproduced in ageing (Beason-Held et al., 2009), both resting and task-based brain networks are dynamic and, much like other facets of ageing, show individual differences over time between and within individuals. Researchers have attempted to explain the changes in the recruitment patterns of fMRI activations and deactivations mentioned previously, which have now been designated ‘classical theories’ (Sala-Llonch et al., 2015). The HAROLD, CRUNCH, and PASA theories were generated from observing regional over-recruitment in different tasks. However, there was no clear consensus whether these mechanisms were the result of a successful strategy to overcome deficits, known as compensation, or simply a reduction of functional specialisation, known as dedifferentiation. Despite the debate around which theory best describes the nature of why frontal over-activation arises with age, the increased activity seen in frontal brain regions has been replicated across multiple executive control tasks (Turner & Spreng, 2012), suggesting this is a robust hallmark of ageing brain function linked to age-related cognitive decline.

With the field at an impasse regarding which model best explains age-related changes in the brain, a multi-modal approach, known as the scaffolding theory of ageing and cognition (STAC), recognises that both structural and functional brain changes are relevant factors in ageing brains and that these factors may have external influences (Park & Reuter-Lorenz, 2009). Crucially, STAC suggests that a supply of neural resources can be added to and diminished via various means. It is proposed that

ageing increases neural challenges and functional deterioration, like atrophy and changes in regional recruitment respectively, and these, in turn, have influences on each other. An individual's scaffolding reserves might be capable of mitigating the effect these have on cognitive function, but this depends on the capacity of the scaffolding resources (such as neurogenesis and regional recruitment strategies for executive functions). Furthermore, these resources can be bolstered by cognitive engagement and exercise, which the authors believed could explain some of the individual differences seen in ageing research. Moreover, this theory was later revised and is known as STAC-R, to recognise the enriching influence of lifestyle factors on neural resources and to reflect new knowledge regarding instruments of neural resource depletion (Reuter-Lorenz & Park, 2014). For instance, physical fitness and enriched leisure activities are neural enriching activities, while depression and stress are neural depleting. This theory has likely been successful because of its compatibility with other theoretical concepts, including cognitive reserve (Stern, 2002, 2009; Tucker & Stern, 2011) and maintenance models (Walhovd et al., 2014).

Cognitive reserve was born from the observation that if two people experienced the same injury, age-related or otherwise inflicted, the severity of the injury is not sufficient to predict long-term clinical outcomes accurately. Early observations suggested that larger brains or more neurons would be better at withstanding insults because there was more substance to deplete before a threshold was passed and deficits begin to materialise. Instead, cognitive reserve theory takes an active view of reserve and instead suggests that an individual's resilience is rooted in the available resources with which they can achieve adaptability to bring about brain changes. The evidence for this includes the somewhat counterintuitive finding that those with better cognitive reserve have been found to have worse outcomes in Alzheimer's disease, such as dying sooner after diagnosis. The apparent contradiction is explained as a delayed diagnosis of the disease because the brain was utilising its reserve materials until they were exhausted, prolonging the onset of the phenotypical outcomes (Stern, 2009). Stern suggests two paradigms that may occur with cognitive

reserve; neural reserve is the ability to progressively adjust the efficiency of recruited regions to match task difficulty, while compensatory reserve describes recruiting alternative or additional networks in disruption, which is highly compatible with STAC-R hypothesis.

Similarly, brain maintenance suggests that the approach to ageing research should be ‘dimensional rather than categorical’ and reflect the multiple systems associated with chronological ageing. Walhovd and colleagues (2014) suggest that brain ageing comes from vulnerability in several systems, such as atrophy of brain structures, aberrant functional network recruitment, and amyloid deposition, which is conceptually similar to age-related diseases developing from multiple cellular mechanisms. This hypothesis rationalises the lack of a single unifying ageing aetiology and the vast heterogeneity in brain function, structure, and cognition in ageing. Importantly, it is both environmental and genetic factors which are likely to be driving forces in cognitive decline, and these can influence ageing at any time across the lifespan, which further demonstrates the necessity of longitudinal studies, rather than cross-sectional, to reduce inhomogeneity of latent variables. Furthermore, it is also compatible with the STAC-R hypothesis, which places an importance on the additive and detrimental effects of neural resources across the lifespan and not just once past an arbitrary ageing threshold.

Resting-state networks

Analysis of resting-state fMRI has facilitated a new window into ageing brain dynamics because it is no longer limited to a cognitive task battery. For instance, despite still demonstrating small-world topology, older participants demonstrated less efficient brain networks at rest, particularly in frontal and temporal cortices and subcortical regions, compared to their younger counterparts (Achard & Bullmore, 2007). Many studies have found decreased connectivity within the DMN (Sala-Llonch et al., 2015). However, Jones et al. (2011) found that DMN’s anterior components showed regional differences in functional connectivity with ageing, such as increases in some areas and decreases to others, but the posterior components

exhibited almost exclusively age-related decreases in connectivity. Other resting-state networks have been examined, but their outcomes have been less conclusive; for example, connectivity within the dorsal attention network (DAN) has been found to increase with age in one study and decrease in another (Andrews-Hanna et al., 2007; Tomasi & Volkow, 2012). Evidence from longitudinal studies are mixed: one finds comparable declines within the DMN and the executive control network (ECN) (Ng et al., 2016) while one found no difference in intra-DMN connectivity over time (Persson et al., 2014).

In addition to a reduction of within network connectivity, older adults conversely display more between network connectivity than younger adults (Damoiseaux, 2017). Age-related integration of brain networks is exemplified by the finding of Spreng and colleagues who compared the complementary relationship between the DMN and DAN, which presents as an anti-correlated couplet in young adults. The anti-correlation was reduced in older adults, suggesting more synchrony between these asynchronous networks, and the within-network correlations derived from the components were reduced, suggesting increased asynchrony within networks (Spreng et al., 2016). As a consequence of more synchrony between opposing networks, it is implied that these networks are less segregated than in young adults. A decrease in network modularity (i.e., the segregation of connections into functional modules) has been reported by several studies, including a longitudinal study (Ng et al., 2016). Grady and colleagues described age-related decreases in all graph-theoretical measures, including modularity and clustering, and greater between-network connectivity in older adults (Grady et al., 2016). Similarly, decreased modularity is reflected as a reduction in local efficiency, suggesting that neighbouring nodes are less-well connected (Geerligs et al., 2015). However, they are not synonymous as it appears that local efficiency appears to follow an inverted-U shape across the lifespan while modularity linearly decreases (Cao et al., 2014). Cao also showed lower rich-club coefficients in older adults than in young adults, which describes the strength of connections between high-degree nodes and thus supports

decreased communication between modules. Furthermore, Cao and colleagues found decreased synchrony between long-range connections, which is replicated by a study that found weaker inter-regional correlations, particularly between the frontal and parietal regions (Sala-Llonch et al., 2014).

However, other findings from the replication by Sala-Llonch and colleagues are contrary to the majority of the literature: they observed more global clustering and increasingly segregated brain architecture in ageing networks, which is epitomised by decreasing global efficiency and longer minimum path. Long path lengths are indicative of decreased global efficiency because traversing the network requires more steps. Chan and colleagues (2014) found an increase of global efficiency with increasing age, which is in agreement with the majority of studies, and also offers an explanation for the opposing findings: there is variability in the method chosen to define nodes. For instance, Sala-Llonch and colleagues used atlas definitions to parcellate nodes and compared inter-regional connectivity. In contrast, Grady restricted their analysis to spherical ROIs placed at a stereotaxic co-ordinate from previous studies of the DMN and DAN and measured inter-network connectivity. Additionally, there is evidence that the number of modules does not change with age, but the composition of modules shift which could account for some of the conflicting evidence (Meunier et al., 2009).

It is notable that Chan and colleagues also found positive associations with functional connectivity and episodic memory scores when controlling for age, suggesting that age does not solely mediate the relationship between functional connectivity and performance on memory tasks. As discussed before, this cross-sectional result should be taken with caution, but it deters the concept that all functional changes are direct consequences of ageing alone.

Structural connectivity

Many age-related network changes have been derived from resting state, or often at least functional MRI, but the transmission of information between spatially distinct

regions relies entirely on physical connections which facilitate interactions. Structural networks can reflect the integrity of these connections by using measures of white matter integrity, i.e., diffusion scalars, as the edge weighting. Networks derived from structural properties recapitulated features of functional networks such as displaying small-world topology independent of age and the high-connectivity of DMN regions such as the precuneus and PCC (Gong et al., 2009). Moreover, diffusion-derived structural networks also display decreases in their rich club integration with increasing age, particularly effecting the frontal hubs (Zhao et al., 2015). While the study by Zhao and colleagues found decreases in both local and global efficiency in older adults compared with young adults, Gong and others found only a significant decrease in lower local efficiency, as global efficiency was preserved in ageing, which would suggest that modular topologies are robust to regional changes.

Similarly, grey matter volume is used as a proxy for the structural integrity of the nodes themselves, rather than the connections in diffusion-derived matrices. Thus, measuring grey matter volume covariance may indicate which regions are undergoing volume change as a consequence of shared mechanisms, such as development and maturation (Alexander-Bloch et al., 2013). In ageing, this covariance was reduced in older adults compared to young counterparts in the DMN, ECN, and language networks (Montembeault et al., 2012).

Integrating structural and functional connectivity

While structural and functional networks share many overlapping characteristics, they are describing complementary but separate information. Reductions in functional connectivity are associated with poor white matter integrity, even in a subset of older adults free of a significant amyloid burden (Andrews-Hanna et al., 2007). Moreover, strong functional connections are observed between regions which are not directly linked by any known white-matter bundle, ostensibly accentuating indirect connections through intermediary regions such as the PCC and other rich-club hubs (Honey et al., 2009). Therefore, these modalities provide independent

information about the system, but joint analysis provides complementary information that is unattainable from either modality alone.

The correlation between structural connectivity and functional connectivity decreases in most regions with increasing age (Zimmermann et al., 2016). A possible mechanism for this loss is that the relationship is perturbed by reorganisation: when comparing microstructural properties of the cingulum bundle and the BOLD covariance between the MPFC and PCC, a correlation between structural and functional connectivity was only observed in adults with low white matter hyperintensity burdens. Individuals who had high WMH burden did not show a relationship between cingulate microstructure and intra-DMN connectivity, and the larger differences between these measures were related to worse memory and executive function test scores (Reijmer et al., 2015). The predictive effect of microstructure on DMN connectivity was specific to the cingulate. Inferior longitudinal fasciculus microstructure, which connects the temporal and parietal lobes, did not replicate the observed correlations.

Summary

Chronological ageing, i.e., the linear accumulation of time, is associated with many brain changes on multiple scales and across multiple systems. Unearthing the mechanisms underpinning the most detrimental of these age-related changes has been impeded by vast inhomogeneity in many aspects. For instance, changes in cognition, structure, and function are apparently dynamic and non-linear, so it is difficult to compare progression between individuals. One particular problem with much of the ageing literature is that it is derived from cross-sectional studies which cannot control for this individual variation and also use categorical groupings of ‘young’ and ‘old’. In response to these points, using longitudinal or quasi-longitudinal study designs may be able to address some of the individual variation points. However, if using cross-sectional designs, age must be controlled as a continuous variable instead of a dichotomy which does not represent the dynamic nature of age-related changes.

There are undoubtedly relationships between chronological ageing, brain structure, brain function, and cognition, but attempts to determine causation or mediation of these relationships have been mostly unsuccessful. While chronological ageing is linear, structure and function of the brain are in a life-long state of flux. Moreover, changes in various cognitive domains seem to be lifelong non-linear declines, particularly in memory and executive functions. Age-related changes appear to be domain-specific, because theory of mind and facets of intelligence do not seem to decline, even in very old age.

Exploring age-related changes in brain structure and function was difficult before MRI, which has proven an invaluable tool to non-invasively explore age-related brain changes *in vivo* that, in turn, can be related to age-related changes in cognition. MRI-derived changes in structure implicate atrophy of the prefrontal regions and hippocampus, with decreased white matter integrity in frontal regions, and increases in CSF compartments. Frontal regions also show declines in functional activations, but when there were several competing hypotheses based on brain activation, none prevailed. However, a well-timed rise in popularity of network analyses helped to change the view of brains from either segregated systems, which carry out particular modular functions, or distributed processing systems, which integrate all their information to small-world networks which economically trade-off integration and segregation in order to minimise energy costs but maximise efficiency. While there are still discrepancies between findings employing these analysis methods, a consensus has emerged that with increasing age there is decreased intra-network connectivity, particularly within the DMN, and more inter-network connectivity, possibly as the brain's modular structure is reorganised to accommodate age-related changes. Crucially, multi-modal imaging provides overlapping, but distinct information which is, therefore, a considerable benefit to incorporating multiple modalities to probe the complicated relationship that occurs between the brain and ageing.

2. IMPLEMENTING A LONGITUDINAL ANIMAL MODEL STUDY

Rodent models of healthy ageing

Animal models in translational science

Animal models are crucial for bridging various scales of neuroscience. Translational experiments with animal models provide an opportunity to combine molecular or cellular neuroscience methods with techniques used in cognitive and behavioural neuroscience, often in the same subject. Translational studies therefore reduce interindividual variability while facilitating direct comparison at multiple scales. For example, animal models are routinely used to answer questions at a cellular or molecular level, by employing techniques such as calcium imaging, autoradiography, or immunohistochemistry. Such techniques are possible with post-mortem human tissue, but it is generally limited to self-selecting individuals who are willing to donate their bodies to science, and this bias can result in an unrepresentative cohort. Moreover, as volunteers who donate their tissue often expect to die from a particular disease, there is a lack of tissue from healthy individuals. However, as described in the previous chapter, it is fundamentally difficult to decide who is healthy by their cognition, and ‘healthy controls’ can be asymptomatic of the age-related structural and functional changes. Furthermore, harvest and preparation of tissue is a long sensitive process, and the conditions under which it can be completed successfully can be more fully controlled in animals; consequently, it can be completed with less damage to the tissue. Conversely, cognitive and behavioural neuroscience frequently apply techniques such as neuroimaging, optogenetics, electrophysiology, to observe brain structure and function during a given behaviour. While these methods can be applied in humans, they are mostly invasive and are not achieved in practice. The controlled conditions of animal model experiments allow the combination of in-depth exploration of the underlying biology using these invasive methods and methods corresponding to those used in humans.

Animal models have been used in scientific endeavours for millennia, but increasingly so in the last 200 years since Darwin's *On the Origin of Species* speculated on common descent of life from observing the evolution of characteristics within species (Fox & Bennett, 2015). Other famous examples include Louis Pasteur's seminal work on acquired immunity that was born from testing inoculations in chickens and cattle, and Ivan Pavlov's renowned experiments with dogs were intended to understand the process of digestion. Common animal species used in modern science include non-human primates, fish, worms, flies, and most frequently, small rodents. The primary factor for determining an effective animal model is the similarity of the process being observed to that in humans, which fundamentally relies on the principle that the humans are mostly comparable to other animals, whether that is at the genomic or a behavioural level. To this end, non-human primates are commonly used in cognitive science because they share many behavioural characteristics with humans, such as their use of 'language parallels' and gestures (Snowdon, 1990). However, a nematode or fly is more likely to be used in genetics (Alexander et al., 2014; Hales et al., 2015), due to the relative simplicity of their genome compared to that of primates. This distinction leads to two distinct use cases for animal models: that of behaviour or disease. Models of behaviour are used to study particular behaviours, such as decision making, anxiety, and memory while animal models of disease are employed to emulate facets of a given condition such as stroke, neuronal degeneration, or cancer. Of course, they can also be combined to measure the resulting change in behaviour phenotype in a given model of disease.

The choice of animal species is imperative for the successful translation of findings to the clinic. While it is true that no animal model will perfectly replicate the human condition, some animal models will be better suited at recapitulating particular aspects of a condition than others (Sams-Dodd, 2006), which is known as face validity. Replication of a condition can be achieved because a given process occurs naturally in the model organism; these are known as spontaneous experimental models, and they are highly-valued because they provide the best characterisation

(Fox & Bennett, 2015). If the condition of interest occurs naturally, it is likely through the same mechanism as is occurring in humans, which will subsequently increase the reliability of any conclusions drawn from the model. For example, many animals display fear in response to a perceived threat or seek rewards while avoiding discomfort, so testing these responses in an animal is likely to be representative of the behaviours in humans. Similarly, certain cancers develop naturally in rodents, and thus features of the growth can be targeted for therapeutics (Guerin et al., 2020).

However, face validity is typically granted by producing the symptoms of a process via scientific intervention, which is known as using an induced experimental model (Hughes & Lang, 1978). For example, rodents do not develop the hallmark pathologies of Alzheimer's disease spontaneously, and while some non-human primates do develop the biological hallmarks, there is no evidence yet that they develop dementia (Edler et al., 2017; Arnsten et al., 2019). Thus, to obtain an animal model of disease, relevant mutations from the human genome are inserted into the rodent genome (Houdebine, 2007). Most transgenic models are being developed to display the same biological hallmarks as humans, which improves the face validity of the model. However, models of disease can only replicate certain aspects; a lack of understanding of the underlying disease mechanisms hinders the recapitulation of many conditions, but even comprehensive understanding of a given disease does not necessitate that induction of all relevant factors would be possible to give rise to identical symptoms. This disconnect is often denounced for the distinction in success for drugs in preclinical versus clinical trials. Continuing with the previous example, therapeutics targeting amyloid-beta aggregation shows promise in transgenic animal models of Alzheimer's disease, but these therapeutics have, so far, failed to produce meaningful results in late-stage trials (Golde et al., 2011).

Animal models of ageing

The process of physiological and cognitive decline with increasing time is near-universal in mammals, so various spontaneous models are easily accessible. Aged non-human primates and rodents have been known to emulate facets of age-related

structural changes on the cellular level, as similarly seen in humans. For instance, age-related changes in the humans are commonly seen in the prefrontal cortex and the hippocampus (see Age-related changes to brain structure), which also occurs in non-human primates and rodents (Morrison & Baxter, 2012; Gallagher & Nicolle, 1993). Moreover, aged rats have fewer synapses compared to young rats in the molecular layer of the dentate gyrus (Geinisman & Bondareff, 1976) and changes in synaptic density occur in the prefrontal cortex of non-human primates and rodents (Peters et al., 2008; Baxter, 2009), which has been demonstrated in a cohort of healthy elderly humans (Henstridge et al., 2015).

Further evidence of shared ageing processes between humans and animal models of ageing is presented by studies that show age-related decline in outcomes across several cognitive domains (Bizon et al., 2012). These cognitive outcomes are related to the previously discussed hippocampal and prefrontal regions (Gallagher & Nicolle, 1993; Baxter, 2009), which appear particularly susceptible to age-related structural changes. For instance, aged rats performing the commonly used spatial memory tests, the Barnes and Morris Water Mazes, repeatedly make more mistakes and take longer to reach success criteria than young rats (Barnes et al., 1980; Gage et al., 1984; Gallagher & Burwell, 1989). Age-related spatial memory decline has also been demonstrated in rhesus monkeys with both the spatial delayed response task and the spatial element of the delayed recognition span task (Baxter et al., 2001). Moreover, aged rats that demonstrated poor spatial memory had fewer perforated axospinous synapses in the dentate gyrus than older animals that performed well or young animals (Geinisman et al., 1986). This finding is congruous with evidence that manifestation of memory deficits is related to hippocampal synaptic density in humans and across models (Burke & Barnes, 2006), which suggests they share similar underlying mechanisms. Moreover, both rats and humans show age-related impairments when discriminating between perceptually similar objects (Johnson et al., 2017; Yassa et al., 2011), which is a form of episodic memory.

Executive function, which is believed to be rooted in prefrontal cortex circuitry, is demonstrably different between old and young subjects (Bizon et al., 2012; Baxter, 2009). Perseverative errors have been found to increase in aged monkeys, rodents, and humans (Moore et al., 2006; Beas et al., 2013; Head et al., 2009) in an attentional set-shifting task that is mediated by the prefrontal cortex in humans and rodents (Anderson et al., 1991; Birrell & Brown, 2000). Furthermore, older humans make decisions in different ways from younger humans, including placing greater value on future rewards (Samanez-Larkin & Knutson, 2015). Old rats similarly preferred to hold out for large rewards regardless of the delay, unlike young rats who begin ‘discounting’ the size of the reward with increasing delay length (Mikels et al., 2015; Simon et al., 2010). Moreover, inhibition of protein kinase A was found to restore executive function abilities in both aged rats and monkeys, suggesting underlying signalling dysregulation is common to ageing across species (Ramos et al., 2003).

In addition to spontaneous animal models of ageing, ageing researchers can additionally manipulate the rate or quality of ageing through scientific interventions. Transgenic models mostly induce accelerated ageing through manipulation of metabolic pathways (Toescu, 2005; Bilkei-Gorzo, 2014), and wild-type experimental models include the use of dietary restriction (DR) which has long been found to extend life span (Carlson & Hoelzel, 1946). Dietary restriction was found to improve motor and cognitive abilities later in life for Wistar rats, particularly those that began the restriction in early adulthood (Todorovic et al., 2018). It has also been found to diminish the decline of spatial cognition in early-aged Sprague-Dawley rats (Geng et al., 2007). Finally, calorically restricted mice showed ameliorated decline of long term memory and reduced anxiety, while their white matter integrity was preserved and there was a slower decline of their cerebral blood flow compared to mice eating ad libitum (Guo et al., 2015; Parikh et al., 2016). It is hypothesised that these diet-related interventions improve lifespan by augmenting energy and oxygen metabolism by improving the resistance to reactive oxygen metabolites (Weindruch & Sohal, 1997), thus protecting against age-related assaults (Martin et al., 2006).

Enriched environments (EE) have been used as an experimental intervention on the lifestyle for over 75 years (Kempermann, 2019), and it generally involves adding extra toys to the cage beyond standard housing. Enriched environments have been found to have many positive effects, including improve long-term memory and ameliorates stressors (Simpson & Kelly, 2011; Melani et al., 2017; Smail et al., 2020). Enriched environments have stimulatory effects, including increased neurogenesis (Rosenzweig & Bennett, 1996), and an increase in glucose metabolism (Gonzalez-Lima et al., 1994). Importantly, for aged rodents, enriched environments are shown to ameliorate the decline in dentate gyrus neurogenesis that occurs with age, restores new cell survival, and improves synaptic maintenance (Kempermann et al., 2002a; Speisman et al., 2013; Leal-Galicia et al., 2008). With these structural changes, the same aged animals were shown to habituate faster to new environments and performed better in spatial memory tests than those placed in isolation or those in group unenriched housing, either by swimming in shorter paths or by reducing the time taken to find the platform, respectively.

Differences in rodent models

Rats are thought to be the first domesticated animal used in scientific endeavours. Consequently, they have been used more frequently as model animals, particularly because they are easy to handle and are receptive to training, which translates to more reliable performances in behavioural tests than other rodents (Modlinska & Pisula, 2020). However, the prevalence of mouse genome editing has seen the frequency of their use as a model organism rise, though there is increasing scope for genetic variations of rats (Ellenbroek & Youn, 2016). The previous section detailed how rodents, and particularly rats, have been used extensively in ageing research (Gill et al., 1989), which is because their process of ageing with no scientific intervention is highly similar to that observed in humans, but in an accelerated manner. Laboratory mice and rats have approximately the same lifespan length, though there are variations between strains and genders; the median lifespan of rats ranges between 24 and 34 months, while for mice the range is 18 and 32 months (Carter et al., 2020).

This time frame is particularly beneficial for facilitating studies that can incorporate an entire lifespan in a time window that is shorter than the lifespan of the researcher, which is not possible in humans studying other humans.

When it comes to using rodents in neuroimaging studies, imaging mice with MRI requires higher spatial resolution; smaller pixel sizes are necessary to resolve the same level of detail for brain structures of rat brains due to their differences in size. Moreover, physiology rates, such as breathing rate, are faster in mice than in rats, impacting head motion correction, which can be influential in humans but is consequently even more so for rodents due to the lowering of the acceptable threshold for movement (Jonckers et al., 2011). Therefore, while rodents are similar in many ways, if studying wild-type models, rats are a sounder choice as a model organism for use in MRI studies.

The Resilient study protocols

Study aims and design

All animal experiments were conducted in accordance with the Home Office (Animals Scientific Procedures Act, UK, 1986) and were approved by the King's College London ethical review committee.

The Resilient study was a longitudinal study following a cohort of rats across a maximum of four age-related time points: 3 months, 5 months, 11 months, and 17 months old, which approximately correspond to late adolescence, young adulthood, middle age, and the beginning of senescence (Sengupta, 2013). At each time point, the rats underwent a battery of behaviour tests, received a comprehensive MRI scanning protocol, and various biological samples, such as blood and faeces, were procured. The longitudinal design was used to minimise the variability that has been observed in healthy ageing studies of humans, and the use of a rodent model of ageing also resulted in greater control over genetic and environmental confounds than would be expected in a comparable human study.

In addition to the longitudinal study design, there was an experimental manipulation in half of the cohort, which was expected to bring about a model of healthy ageing. Previous pilot data from our group (Cash et al., King's College London, unpublished data) has found that the combination of environmental enrichment and dietary restriction (EEDR) is more effective at stimulating adult neurogenesis in the hippocampus than the additive effect of the two conditions. Therefore, EEDR was hypothesised to engender a group of 'healthy agers' which would be compared to the 'normal agers' that were raised in standard conditions.

The overall aim of the study was to explore age-related changes in brain structure and function, while also identifying any potential differences associated with an experimental model intended to emulate facets of a healthy lifestyle.

Choice of rat breed

A pilot study was conducted to quantitatively determine if there are differences in behaviour between two common breeds of laboratory rat, Sprague-Dawley (SD) and Lister Hooded (LH), to choose the strain for the subsequent longitudinal study. Eight rats from each breed were tested using novel object recognition, which would be used in the study. Novel object recognition is explained in greater detail elsewhere (see Locomotor activity and Novel Object Paradigms), but, in brief, animals are placed in an open-field arena, to which they have been previously habituated, with two identical objects for five minutes. Animals are returned to the same arena after one hour, where one of the two objects are switched for a distinct novel object, and the exploration of the familiar and the novel objects is measured. All statistical analysis utilised Prism 8 (Graphpad, California, USA), and results report the mean $\pm$ standard deviation.

Initially, only automated analysis by Cleversys Topscan Suite (Cleversys Inc., Virginia, USA) was considered to measure the exploration. However, registration problems emerged, apparently caused by the lack of sufficient contrast difference between the black heads of the LH rats and the black open-field arena. Thus, manual tracking, where an experimenter measures each rat's exploration of each object in the arena, was compared to the Cleversys analysis.

The first finding, as seen in figure 2-1A, was that Cleversys software is more constrained than manual tracking; the mean time spent exploring according to Cleversys was 18.483 ± 12.145 seconds, while the manual time calculated for exploring was 50.750 ± 5.127 . Cleversys also shows less total exploration for LH (9.895 ± 6.262 seconds) compared to SD (27.070 ± 14.510 seconds), which was expected given the registration problems.

Figure 2-1: Novel object recognition results between rat strains. Comparison of A) total exploration time (i.e., the summation of time spent with both novel and familiar objects) for all rats across measurement methods; B) total exploration time across breeds; C) times spent exploring either a novel or familiar object across breeds; D) the discrimination index (difference in time spent exploring the novel object compared to the familiar, as a percentage of total exploration time) across breeds. All breed comparisons use manual exploration time calculations. SD rats are denoted by black circles, while pink squares represent LH rats. Lines represent mean and standard deviation for each group.

A two-way repeated-measures ANOVA was implemented to test the effects of breed, scoring method, and the interaction between them on exploration time. Breed alone was not found to have an effect on exploration time ($F(1,14) = 0.4737$, $p = 0.5025$), but the scoring method did ($F(1,14) = 88.12$, $p < 0.0001$). Moreover, there was a significant interaction between scoring method and breed ($F(1,14) = 12.62$, $p = 0.0032$). The interaction was further explored with post-hoc comparisons: the mean difference between manual and Cleversys for the SD was found to be 20.06 seconds (95 % CI: 7.888 to 32.22; $p = 0.0021$). On the other hand, the mean difference between the two methods was found to be 44.48 seconds for the LH rats (95 % CI: 32.31 to 56.65, $p < 0.0001$).

Given the problematic discrepancy between breeds, only the manual method was considered for further comparison of the two breeds' object exploration. No animals were excluded due to lack of exploration (total exploration less than 10 seconds; see figure 2-1B) or complete fixation on only one of the two objects.

The time spent exploring the two objects was not significantly different between the two breeds, as determined by an unpaired two-tailed t-test (7.250 ± 9.819 seconds, 95 % CI: -13.81 to 28.31 seconds, $p = 0.4725$). On average, the SD rats explored both objects for 47.13 ± 8.474 seconds (95 % CI: 27.09 – 67.16), while the LH rats explored the objects for 54.38 ± 14.03 seconds (95 % CI: 42.65 – 66.10), as seen in figure 2-1B.

The hypothesis that the rats showed a preference for exploring the novel object compared to the familiar object was tested by breaking down exploration time by object familiarity (figure 2-1C). The novel object was explored for 28.625 ± 17.212 seconds for the SD group, and 33.875 ± 9.963 seconds for the LH. In comparison, the SD group explored the familiar object for 18.5 ± 10.461 seconds, while the LH explored the familiar object for 20.500 ± 9.739 seconds.

A two-way repeated-measures ANOVA tested the effects of breed, object familiarity, and the interaction between the two terms on exploration time. The model estimated

that the novel object was explored for 11.75 ± 3.659 seconds longer than the familiar object, and there is a significant effect of object familiarity on exploration time ($F(1,14) = 10.31, p = 0.0063$). There was no significant effect of breed on exploration time ($F(1,14) = 0.5452, p = 0.4725$), and there was no significant interaction between the two terms ($F(1,14) = 0.1972, p = 0.6637$). Post-hoc comparison of object familiarity shows that both SD (mean difference = 10.13 ± 5.174 seconds) and LH (mean difference = 13.38 ± 5.174 seconds) spend more time exploring the novel object. However, while the difference between the two times surpassed significance for LH rats (95 % CI: 0.4236 to 25.33 seconds, $p = 0.0427$), the time SD rats spent at the novel object did not significantly differ from the familiar (95 % CI: - 2.826 to 23.08 seconds, $p = 0.13633$).

Accounting for individual differences in overall exploration time, within and between studies, is achieved by utilising the prevalent discrimination index (as described by Antunes & Biala, 2012), which is calculated as:

$$\frac{time_{novel} - time_{familiar}}{time_{novel} + time_{familiar}} \times 100$$

A positive discrimination index indicates a preference for the novel object over the familiar, while a negative index indicates the reverse. A discrimination index of 0 indicates the rats spent equal time at both objects, which would be expected by chance.

The discrimination index was calculated for all rats using the manually derived exploration times (figure 2-1D). While the SD rats had a lower discrimination index on average (18.48 ± 8.399 %) compared to the LH (25.41 ± 9.345 %), an unpaired two-tailed t-test found the difference in mean discrimination index between the two breeds was not statistically different (6.931 ± 12.56 %, 95 % CI: -20.02 to 33.88 %, $p = 0.5899$).

In summary, although it appears the LH rats explore the objects slightly more than SD, there was no statistical significance between the two rat breeds. Cleversys is the

preferred measurement method because it is more constrained in its estimation, less time-expensive, not prone to human error, and, therefore, increases reproducibility of the results when compared to manual tracking. Therefore, as the preferred method did not work successfully in the LH rats, and there is ostensibly no statistically significant behavioural difference, the SD rats were chosen for the study.

Animals

Male Sprague Dawley rats ($N = 48$) were received from Charles Rivers, UK at 1-month-old ± 4 days. All rats were housed in cages of four until the largest rat in the cage was above 600 grams when they were then housed in pairs, due to legal requirements for minimum cage space per rat. As lone-housing is a known stressor in rats (Nyuyki et al., 2012), and loneliness is known to have negative impacts on healthy ageing in humans (Cacioppo et al., 2014), lone housing was considered a confound. Any animal that was housed alone and could not be rehomed with another rat was withdrawn from the study. Cages were kept in a temperature (21 ± 2 °C) and humidity (55 ± 15 %) controlled room, on a 12:12 light: dark cycle. Each cage had ad libitum access to water at all times and was supplied with sawdust bedding, nesting material, a cardboard tube, and wooden chewsticks as standard.

Environmental Enrichment Dietary Restriction model

Elements of a healthy lifestyle were modelled in half of the rats ($N = 24$) by providing environmental enrichment and dietary restriction (EEDR) following the first scan time point at 3 months old. The control rats continued to be housed in standard cages and given ad libitum access to standard chow. Environmental enrichment comprised of four changing sets of toys, a list of which can be found in appendix 2.a, that were replaced on a weekly rota which coincided with the weekly weigh-in. The control rats were weighed and had their cages changed at the same time but did not receive environmental enrichment beyond the standard caging. Dietary restriction took the form of intermittent fasting, where EEDR rats were fasted three non-consecutive days a week for 24 hours before giving access to ad libitum food

for 24 hours. Dietary restriction was first implemented immediately after the first scanning time point, at the same time as environmental enrichment. While environmental enrichment lasted the length of the study, the dietary restriction was discontinued six weeks before the final behaviour battery, which was approximately two months before the end of the study. The decision to end dietary restriction was driven by the attentional set-shifting paradigm, which required a reduction in caloric intake prior to training and testing, and would therefore not have been possible to combine with intermittent fasting.

Behaviour testing

Figure 2-2: Diagram of longitudinal design

All behaviour tests took place approximately one week before the rats underwent MRI scanning for the time point. Cognition can be severely impacted following the administration of general anaesthesia (Brioni et al., 2017), so behaviour tests were always carried out before MRI to minimise the impact on cognition.

Locomotor activity and Novel Object Paradigms

An open field test was used to measure spontaneous locomotor activity and anxiety (Seibenhener & Wooten, 2015). This test has been widely used in some form for almost a century (Hall & Ballachey, 1932), and leverages the aversion that rodents show of open areas. Therefore, the time spent in the centre of the field is considered a measure of anxiety. As anxiety is overcome, the area explored, and distance travelled, increase. This test is paired with the novel object paradigm to test the hippocampal-dependent recognition memory (Broadbent et al., 2010; Cohen & Stackman Jr., 2015).

All tests took place in the same room, under the same red-light conditions. At most, four rats were tested at a time, which was recorded using an overhead camera to measure the time spent exploring both objects. The arenas were built inhouse from black Perspex, and each of the four individual arenas measured 60 x 60 cm.

Before the first session, rats were habituated to the open field arenas in 30-minute sessions over four consecutive days. However, for behaviour testing of subsequent

time points, only one day of habituation was given. The final or only day of habituation (for first or subsequent time points, respectively), also served as a measure of locomotor activity by quantifying the movement of the rats in the arena over the 30-minute habituation period (figure 2-3a). In the novel object recognition paradigm, the rats were presented with two identical 3D printed shapes, secured in the arena with Velcro. In the familiarisation stage, rats are given five minutes to freely explore the arena and the objects, before being returned to the home cage for one hour. The rats were returned to the arena for a further five minutes in the test stage, where one of the objects was replaced with a novel object (figure 2-3b).

Figure 2-3: Representations of the open field test. A) In a locomotor activity assessment, the time spent in and the distance travelled is measured for each zone, which become increasingly more exposed. B) novel object recognition diagram. In the familiarisation phase, the rat is placed in the open field with two identical objects for five minutes. This is followed by the test phase one hour later when the rat is placed in the arena with the previous object and a new object for five minutes

All open field videos were analysed using Cleversys Topscan Suite (CleverSys Inc., Virginia, USA). Open field analysis measured several metrics regarding zones outlined for three zones in each arena: the outer arena, which was delineated by the boundary of the arena, the inner arena which consisted of the inner 20 % of the outer zone, and the centre which consisted of the inner 20 % of the inner zone. Metrics measured for each arena included the number of entries, duration in zone, latency to enter, distance travelled, and velocity of travelling in each zone. Novel object analysis measured the movement in three zones: the arena, and the outline of both objects

dilated by six pixels. Metrics measured for novel object tests included locomotion within the arena as in the open field test, and bouts sniffing each object, duration sniffing each object, latency to sniff each object.

Attentional Set-Shifting

Cognitive flexibility is a prefrontal dependent executive function that is often tested with the Wisconsin Card Sorting Test (Berg, 1948), where a subject is required to determine a rule for distinguishing between sets and apply it correctly in changing contexts. In humans, this rule is determined from cards with different colours and symbols, where only one dimension (e.g., symbol colour) is decisive to receive the reward. An intra-dimensional shift is when new colours and shapes are used, but the colour is still the deciding dimension for the reward. Conversely, an extra-dimensional shift occurs when the shapes are now used to determine the provision of rewards, and the colour is now irrelevant. This test has been adapted from humans for rodents using a protocol that replaces cards with shapes and colours with pots with scents and media (Birrell & Brown, 2000).

Before testing, animals were given ad libitum access to water and maintained on a restricted diet for 14 days before testing, with each rat being given 17 grams of standard food pellets per day. Food restriction before testing was necessary to engender the motivation to find the food reward given in the task, and the mass of food was calculated to maintain bodyweight at 80-85 % of their free-feeding counterparts. Animals were fasted overnight before habituation and testing.

Figure 2-4: Attentional set-shifting diagram. The rat is taught to dig in two pots placed in either decision chamber. In the test phase, only one pot is baited with a food reward, and the rat should use the distinct digging media and odours to distinguish between the baited and non-baited pots.

Test	Dimension		Combinations	
	Salient	Irrelevant	Rewarded	Unrewarded
Simple	Odour	Digging Media	Tea Tree	Frankincense
Compound			Tea Tree / Pebbles	Frankincense / Pebbles
			Tea Tree / Wood Chips	Frankincense / Wood Chips
Reversal 1			Frankincense / Pebbles	Tea Tree / Pebbles
			Frankincense / Wood Chips	Tea Tree / Wood Chips
Intra-dimensional			Peppermint / Tetrabed	Cinnamon / Tetrabed
			Peppermint / Wood Shavings	Cinnamon / Wood Shavings
Reversal 2			Cinnamon / Tetrabed	Peppermint / Tetrabed
	Cinnamon / Wood Shavings	Peppermint / Wood Shavings		
Extra-dimensional	Digging Media	Odour	White Paper / Eucalyptus	Brown Paper / Eucalyptus
Reversal 3			White Paper / Rosemary	Brown Paper / Rosemary
			Brown Paper / Eucalyptus	White Paper / Eucalyptus
			Brown Paper / Rosemary	White Paper / Rosemary

Table 2-1: Example combinations for attentional set-shifting trials. Bold text denotes the salient dimension choice.

The testing procedure and equipment are based on those described by Birrell & Brown (2000). A custom black opaque Plexiglas arena (40 x 60 x 30 cm) was made in-house. The arena was split into thirds, where the choice area was a third of the length of the arena, and the start area was two-thirds of the length. The choice area was split by a Plexiglas divider, giving two choice sections measuring 20 x 20 cm. Digging bowls, made from plastic plant pots are placed in cut-outs in these sections, so only a small part of the top 2 cm of the plant pot is showing. Each choice area is separated from the start arena by a removable Plexiglas divider.

Training and testing occurred in the same room under the same low-light conditions. On the training day, rats are habituated to the arena and trained to dig in the submerged pots filled with sawdust to retrieve a reward (Honey Cheerios, Nestle, UK). Once the rat is digging reliably, it is administered a simple discrimination test. The subject must learn to discriminate between two distinct elements of the same medium (i.e., two different scents or digging media), only one of which has the food reward. Media types and the salient element of the media were counterbalanced across groups, while left and right choice area was counterbalanced across trials. Rats are permitted four discovery trials, after which they are required to achieve six consecutive correct trials in a row. All trials begin with the rat in the start area, as the dividers to both choice areas are simultaneously removed. A discovery trial ends after ten minutes, or once a rat has tried digging in both pots. A test trial ends, either with the choice of an area or after ten minutes. A correct trial is when the rat digs in the correct choice area, whereas an incorrect trial is when the rat digs in the incorrect choice area or does not dig at all. Metrics recorded for this test include trials to criterion, length of test, and the number of mistakes made prior to test completion.

Testing took place the day following habituation, and as with the habituation protocol, trials begin with access to the choice arenas, where only one of the bowls contains a food reward. At the start of each test, the first four trials are discovery trials, and the test trials are terminated once a rat digs in a bowl or after ten minutes with no digging. Trials continue until the criterion of six consecutive trials is reached

or until 25 trials have been run. The order of testing sessions is consistent between rats, and an example is available in table 2-1. Rats begin with a simple discrimination with only one dimension (odour or media, depending on the habituation protocol). The rat then moves on to compound discrimination, where a second dimension is introduced, but the salient dimension remains the same as the simple discrimination. A reversal trial is run where the salient dimension remains the same, but the correct choice becomes incorrect, and the reward is located in the previously incorrect choice. An intra-dimensional shift uses new media and odours, but maintain the correct dimension as previously tested in the compound discrimination. An extra-dimensional shift also uses new media and odours, but the previously correct dimension is now incorrect, and the previously incorrect dimension now contains the reward. Analogous to the habituation trials, the number of trials to criterion, number of incorrect trials, and the length of the test were measured.

The test was also completed with an independent cohort to assess the effect of age on this behaviour. Ten male Sprague Dawley rats that were approximately 3 months old at the time of testing were observed to this end.

Clinical Frailty Index

Frailty is ‘a condition of greater risk of adverse health outcomes’ that increases in likelihood with increasing age (Collard et al., 2012). Healthy ageing is associated with minimal frailty, so quantifying the degree of frailty is essential for describing heterogeneity in health outcomes in ageing studies. There are multiple metrics for quantifying frailty in humans (de Vries et al., 2011), and so attempts have been made to translate these metrics to rodents (Miller et al., 2017; Yorke et al., 2017). A truncated version of the criteria put forward by Yorke et al. (2017) was used for this study, which is included in appendix 2.2. The battery was completed before the 11-month and 17-month scanning time points. Rats were placed in a clean cage to carry out the assessment, in the same room as their home cage.

While many of the items involve physical assessment following the same protocol as that detailed by the original paper, a notable exception was the hearing loss test. The

startle reflex was assessed by clicking an unloaded stapler approximately 15 cm from the subject three times with 30 seconds between each click. The stapler was kept out of sight of the rat to minimise inducing a startle reflex via visual stimuli rather than auditory. A point was given to rats that startled at all three clicks, a half-point was given to rats that startled to at least one click, and no points were given to rats that did not exhibit a startle reflex to any of the clicks.

Grip Strength test

Grip strength has been proposed as a biomarker for frailty in older humans because it is more predictive of mortality than chronological age alone (Syddall et al., 2003), which suggests that it is indicative of underlying ageing health. Grip strength was measured using a Grip Strength Meter (TSE Systems GmbH, Bad Homburg, Germany) twice: at the first and last time point, after the behaviour battery but before scanning. Rats were supported around their waist and hind limbs, with their forelimbs resting on the grip bar. They were gently pulled horizontally until the rat released its grip on the bar, at which point the maximum exerted force is displayed. This protocol was repeated three times for each rat.

Scanning protocols

Physiological considerations

Scanning rodents requires specific adjustments with regards to protocols developed with humans. The most pertinent adjustments relate to getting the animal in the scanner. Although it is possible to scan animals awake, it is particularly problematic. It requires large amounts of time-consuming training to habituate the animal to stay still in the bed for long periods amongst the loud distractions of the scanner (Chang et al., 2016). Additionally, the noises are likely to be stressful, which aside from being a potential confound in fMRI experiments that are already sensitive to small changes in physiology, it is unethical to subject animals to unnecessary levels of stress.

Consequently, it is much more common to administer general anaesthetics and sedatives to animals, but this has the unintended effect of having to be much more cognisant of animal physiology than in humans (Vutskits & Xie, 2016). Firstly, as the Resilient study is a longitudinal study, the choice of anaesthetic must be minimally invasive and facilitate easy recovery. The choice of anaesthetic is known to impact the accuracy of the resting state signal. For instance, Paasonen et al. (2018) found that propofol, urethane, and α -chloralose to be closest to awake scanning, but these are also very difficult to incorporate into a longitudinal design. While isoflurane exposure can cause inflammation and cellular apoptosis (Vutskits & Xie, 2016), and the central nervous system is particularly vulnerable to exposure in late-life (Strøm et al., 2014), it is commonly used as an inhalational anaesthetic, including successful replication of BOLD-derived resting-state networks (Westphal et al., 2017). Furthermore, the duration and depth of anaesthesia are reproducible, and it is safer than other options for repeated exposure (Albrecht et al., 2014). Therefore, isoflurane was the anaesthetic of choice given the longitudinal design, to augment the rats' recovery, but the volume delivered was minimised, especially during the functional scanning protocol.

The importance of physiology relies on stable animal physiology to differing degrees in structural and functional MRI, as the faster temporal resolution of the latter is primed for measuring physiological changes. As already demonstrated, physiology can be greatly impacted by choice of anaesthetic, but it can also be impacted due to the volatility of the vascular system and the potential for neurovascular and metabolic uncoupling (Pan et al., 2015). Therefore, care was taken to monitor breathing rate, maximise arterial blood oxygenation readings, and stabilise internal temperature between 36 and 38 °C throughout scanning, but particularly during the functional paradigms. Functional MRI can also be highly corrupted by motion, which is one of the principal reasons for anaesthetising rats. Therefore, rats were secured in the nose cone with bite bar and ear bars, which immobilises the head and minimises breathing-related motion artefacts.

Scanning preparation

Rats were weighed and then placed in a ventilated chamber heated to 40 °C for 20 minutes to promote vasodilation and facilitate intravenous blood sampling, before being transferred to an anaesthetic induction chamber. Anaesthesia was induced with 5 % isoflurane in a 30:70 mixture of Oxygen in medical air, flowing at approximately 1 litre per minute. Induction lasted up to 3 minutes before isoflurane was reduced to 2.5 % for sampling and transferring the rat into the scanner. Approximately 0.3 millilitres of blood from the tail vein and faecal samples were collected during this initial anaesthetic process, for additional microbiome and blood plasma testing that is beyond the scope of this work. The animal was transferred to the scan bed, which has a built-in heating system using hot water that was supplemented with a tube supplying thermostatically controlled hot air using the reading from a rectal thermometer set at 37 °C. The thermometer and other physiological monitoring equipment, including pulse-oximetry placed on the hind paw and a respiration balloon under the chest to measure respiration rate, was made by Small Animal Instruments, Inc., NY, USA. Physiology was monitored and recorded throughout using PC-SAM (Small Animal Instruments, Inc, NY, USA)

software, and anaesthesia was adjusted to keep physiology as stable as possible. The bed was fitted with a four-channel array receiver coil that was placed on the superior head surface, and a transmit/receiver ASL coil, which sat inferior to the rat's neck to transmit labelling pulses for cerebral blood flow measurements. The animal was then loaded, head prone, into the 86 mm volume transmission coil in the 9.4 T Bruker Biospec MR scanner.

Scan protocol

Figure 2-5: Diagrammatic representation of the scanning protocol

Scanning protocols were saved in and implemented using Paravision 6.0.1 (Bruker Corp., Ettlingen, Germany). Overall, the scan outline is available in figure 2-5, with raw images available in figure 2-6. Scanning took approximately two hours from starting the first scan to the end of the last.

Impedance matching of the receiver coils was carried out once a day before scanning proceeded for the first rat. The rat's position in the scanner was verified and adjusted using localiser scans for both the array and arterial spin labelling (ASL) coils before a field shim was generated and a B₀ map was acquired. Isoflurane was reduced to 1.8 % during the field shim and remained at that level for the length of the functional scanning provided the animal was physiologically stable.

First, two spin-echo EPI scans were run in order to calculate the distortion field for blip-up blip-down correction (e.g., as implemented with FSL's top up) which was immediately followed by a multi-echo resting-state scan. The position of the carotids was evaluated in case of movement during the resting-state scan, before calculating the necessary power required for tagging of the carotid for continuous ASL (CASL). The reference power derived from the CASL coil power was multiplied by 0.725 and used as the pulse power for the CASL scan. The tagging slice was calculated as 5 mm

before the most caudal split of the carotid arteries, as viewed using the ASL coil localiser scan. After the CASL scan has run, a duplicate CASL scan with pulse power 0 is run after T_1 relaxation, and thus approximately 10 seconds after the previous scan finishes. This scan acts as a T_1 -weighted reference scan for quantifying Cerebral Blood Flow. A phase reversed copy of the reference scan is also run for distortion correction.

The isoflurane was raised after the completion of the CASL scans and remained at 2-2.5 % for the anatomical scanning protocol. A 3D ultra-short echo was acquired for combining with the MP2RAGE structural scan. The structural scans were followed by the acquisition of a 3D field map, diffusion-weighted EPI scan, and a reversed-phase copy of the diffusion scan, before concluding with T_2 - and T_2^* -weighted images.

Following scanning, rats were administered a subcutaneous saline injection to combat dehydration and placed in a warm ventilated chamber until recovered from anaesthesia, when they were placed back into their home cages.

Figure 2-6: Example raw scans from protocol

Generation of analysis resources

Despite the need for preclinical imaging to expedite laboratory research to potential therapeutics, the preclinical imaging community is not served by ample foundational resources and software instruments compared to human imaging. For example, the neuroimaging software used to process human MRI data relies heavily on prior information. Priors such as average templates and tissue probability maps are species-specific, and therefore almost always inappropriate for preclinical data. Thus, researchers make ad-hoc adaptations of existing software tools so that they work on their data. These local, in-house customisations move the field further from its human counterpart in transparency and, more worryingly, the ability to generalise across studies, samples, or modalities.

The software and resources detailed in the following section will be openly released, equipping researchers with adequate tools and reliable prior knowledge to integrate into their workflows, thus expanding the availability of fundamental resources required by the field.

Age-specific population templates and probability maps

As the brain's morphology is known to change over the lifespan (see Measuring correlates of ageing with MRI), population-specific averages have been implemented in human ageing studies to reduce inter-subject variability particularly in older populations (Huang et al., 2010; Fillmore et al., 2015). Age-specific MRI resources are available for developmental rodent studies until postnatal day 80 (Rumple et al., 2013; Calabrese et al., 2013), which is understandable as the brain changes rapidly during this time, but to the best of our knowledge, there are no resources available for rodent imaging that reflect age-specific brain changes occurring in adulthood and old age. Brain size and the relative volumes of different tissues were hypothesised to change over time, so these resources were created as part of the Resilient study.

Figure 2-7: Example resources generated for the Resilient study. Column A) template; column B) tissue label image; dark grey represents grey matter, light grey represents white matter, and white represents CSF; column C) final tissue probability maps for one subject.

A previously delineated rat atlas (Papp et al., 2014) was categorised into tissue type, i.e., grey matter, white matter, and CSF. Regions with non-homogenous tissue compositions, which were most susceptible to partial volume effects, were excluded from the list of regions, leaving only regions that were highly representative of their tissue class. A label image (as seen in figure 2-7B) was generated from this list, where each voxel was assigned an integer value that describes the tissue composition. Put simply, all voxels that were highly likely to contain grey matter were given the value 1, voxels likely to contain white matter were valued 2, and CSF voxels were given the value 3.

At each time point, all available 3D T_1 -weighted (T_{1w}) images were combined across groups, to create four age-specific, average *in vivo* MRI templates using the `antsMultivariateTemplate2.sh` script distributed with ANTs. All templates, an example of which is available in figure 2-7A, were spatially normalised to the high-resolution reference image associated with the rat atlas, enabling the label image to be projected to the time point space, by applying the inverse transformation.

The excellent tissue contrast provided by T₁w images was leveraged to segment all voxels in the brain into three compartments with the tissue label image using ANTs Atropos tool. As the label image only included homogenous regions, the label classification was propagated throughout a template space brain mask (see Atlas-based brain extraction), producing a voxelwise probability for each corresponding tissue class. Tissue probability maps were created for each subject by transforming their T₁w image from the native space to the time point space and using the time point probability maps for segmentation. All subject segmentations in the time point template space were averaged to create the final probability maps. The volume of each tissue class, examples of which are available in figure 2-7C, was calculated as the product of voxel size and the sum of values in each Jacobian-modulated tissue compartment image, and the sum of the three values was taken as a measure of total intracranial volume (TIV).

Linear mixed-effects models were employed to demonstrate the requirement of age-specific resources by testing the hypothesis that brain volumes change with age in the adult rat. Thus, separate models were fit to predict each brain volume measure (TIV, grey matter, white matter, and CSF) from age. To test the hypothesis that tissue classes have distinct trajectories, and thus represent different proportions of brain volume with age, the volumes were additionally normalised to a ratio of TIV, and a linear mixed-effects model was fit to predict the ratio of brain volume from tissue class and age. Post-hoc analysis of the volume models was carried out by calculating the estimated marginal means, and the estimated marginal trends were calculated per tissue class from the ratio model. Pairwise comparisons of the estimated marginal means across age and the gradient for each tissue type's estimated marginal trend are reported.

The results of the volume models suggest that TIV, grey matter, and white matter increase with age, while CSF does not change. Post-hoc analyses show TIV significantly increases at every time point, with the smallest increase between 5 and 11 months ($47 \pm 8.21 \text{ mm}^3$, $p < 0.0001$). Grey matter volume significantly increases

at each time point, with the smallest increase also between 5 months and 11 months ($41.5 \pm 12.1 \text{ mm}^3$, $p = 0.0047$). In contrast, white matter volume does not significantly differ between 5 months and 11 months ($3.14 \pm 6.53 \text{ mm}^3$, $p = 0.96$), though all other comparisons are significantly different, with the smallest increase between 3 and 5 months ($21.99 \pm 5.69 \text{ mm}^3$, $p = 0.0012$). Furthermore, both age and tissue class are significant predictors in the ratio model, as is the interaction between the two predictors, indicating that the slopes for each tissue class (grey matter: $0.36 \pm 0.03 \%$, white matter: $0.18 \pm 0.03 \%$, CSF: $-0.02 \pm 0.03 \%$) are significantly different, with the smallest difference between grey and white matter ($0.18 \pm 0.04 \%$, $p = 0.0002$). Thus, rodent brain volumes change over time, and tissue classes have distinct trajectories; thus, preclinical imaging studies may benefit from age-specific resources.

Atlas-based brain extraction

Masking non-brain tissue in an image is known as brain extraction, and it is an essential preprocessing step. Brain extraction improves registration and reduces the influence of non-brain signal, which introduces variability between subjects. Although human MRI data has many reliable tools for brain extraction, these rely on prior information, such as shape models or probability maps, that is inappropriate for use with rodent data. Researchers can manually delineate the brain mask for each image, but the results can vary depending on the researcher carrying out the delineation and is a time-consuming exercise which is routinely being discounted with increasing MRI study sizes.

Automated tools increase reproducibility and reduce human error, but the existing neuroimaging software almost always expects human brains as default, which is problematic for rodent images for several reasons. First, there are large differences in the shape of human and rodent heads, and therefore their brains. Consequently, methods based on quasi-spherical shapes, such as FSL's BET (Smith, 2002) are inappropriate for use with rodents. Atlas-based methods, such as `antsBrainExtraction.sh` which is distributed with ANTs (Avants et al., 2008), have been found to outperform

other methods applied in human imaging (Glasser et al., 2013; Esteban et al., 2019) but they rely on human population averages, which do not generalise to rodents for several reasons. In rodents, the anterior-posterior plane is elongated relative to the left-right and superior-inferior planes, which makes the brain more ellipsoid than the human sphere. There is more non-brain tissue surrounding the brain in rodents, which can divert signal-seeking algorithms such as clustering and region growing methods like `mia-3dbrainextractT1` (Wollny et al., 2013). Additionally, rodents have a smaller gap between the brain and the skull, and the field-of-view is generally very narrowly defined around the brain with less background area surrounding the head, which precludes watershed strategies (Ségonne et al., 2004). Animals are often placed with their anterior-posterior plane along the direction of the magnetic gradient, whereas human scanners usually accommodate their subjects with the superior-inferior plane along the direction of the bore. Moreover, while humans are almost always supine, animals are more likely to be prone, although can be supine depending on the idiosyncrasies of the study. This difference can confuse orientation to the standard Right Anterior Superior (RAS) alignment, due to the inconsistencies of whether RAS relates to the scanning subject or the scanner. Second, the standard structural image in human imaging is T_1 -weighted, but in preclinical imaging, it is more common to acquire T_2 -weighted structural images. This difference in contrast-preference generally hinders efforts to extract the brain based on histogram thresholding methods. Finally, preclinical imaging usually relies on stronger magnetic fields to increase image resolution required to image smaller brains. The range of common magnetic field strengths is larger for rodent scanning, which ranges between 4.7 T to 15 T, than with human scans, which are usually 1.5 T or 3 T, while 7 T is less common. With stronger magnetic fields, the effects of field inhomogeneities and artefacts on images are also stronger.

The preclinical imaging community have designed several brain extraction tools to address these issues (Chou et al., 2011; Wood et al., 2013; Oguz et al., 2014; Hsu et al., 2020; Liu et al., 2020), which depend on the image contrast to outline the shape

of the brain. Therefore, the success is constrained to images acquired with standard parameters and with uniform signal and they can perform poorly across datasets, which are often idiosyncratic to the MR system used and the purpose of the study, because these factors determine the image resolution, contrasts, and field-of-view (FOV).

However, the shape of the brain is mostly homogeneous between rodents. Therefore, atlas-based brain extraction should be an ideal approach to apply, but it has been challenging to implement in rodent pipelines, because of the paradoxical requirement of aligning an image to a template to get a mask, while also requiring a mask to achieve the best alignment results. It was successfully applied in six mice as part of the quantification of atlasBEX (Lohmeier et al., 2019), but this implementation required a group template to be manually delineated and had high computational demand.

A brain masking tool, *artsBrainExtraction* (where *arts* stands for *ANTs-based Rodents Tools*), was developed to adapt the *antsBrainExtraction.sh* workflow specifically to the idiosyncrasies of rodent imaging, including the template, resolution, and variability of image contrast. This approach to brain extraction is agnostic to image contrast and leverages the similarity between individuals by augmenting the initial alignment of a target image without a mask to a template. Furthermore, it is made available as part of the *niRodents* Python package (www.github.com/nipreps/nirodents) of neuroimaging processing tools for rodents, so it is available across operating systems and without proprietary software.

Atlas-based brain extraction fundamentally relies on a precise spatial mapping between the target specimen and a reference population-average (or template), in which a precise delineation (binary or probabilistic) of the brain is defined. The estimation of the spatial mapping between individual and template is typically resolved with a non-linear image registration framework. Once the mapping is estimated, the brain mask (or probability map) defined on the template space can then be projected into the individual's image space, effectively masking out the non-

brain tissue. Atlas-based methods are less reliant on prior information, and thus typically more robust, than their counterparts. For instance, methods that define the mask using image contrast may exclude regions with typically low signal, such as the ventral surface of the brain (Liu et al., 2020). Moreover, existing methods (for both humans and rodents) generally require specific preprocessing steps and parameter adjustment (Oguz et al., 2014; Liu et al., 2020), while atlas-based brain extraction requires minimal user input.

The `antsBrainExtraction.sh` workflow is a mature implementation of an atlas-based approach, which uses an MRI template derived from the scans of 20 adult humans. The workflow first manipulates the images to facilitate the convergence of the registration process: the image histogram is truncated to discard signal intensity outliers, and the images are corrected for non-uniform intensity (Tustison et al., 2010). Both images are Laplacian filtered, which demarcates rapid changes in signal, such as at the brain boundary. The transformation is estimated with a three-stage multi-modal registration is implemented with rigid-body, affine, and non-linear registrations between both the bias-corrected image and the Laplacian images. Finally, the template mask is projected into the individual’s space.

Figure 2-8: Illustration of the default `antsBrainExtraction` algorithm

The proposed default `antsBrainExtraction` workflow recapitulates the ANTs approach with critical changes implemented for non-human imaging, as presented in figure 2-8. The

most fundamental change is the reference template; `TemplateFlow` (www.templateflow.org) is integrated into the workflow, which equips the workflow with programmatic access to a bank of population templates to choose from, such as the Fischer344 template (Goerzen et al., 2020), which `artsBrainExtraction` uses by default. Other adaptations to the ANTs' implementation include building the workflow with `Nipype` (Gorgolewski et al., 2011), and denoising the moving image using a non-local means filter (Manjón et al., 2010) before it is corrected for signal inhomogeneity (Tustison et al., 2010).

Further improvements include the compatibility of parameter values for rodent images (e.g., the distance between B-Spline control points in the inhomogeneity correction step) and the implementation of select ANTs tools to ensure consistent orientation information is stored in the headers of intermediate and final images. For particularly challenging head positions, the original human workflow includes a step to initialise the nonlinear registration more robustly. Because FOV prescription and head tilting are generally more consistent between rodent images, due to fixation frames and specialised beds, this extra step is disabled by default in `artsBrainExtraction`.

It is important to note that some changes were made to the default algorithm for application to the Resilient dataset, because the original algorithm was developed for T_2w images, and thus could be enriched for the target images in two ways. First, the Waxholm Space template (Papp et al., 2014), which is derived from a Sprague Dawley rat, was chosen as the target image to match previous preprocessing endeavours. Secondly, the workflow was truncated because the target images were T_1w images derived from an MP2RAGE protocol. As these images were partially processed at the time of masking, they did not require the denoising or bias field correction steps that are necessary for raw images and would require a mask much earlier in the processing pipeline.

3. VARIANCE AND COVARIANCE OF BRAIN STRUCTURE IN HEALTHY AGEING

Rationale

This chapter describes the analysis of the structural imaging data collected during the Resilient study. The work aims to measure the structural variance and covariance patterns between control rats and those that received EEDR, across multiple structural imaging modalities, which are analogous to those collected in human ageing studies, at all time points. Structural metrics are more straightforward in their interpretation, at least compared to functional imaging because structural metrics are less dependent on fluctuations in physiology. Physiology can be more volatile in preclinical studies because of factors that are required for preclinical imaging but less common in human studies, such as anaesthesia (Becq et al., 2020; Grandjean et al., 2014). Although there is a lack of evidence to conclude the directionality of the relationship between age, brain structure, and brain function changes observed in ageing, in the design of this particular experiment but also the field more generally, coordinated brain function is inherently underpinned by structural architecture, so it is intuitive to examine the potential differences in brain structure in the first instance.

As discussed previously, the ageing brain demonstrates ostensibly non-linear individual trajectories of structural changes. As ageing is a continuous trait rather than a series of categorical outcomes, it follows that the effects of ageing can be observed as small but widely distributed differences across the whole brain, rather than in particular focal regions. For example, while there are varying degrees of brain injury, an infarct can be categorically present or absent. Ageing, in comparison, is a dynamic, non-linear trait which should be assessed appropriately. Moreover, as it is unlikely that the degree of ageing is focalised to a single region, neither is it likely that a single modality is best placed to capture the multitudes of subtle influences. Multi-modal imaging probes different facets of brain structure in the same subject which are likely related but not equivalent (Hao et al., 2013). The design of the Resilient study included multiple different structural imaging protocols (including

MP2RAGE, multi-gradient echo, multi-spin echo, and diffusion tensor spin-echo echo-planar imaging (EPI) images) to capture a variety of quantitative metrics in the ageing rat brain that encompass tissue volume, tissue concentration, relaxometry, and diffusion characteristics. Example images of these modalities are available in figure 3-1. To this end, it was hypothesised that utilising multiple complementary information sources could better inform future hypotheses about the differences in brain structure during ageing than any alone.

Figure 3-1: Example derived structural images

Due to its short TR and TE, T_1 -weighted (T_{1w}) images provide good contrast between grey and white matter with short acquisition times, so have traditionally been used to derive high-resolution brain volumetric measurements (Good et al., 2001). The voxel values represent the signal and so are relative; therefore, values from T_{1w} images can only be compared when modulated to infer volumetric information. Estimates of tissue volume and concentrations are canonically derived from T_{1w} images using a process known as voxel-based morphometry (VBM; Ashburner & Friston, 2001), which requires the image to be divided into anatomically meaningful segments. Segmentation exploits features of the image, such as the visible contrast differences between tissues (which facilitates manual delineation of the segments of interest), signal value of the voxel (which are known as histogram-based

methods), or the shape of the image (in conjunction with an *a priori* atlas of brain regions or tissues). The latter case is increasingly common, due to the abundance of automated tools that can expedite image processing, which is valuable when dealing with large datasets (Feo & Giove, 2019). However, for shape-based automated segmentation to be successful, the atlas and image must share a common space, which likely involves applying a function to spatially transform one image matrix to the space of the other. In this process, the Jacobian determinant, which describes the linear transformation from an image's native space to another, provides scaling information that allows volumes and concentrations to be calculated in VBM. Moreover, Jacobian determinants derived from the transformation of the T_1w images to the common space are commonly used as inputs for volume and tensor-based morphometry (TBM) which measures the magnitude and direction of enlargement or shrinking required for spatial normalisation relative to a common-space (Ashburner & Friston, 2003; Chung et al., 2001).

In comparison to the relative values in T_1w images, relaxometry maps and diffusion scalars provide quantitative values which are a direct measurement of the micro-environment of a particular voxel. For instance, T_1 maps provide the exponential time constant for the longitudinal relaxation at each voxel, which represents the time taken for the longitudinal magnetisation to return to 63 % of its original value as nuclei in that voxel return from the transverse plane after the delivery of a resonating radiofrequency pulse. Similarly, voxel values in T_2 maps represent the exponential time constant in that voxel, thus providing the expected time for the transverse relaxation to decay to approximately 37 % of the maximum magnitude. However, as discussed previously (see: Basic concepts of MRI), although T_2 decay is expected to be exponential, it is never observed due to inhomogeneities in the magnetic field. Therefore, T_2^* represents the observed time for the transverse magnetisation magnitude to decay to 37 % of its original magnitude, while accounting for the 'true' decay and field inhomogeneities. T_1 relaxometry is often associated with differences in white matter myelination but is also an accurate marker of new tissue growth

(Yeatman et al., 2014), T_2 maps can show a wide array of pathologies because it is highly sensitive to changes in water distributions (MacKay et al., 2006), and T_2^* has long been used as a marker of iron accumulation (Haacke et al., 2005). While the underlying biology fundamentally drives these values, they can be challenging to interpret in isolation; for instance, a difference in relaxometry values may be due to the changes to the number of nuclei present in the voxel or their arrangement in tissues. Combining these modalities presents an opportunity to observe the underlying microstructural changes (Deoni, 2010).

Similarly, diffusion MRI uses various gradient directions to estimate the direction of diffusion for protons in water. With enough gradient directions, it is possible to fit various tensor models to interpret the underlying microstructure, so understanding each metric's contributions to brain structure depends on the model fit to the data. Fractional anisotropy (FA), which denotes the degree of directionality of water diffusion, and mean diffusivity (MD), which denotes the magnitude of water diffusion, are fit from a diffusion tensor model (Alexander et al., 2007). In comparison, orientation dispersion index (ODI), the intracellular volume fraction (V_{ic}), and the isometric volume fraction (V_{iso}) are fit by the NODDI 3-compartment model (Zhang et al., 2012), which requires high angular resolution diffusion imaging. The variation of diffusion angles is measured with ODI, which allegedly represents the fanning of axonal tracts. V_{ic} represents the proportion of a voxel which is explained by the restricted diffusion model and therefore is a marker of neurite density. V_{iso} measures the fraction of a voxel which is explained by the free diffusion compartment and thus is a measure of the unrestricted water content of a voxel.

The first aim of this analysis was to compare the variance of multiple structural metrics between the two groups at each time point. As subtle, widely distributed changes were hypothesised, each modality was compared between groups using a whole brain, voxelwise approach rather than determining measures from *a priori* regions. While the locality of differences was likely to overlap in regions with the largest brain structure differences between divergent modalities, it was hypothesised

that using multiple distinct modalities would provide complementary changes which show little overlap, as each modality teases out different facets of tissue composition. However, small effect sizes require a large statistical sensitivity to expose them and searching for small differences across the whole brain would require greater sensitivity still. For a given modality, the hypothesised differences may not meet test criteria so, an additional dimensionality reduction approach was applied to condense multiple, salient metrics into one compound measure. As each metric describes a different facet of the microstructural processes occurring in a given region, combining them may better characterise the underlying brain structure than any of the modalities alone. Thus, the second aim of this analysis was to measure the regional covariance, as characterised by a multivariate array of morphological features (Seidlitz et al., 2018), between groups and across time. As morphometrically similar regions are more likely to be physically connected (Goulas et al., 2017), the covariance of morphometric characteristics was used to define structural networks. It was hypothesised that, due to the benefits of using multivariate data, this approach would better incorporate the complexity of the ageing brain and better recapitulate the dynamic and continuous traits associated with ageing, compared to the univariate variance analysis.

Methods

Image acquisition and pre-processing

T₁-weighted image and T₁ mapping

T₁w images and quantitative T₁ relaxometry maps were derived from a 3D MP2RAGE sequence which utilises multiple inversion times (TIs = 900 & 3500 ms; flip angle = 7 ° & 9 °). The sequence provided an image with 0.19 mm³ voxel size and a 160x160x128 matrix with TR of 9000 ms, TE of 2.695 ms and echo repetition times of 7.025 ms.

The MP2RAGE images were processed using QUantitative Imaging Tools (QUIT; Wood, 2018). In brief, the phase images from each coil channel were combined from

a reference ultra-short TE image (i.e. TE less than 0.04 ms, with 3.75 ms TR, 4 ° flip angle, and a 128³ matrix to give an image volume of 58 mm³) and the complex reconstruction of the MP2RAGE images (Robinson et al., 2017). The combined image was input to `qi_mp2rage`, which uses the signal at the two inversion times to produce a T₁ map and a T₁w image that is inherently corrected for inhomogeneity of signal intensity across the magnetic field.

Age-specific templates and tissue probability maps were generated using the T₁w images as described previously (see Age-specific population templates and probability maps). Briefly, the T₁w images for each subject were used to generate an average time point template with Advanced Normalisation Tools (ANTs) `antsMultivariateTemplateConstruction2.sh` pipeline (Avants et al., 2011). This template was spatially normalised to a standard template with a manually delineated mask and atlas (Papp et al., 2014), in order to implement an atlas-based brain masking method based on `antsBrainExtraction` (Avants et al., 2010) which removes non-brain tissue (e.g., skull) from each image. The normalisation process requires linearly concatenating the linear and non-linear warp information for both normalisation processes (subject to the average template and average template to the atlas reference image) and applying it to the binary brain mask, thus providing a native-space brain mask for each image. Additionally, Jacobian determinants, which describe the voxelwise deformation of the subject image to the average template space, were created for both the non-linear and affine transforms (absolute Jacobian determinant) and the non-linear alone (relative Jacobian determinant) using ANTs' `CreateJacobianDeterminant` command.

Tissue volumes were estimated by segmenting the T₁w image into three compartments, representing grey matter, white matter, and CSF, using ANTs `Atropos` (Avants et al., 2011b). A label image was generated from the standard template atlas, using regions that are unlikely to be affected by partial volume effects, and thus have signal intensities that are representative of their tissue type. `Atropos` then propagates the label classification throughout the native-space brain mask and

provides a voxelwise probability of being a particular tissue class. Each tissue probability map was qualitatively assessed.

The T_1 map and tissue compartment images were transformed to the time point average space to complete voxelwise comparisons using the linear and non-linear warps provided by the template generation. The time point normalised probabilistic compartment images were subsequently modulated by the absolute Jacobian determinant in order to preserve tissue volume (e.g., grey matter volume, GMV) information at each voxel. The global volume for each tissue class was calculated as the product of voxel size and the sum of values in each Jacobian-modulated tissue compartment image; the sum of the three global values was taken as a measure of total intracranial volume (TIV). Additionally, GMV images were normalised by accounting for GMV fraction per unit of TIV to provide the concentration of grey matter (GMC) at each voxel.

T_2 and T_2^* mapping

T_2 signal was measured using a Bruker 2D multi-slice-multi-echo (MSME) spin-echo protocol, with 12 echoes each with TE of 8 ms and an overall TR of 3600 ms, and a 90° flip angle. The in-plane resolution was 0.156 mm^2 with 0.7 mm slices spaced 1 mm apart, giving a 180×180 matrix which covers 29 mm^2 .

T_2^* signal was measured using Bruker's 2D multiple gradient echo (MGE) sequence, with ten echoes each with TE of 3.5 ms and an overall TR of 1660 ms and a 68° flip angle. The in-plane resolution was also 0.156 mm^2 with 0.7 mm slices spaced 1 mm apart, but the matrix was 180×128 , and the resulting image volume was $29 \times 20 \text{ mm}$.

Maps of both T_2 and T_2^* were generated by fitting a monoexponential decay function to the signal from the multiple echo times using QUIT's `qi_multiecho`, which implements the Auto-Regression on Linear Operations (ARLO) method (Pei et al., 2015). Each map was then separately aligned to the T_1w image by rigid-body registration, thus enabling transformation to the template space by concatenating the

rigid body warp from this stage with the transforms created by normalising the T_{1w} image to the template space.

Diffusion imaging

Diffusion images were created using a 2D DTI spin-echo EPI sequence, which acquired three shells and 60 directions with TE: 19.2 ms and TR 9000 ms, with 90° flip angle. The in-plane resolution was 0.35 mm² with a 0.35 mm slice thickness, giving a FOV of 24 x 19.5.

First, images were corrected for susceptibility induced distortions using FSL's `topup` (Andersson et al., 2003). The original images and images with the phase-encoding direction flipped were used to estimate the distortion warps which were applied to the original image. The susceptibility-corrected images were then aligned to the T_{1w} image with a rigid body registration, to apply the subject-space mask to the diffusion images but also for future normalisation to the time point template, as previously described. Eddy-current and movement distortions were corrected using FSL's `eddy` command (Andersson & Sotiropoulos, 2016).

Voxelwise maps of diffusion tensor scalars (i.e., fractional anisotropy, FA, and mean diffusivity, MD) were created by `dtifit` from FMRIB's Diffusion Toolbox (FDT), which fits a tensor at each voxel of the corrected data. Additionally, a neurite orientation dispersion and density imaging (NODDI) model was used to fit three compartments with distinct microstructural properties: intra-cellular, extra-cellular, and CSF diffusion (Zhang et al., 2012). Voxelwise maps of the fractional volumes representing the intra-cellular (V_{ic}) and CSF (V_{iso}) compartments were produced, in addition to a map of the orientation dispersion index (ODI).

All voxelwise maps were normalised to the appropriate time point template by concatenating the affine transform to the T_{1w} image, and the affine and non-linear transforms to the time point space.

Image analysis

Global analysis of volumetric metrics

Global tissue volumes were calculated by summing the modulated probability images for each tissue class, thus providing volume for each tissue in mm³. The sum of all tissue classes was used to calculate total intracranial volume, which was also used to generate proportions of tissue composition relative to overall brain volume. As there were a large number of missing values, linear mixed effect models were generated using `lme4` and `lmerTest` for R. Conditional growth models with random slopes and intercepts were fit for total intracranial volume, grey matter volume, white matter volume, and CSF volume. Supplementary conditional growth models with random slopes were fitted to predict tissue ratio for grey matter, white matter, and CSF ratios. All dependent variables were predicted by the fixed effects of age and group, and the two-way interaction between age and group, while also controlling for the random effects of age and subject. T-tests used Satterthwaite's method for calculating degrees of freedom. Results are presented as the estimate $\pm$ standard error.

Post-hoc analysis of the mixed models was carried out by calculating the estimated marginal means for each of the fixed effects with `emmeans`, before carrying out pairwise comparisons (control – EEDR) between the group estimated marginal mean for each time point. P-values were adjusted for multiple comparisons using Tukey's multiplicity adjustment.

Voxelwise analysis of volumetric and microstructural metrics

The normalised images for each set of derived metric images (see: Image acquisition and pre-processing) were tested with a mass univariate voxelwise test using FSL's `randomise`. Non-parametric two-sample F-tests were conducted for each metric at each time point to assess for any differences between the EEDR and control groups. To this end, the group labels were exchanged 10,000 times to generate a null distribution, against which the observed differences were tested.

Results figures utilise the dual-coding approach (Allen et al., 2012) which plot both the effect size and the statistical significance of a result to obtain a more comprehensive understanding of the results than standard threshold-based methods. In dual-coded figures, made using the `nanslice` python package (www.github.com/spinacist/nanslice), the effect size defines the colour of the overlay, while the F-test derived p-value determines the opacity of the overlay. An upper threshold of $p = 0.2$ was set to include effects that occur below statistical significance, but regions which surpass the critical threshold are delineated with a black contour line. Voxelwise p-values were derived using threshold-free cluster enhancement (Smith & Nichols, 2009), and uncorrected p-values equal to or less than 0.001 are reported as statistically significant. QUIT's `qi_glmcontrasts` generated the contrast images used to determine the effect size.

Generation of morphometric similarity networks

Morphometric similarity networks were generated with code written in R (version 3.6.1) using metrics derived from T_1w images (GMC and GMV), relaxometry maps (T_1 , T_2 , and T_2^*), and diffusion images (FA, MD, V_{ic} , V_{iso} , and ODI). Each image was warped to the template space, and the mean value was extracted from each grey matter region from an in-house atlas, which was generated by merging two established atlases (Papp et al., 2014; Valdés-Hernández et al., 2011). By only using ROIs which correspond to grey matter regions, there were 154 lateralised regions. The values for each modality were standardised using a robust alternative to z-scores using the following equation:

$$Z_i = \frac{x_i - \tilde{x}}{\text{Median Absolute Deviation}}$$

where $\tilde{x}$ is the sample median value for a given modality as measured across all atlas regions for a subject, and the Median Absolute Deviation is given by calculating the median of the residuals from the sample median. No assumption of normality was inferred on the modality scores since many scores represent a fraction that has a limited range and could exhibit skewed, ceiling, or floor values. Moreover, this

method is robust to outliers (Malo et al., 2006). Once each modality was standardised, regional morphometric arrays were correlated pairwise to create an adjacency matrix per subject, where each edge represents the morphometric similarity between two of the 154 nodes. An example adjacency matrix is available in appendix 3-1.

Network-based statistics

Statistical analysis of the edges in adjacency matrices often requires controlling for a large number of multiple comparisons. A popular tool to control the family-wise error rate is network-based statistics (NBS; Zalesky et al., 2010), which utilises non-parametric permutation testing to perform mass univariate hypothesis testing on the edges of an adjacency matrix. NBS clusters graph edges in a topological space to create graph components which can be tested using standard null-hypothesis tests.

The adjacency matrices, containing Pearson's R correlations, were standardised using Fisher's transform, and input into NBS version 1.2 in MATLAB version 9.8.0. The hypothesis of interest was tested, and an F-statistic assigned at every edge of the adjacency matrix. The graph component with the largest number of connections (i.e., the extent of the component) was established using only the edges that surpass an initial threshold, which was calculated with the aim of detecting an effect size of either $d = 0.2$ or $d = 0.5$ based on the total number of subjects available at each time point. The size of the component was tested against the null hypothesis by exchanging the group labels 5000 times, and p-values less than 0.05 were considered significant.

Graph-theoretical analyses

Thresholding

A second method to control for the number of comparisons in a covariance matrix is to threshold the edges of a graph to reduce the likelihood of spurious correlations influencing the topology of the network. In its simplest form, thresholding is the process of keeping the value of edges that surpass a given threshold while setting all other values to zero, which generates a weighted graph, or by setting all values above

the threshold to one and all values below the threshold to zero, thus creating a binary graph. Thresholds can be absolute, where a threshold is a given value (e.g., $E \geq n$), but this can result in differences in numbers of edges between graphs, which can impact many commonly measured graph metrics (van Wijk et al., 2010). Thus, setting a relative threshold preserves the ratio of edges, by changing the threshold value relative to the total number of possible edges in the network, which is otherwise known as fixing a graph's sparsity. Whichever thresholding method is used, the choice of threshold boundary can be arbitrary but not trivial, since a lenient threshold will increase the likelihood of spurious connections, but a threshold that is too restrictive will disconnect nodes and alter the topological properties of the graph. Consequently, binary graphs were constructed over a series of relative thresholds from 5 % to 50 % of the strongest connections at intervals of 5 %.

The conclusive threshold was chosen to be the sparsity that maximised cost-efficiency ratios and small-worldness. Global cost efficiency is defined as the network's global efficiency at a given sparsity minus the sparsity (Bassett et al., 2009). The global efficiency of a network is the mean of all nodal efficiencies, which can be calculated for a given node, i , as:

$$E_i = \frac{1}{V-1} \sum_{j \in G} \frac{1}{d_{ij}}$$

where V is the number of nodes in a graph (G), and d_{ij} is the shortest path between i and another given node, j . Nodal efficiency illustrates how easily a node can trade information with another node, with a high efficiency indicating a low characteristic path length between a given node and others (Latora & Marchiori, 2001). The cost of a network is directly related to how many edges there are; a sparse graph indicates a low-cost network, while a dense graph represents a high-cost network. The global cost-efficiency was maximised because brain networks display trade-offs between wiring costs and network efficiency, enduring high-cost networks to achieve the minimum efficiency for effective information transfer (Bullmore & Sporns, 2012).

The sparsity intervals were also assessed for their small-worldness, which has several metrics available to characterise it. The first and most widely used metric, σ , described in a previous chapter (see 'The brain as networks'), compares the observed clustering coefficients and characteristic path lengths to those of randomly rewired networks. However, this value is heavily contingent on the density of edges, and therefore it was inappropriate for comparison across various sparsity thresholds (Neal, 2017). Instead, small-worldness was assessed using the small-world propensity (Muldoon et al., 2016), a measure which compares characteristic path length and clustering coefficient to both lattice and random networks with equivalent features, since brain networks display a topology which is neither random nor regular. Small-world propensity can be formally characterised as:

$$\varphi = 1 - \sqrt{\frac{\Delta_C^2 + \Delta_L^2}{2}}$$

where

$$\Delta_C = \frac{C_{latt} - C_{obs}}{C_{latt} - C_{rand}}$$

and

$$\Delta_L = \frac{L_{obs} - L_{rand}}{L_{latt} - L_{rand}}$$

The observed network was rewired to a lattice using an algorithm provided with the original paper, which was ported from MATLAB for compatibility with R. The random network was generated by using the rewire feature from the **igraph** package (Csardi & Nepusz, 2006) to iteratively substituting pairs of edges in the graph using the algorithm detailed by Maslov & Sneppen (2002). The clustering coefficients, characteristic path length, and efficiency of all networks were calculated using functions provided by **igraph** and **brainGraph** (Watson, 2018). As in the original paper, Δ_C and Δ_L were bound between 0 and 1, to ensure φ was also between 0 and 1; therefore, when $\varphi = 1$, it can be concluded that a network demonstrates high small-

world characteristics (i.e., low differences in characteristic path length to a lattice network, and large differences in characteristic path length compared to a random network). However, small-worldness is not categorical, and so Muldoon and colleagues recommend that a network with $0.6 < \phi < 1$ can be considered as displaying small-worldness.

Measures

Once the graph sparsity was selected, the minimum edge weight value for that sparsity was calculated, and adjacency matrix values below the pertinent value were set to zero while the remaining values were set to one, resulting in binary undirected morphometric similarity graph for each subject. Various global graph measures were calculated with `igraph` and `brainGraph` in R, including the small-world propensity (ϕ), global clustering coefficient (C), and global efficiency (E_{glob}) which, as described previously, depends on the characteristic path length (Latora & Marchiori, 2001). Additionally, the graph was clustered using the Leiden algorithm (Traag et al., 2019), which augments the popular Louvain algorithm (Blondel et al., 2008), to obtain the maximum modularity (Q) from the resulting node communities. As modular communities are, by definition, more likely to have links between themselves than expected by random connections, Q quantifies the difference in the number of observed edges between nodes in a module and what is expected by chance in a network. Modularity can be defined formally as

$$Q = \frac{1}{4E} \sum_{ij \in G} \left[A_{ij} - \frac{k_i k_j}{2E} \right] \delta(c_i, c_j)$$

where E is the number of unique edges in a graph, A_{ij} represents the actual number of edges between nodes i and j , k_i and k_j denote the respective degrees of edges for nodes i and j which provides the expected number of edges between them at random, and δ -function is 1 if the community for i , c_i , is the same as the community for j , c_j , and 0 if not (Newman, 2006). If a network's $Q \approx 0$, edges within the community are no stronger than expected by chance. If $Q < 0$, there is no modular structure in a

network. When $Q \approx 1$, the edges show highly modular characteristics, suggesting that nodes in the module communicate predominantly within the community. Thus, an intermediate, positive Q (i.e., $0.3 < Q < 1$) provides evidence of distinct modules with strong intra-community connections, but which still have inter-community connections that mediate transfer of information between modules.

These metrics were calculated for each subject at each time point, and similarly to the global volume data, linear mixed-effects models were generated using `lme4` and `lmerTest`. Conditional growth models with random slopes and random intercepts were fit for each metric separately. These values were predicted by the fixed effects of age and group, the two-way interaction between them, while also controlling for the random effects of age and subject. Further comparison of the groups at each time point was carried out by post-hoc pairwise (control – EEDR) comparison of the estimated marginal means using p-values that were adjusted for multiple comparisons using Tukey’s multiplicity adjustment.

Results

Global volumetric measures

Linear mixed-effects models were fit with the volume of interest (intracranial, grey matter, white matter, and CSF) as the respective outcome variables, and each model used group, age, and the two-way interaction as fixed effect predictors. The random effects of the individual subjects were controlled by including a random slope and intercept for age:

$$Volume_{ij} = (\beta_0 + u_{0ij}) + \beta_1 group_{ij} + (\beta_2 + u_{2ij})age_{ij} + \beta_3 group_{ij}age_{ij} + \varepsilon_{ij}$$

where i and j refer to the i th subject in the j th group. The coefficients for these models are summarised in the table 3-1 and the group trends over time are available in figure 3-2.

Total intracranial volume

The intercept, which refers to the control group at baseline, estimated the control's baseline TIV as $2578.86 \pm 22.37 \text{ mm}^3$. There were significant increases in TIV between baseline and 5 months ($\beta = 102.56 \pm 1.234 \text{ mm}^3$), 11 months ($\beta = 156.71 \pm 11.50 \text{ mm}^3$), and 17 months ($232.25 \pm 17.15 \text{ mm}^3$). There was no significant difference between the volumes of the two groups at baseline ($\beta = -59.02 \pm 31.30 \text{ mm}^3$), but there were significant interaction effects, which indicate that the EEDR group had smaller differences between their baseline and 5-month ($\beta = -67.58 \pm 12.26 \text{ mm}^3$), 11-month ($\beta = -75.77 \pm 15.19 \text{ mm}^3$), and 17-month ($\beta = -82.01 \pm 21.23 \text{ mm}^3$) estimates relative to the same comparison of the control group.

Pairwise comparisons of each group's estimated marginal means at each time point confirms that the volume estimates for the two groups are not significantly different at the 3-month baseline (contrast estimate = $59 \pm 31.3 \text{ mm}^3$, $p = 0.0654$), but the control rats have larger TIV estimates at 5 (contrast estimate = $127 \pm 31.5 \text{ mm}^3$, $p = 0.0002$), 11 (contrast estimate = $135 \pm 33.2 \text{ mm}^3$, $p = 0.0001$) and 17 months (contrast estimate = $141 \pm 36.8 \text{ mm}^3$, $p = 0.0003$).

Grey matter volume

The model predicted GMV for the control baseline as $1362.64 \pm 15.11 \text{ mm}^3$. The grey matter model shows the control group's GMV increased on average relative to baseline at the 5-month ($\beta = 109.12 \pm 13.70 \text{ mm}^3$), 11-month ($\beta = 160.64 \pm 16.97 \text{ mm}^3$), and 17-month ($\beta = 186.55 \pm 24.02 \text{ mm}^3$) time point estimates. There were no significant differences in GMV between the groups at baseline ($\beta = -25.46 \pm 21.14 \text{ mm}^3$), but significant interaction effects denote that the difference between the EEDR group baseline volume and 5 ($\beta = -94.35 \pm 19.17 \text{ mm}^3$), 11 ($\beta = -104.43 \pm 22.45 \text{ mm}^3$), and 17-

month ($\beta = -6.32 \pm 29.62 \text{ mm}^3$) time point estimates were smaller than the equivalent differences for the control group.

Post-hoc comparisons confirm that there are no statistically significant differences between the groups' GMV at baseline (contrast estimate = $25.5 \pm 21.1 \text{ mm}^3$, $p = 0.2327$), but the controls had more grey matter than the EEDR group at 5 (contrast estimate = $119.8 \pm 21.5 \text{ mm}^3$, $p < 0.0001$), 11 (contrast estimate = $129.9 \pm 25.2 \text{ mm}^3$, $p < 0.0001$) and 17 months (contrast estimate = $91.8 \pm 32.5 \text{ mm}^3$, $p = 0.0068$).

White matter volume

Baseline white matter volume was estimated for the control rats as $591.91 \pm 6.48 \text{ mm}^3$, which significantly increased relative to baseline at 5 months ($\beta = 45.35 \pm 7.46 \text{ mm}^3$), 11 months ($\beta = 52.73 \pm 9.48 \text{ mm}^3$), and 17 months ($\beta = 84.49 \pm 13.82 \text{ mm}^3$). There were no significant differences between the groups' estimated white matter volume at baseline ($\beta = -0.76 \pm 9.07 \text{ mm}^3$), but the EEDR rats were estimated to have smaller differences between their baseline and 5-month time point ($\beta = -45.75 \pm 10.44 \text{ mm}^3$), and the baseline and 11-month time point ($\beta = -49.55 \pm 12.54 \text{ mm}^3$), compared to the same differences in the control group. However, the difference in white matter volumes between baseline and the 17 months was not significantly different between the two groups ($\beta = -0.78 \pm 17.07 \text{ mm}^3$).

Post-hoc comparisons confirm these findings; there were no significant differences in the white matter volume between the two groups at the baseline (contrast estimate = $0.758 \pm 9.07 \text{ mm}^3$, $p = 0.9336$) or the final time point (contrast estimate = $31.540 \pm 17.46 \text{ mm}^3$, $p = 0.0769$), but the controls had significantly more white matter volume at 5 months (contrast estimate $46.507 \pm 9.43 \text{ mm}^3$, $p < 0.0001$) and 11 months (contrast estimate = $50.308 \pm 12.35 \text{ mm}^3$, $p = 0.0001$).

CSF volume

Finally, with respect to the volume of CSF, the model estimated that the baseline CSF volume for the control group was $589.94 \pm 10 \text{ mm}^3$. Unlike the previous models described, the control group CSF volume did not differ between baseline and 5 months ($\beta = 1.58 \pm 10.77 \text{ mm}^3$), baseline and 11 months ($\beta = -6.47 \pm 12.52 \text{ mm}^3$), or baseline and 17 months ($\beta = -3.99 \pm 16.36 \text{ mm}^3$). CSF volume at baseline is not significantly different between the groups ($\beta = -20.37 \pm 13.99 \text{ mm}^3$). Moreover, there are no significant interaction effects, which suggests that the differences between baseline and 5 months ($\beta = -14.12 \pm 15.06 \text{ mm}^3$), baseline and 11 months ($\beta = -12.10 \pm 16.62 \text{ mm}^3$), and baseline and 17 months ($\beta = -14.72 \pm 20.18 \text{ mm}^3$) for the EEDR group are similar to those observed between the same time points in the control group. Due to the lack of significant effects, post-hoc comparisons were not pursued for this model.

TIV model					
	β	SE	df	t value	P value
Intercept: difference from 0					
Control at 3 months	2578.86	22.37	47.66	115.28	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-59.02	31.31	47.66	-1.89	0.0655
Control at 5 months	102.56	8.76	78.27	11.71	<0.0001
Control at 11 months	156.71	11.50	91.65	13.63	<0.0001
Control at 17 months	232.25	17.15	38.50	13.54	<0.0001
Interaction effects: difference from control change between 3 months and time point					
EEDR difference in 3 to 5 months	-67.57	12.26	78.27	-5.51	<0.0001
EEDR difference in 3 to 11 months	-75.77	15.19	91.65	-4.99	<0.0001
EEDR difference in 3 to 17 months	-82.01	21.23	35.38	-3.86	0.0005
Grey matter model					
Intercept: difference from 0					
Control at 3 months	1362.64	15.11	61.08	90.21	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-25.46	21.14	61.08	-1.20	0.2330
Control at 5 months	109.12	13.70	93.90	7.97	<0.0001
Control at 11 months	160.64	16.97	101.66	9.47	<0.0001
Control at 17 months	186.55	24.02	48.44	7.77	<0.0001
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	-94.35	19.17	93.90	-4.92	<0.0001
EEDR difference in 3 to 11 months	-104.43	22.45	102.04	-4.65	<0.0001
EEDR difference in 3 to 17 months	-66.32	29.62	46.55	-2.24	0.0300
White matter model					
Intercept: difference from 0					
Control at 3 months	591.91	6.48	78.48	91.31	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-0.76	9.07	78.48	-0.08	0.9336
Control at 5 months	45.35	7.46	91.74	6.08	<0.0001
Control at 11 months	52.73	9.48	103.91	5.56	<0.0001
Control at 17 months	84.49	13.82	62.36	6.11	<0.0001
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	-45.75	10.44	91.74	-4.38	<0.0001
EEDR difference in 3 to 11 months	-49.55	12.54	104.41	-3.95	0.0001
EEDR difference in 3 to 17 months	-30.78	17.07	59.18	-1.80	0.0765
CSF model					
Intercept: difference from 0					
Control at 3 months	589.94	10.00	89.24	59.00	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-20.37	14.00	89.24	-1.46	0.1490
Control at 5 months	1.58	10.76	86.61	0.15	0.8837
Control at 11 months	-6.47	12.52	114.82	-0.52	0.6062
Control at 17 months	-3.99	16.36	67.98	-0.24	0.8083
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	-14.12	15.06	86.61	-0.94	0.3512
EEDR difference in 3 to 11 months	-12.11	16.63	112.52	-0.73	0.4680
EEDR difference in 3 to 17 months	-14.73	20.18	69.75	-0.73	0.4679

Table 3-1: Coefficients for the brain volume models, predicting total intracranial (TIV), grey matter (GM), white matter (WM), and CSF volume from group, age, and the interaction between group and age

Figure 3-2: Median global tissue volume measures for A) intracranial volume, B) grey matter volume C) white matter volume, D) CSF volume per group. Tissue volume measures across study are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

Tissue proportions of brain volume

The proportion of each tissue relative to TIV was calculated to test if the observed change in TIV was driving the differences in individual tissue classes. The trend of these changes can be observed in figure 3-3, and similar linear mixed-effects models were fit as for the absolute volume data:

$$\text{Tissue proportion}_{ij} = \beta_0 + \beta_1 \text{group}_{ij} + (\beta_2 + u_{2ij}) \text{age}_{ij} + \beta_3 \text{group}_{ij} \text{age}_{ij} + \varepsilon_{ij}$$

where i and j refer to the i th subject in the j th group.

Grey matter proportion

The grey matter proportion for the control baseline was estimated as $53.6 \pm 0.2 \%$. The grey matter proportion increased on average relative to baseline at the 5-month ($\beta = 0.9 \pm 0.3 \%$), 11-month ($\beta = 1.8 \pm 0.3 \%$), and 17-month ($\beta = 1.4 \pm 0.4 \%$) time point estimates. There were no significant differences in grey matter proportion between the groups at baseline ($\beta = -0.0 \pm 0.3 \%$), and there were no significant interaction effects, indicating that the magnitude of difference between the baseline grey matter proportion and 5 ($\beta = 0.4 \pm 0.4 \%$), 11 ($\beta = -0.5 \pm 0.4 \%$), and 17-month ($\beta = -0.0 \pm 0.5 \%$) time point estimates were the same for both groups.

White matter proportion

The white matter proportion for the control baseline was estimated as $23.3 \pm 0.2 \%$, and it did not increase on average relative to baseline at 5 months ($\beta = 0.3 \pm 0.3 \%$), 11-month ($\beta = 0.1 \pm 0.3 \%$), or 17-month ($\beta = 0.7 \pm 0.5 \%$) time points. Furthermore, there were no significant differences in white matter proportion between the groups at baseline ($\beta = 0.4 \pm 0.3 \%$), and there were no significant interaction effects, indicating that the magnitude of difference between the baseline white matter proportion and 5 ($\beta = -0.3 \pm 0.4 \%$), 11 ($\beta = -0.4 \pm 0.4 \%$), and 17-month ($\beta = -0.1 \pm 0.6 \%$) time point estimates were the same for both groups.

CSF proportion

Finally, the CSF proportion was estimated as $23.2 \pm 0.3 \%$ for the control rats at baseline. This proportion decreased relative to baseline at 5 ($\beta = -1.3 \pm 0.4 \%$), 11 ($\beta = -1.9 \pm 0.5 \%$), and 17 months old ($\beta = -2.2 \pm 0.6 \%$). However, there was no significant differences in CSF proportion between the groups at baseline ($\beta = -0.4 \pm 0.4 \%$), and there were no significant interaction effects, indicating that the magnitude of difference between the baseline CSF proportion and 5 ($\beta = 0.7 \pm 0.5 \%$), 11 ($\beta = 0.8 \pm 0.6 \%$), and 17-month ($\beta = 0.2 \pm 0.7 \%$) time point estimates were the same for both groups.

Figure 3-3: Proportion of tissue volumes relative to total intracranial volume for both control (left) and EEDR (right) rats across the length of the study, which is fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

Voxelwise volumetric measures

To visualise the locality of the volumetric differences observed in the previous section, voxel-based morphometry of grey matter and tensor-based morphometry of the log Jacobian determinants were compared on a voxelwise basis between the two groups at each time point. As the global volume models indicated that there were no significant differences between the groups at baseline, these are not described here but instead are included in appendix 3-2A.

Voxel-based morphometry of grey matter

Grey matter volume

The widespread increase in GMV in the control rats, compared to the EEDR rats, at 5 months can be observed in figure 3-4A. The regions with the largest increases in GMV for the control rats, both in terms of effect size and statistical significance, appear to be the dorsal cortical surface. However, there are many symmetrical subcortical clusters, in regions including the superior parts of the dorsal hippocampus, amygdala, globus pallidus, and the thalamus, where the control rats have more grey matter than the EEDR rats. There are far fewer regions where the EEDR rats had higher GMV compared to controls at the 5-month-old time point following EEDR intervention. Nevertheless, there are small clusters, located in the caudal parietal cortex and inferior dorsal hippocampus, although they mostly do not surpass significance criteria.

Many of the differences in GMV observed at the third time point (figure 3-4B) are more widespread in their effects but less statistically significant than the previous time point. The cortical clusters are still the most significant difference between the control and EEDR rats, denoting an increase of GMV at the cortex of the control group rats. The subcortical clusters mentioned previously are still present, and appear more widespread in terms of effect size, but do not reach significance in this instance. With regards to the increased GMV in EEDR rats relative to controls, the cluster denoting an increase in the caudal parietal cortex and dorsal hippocampus in EEDR rats is still present, but no longer surpass significance criteria.

However, almost all localised GMV differences between the two groups disappeared by the last time point (figure 3-4D), and only weak clusters appear, the direction of which suggests an increase in GMV in the controls with almost no evidence of an increase in the EEDR rats.

Grey matter concentration

In comparison to GMV, which uses absolute grey matter differences regardless of starting brain size, GMC accounts for differences in TIV. Thus, GMC shows regions where there are changes in grey matter despite differences in overall brain size.

At the first time point following EEDR intervention (figure 3-4D), many clusters mirror the GMV results; there are statistically significant clusters indicating increases in GMC in the dorsal cortical surface and amygdala for the control rats and increased GMC in the caudal parietal cortex in the EEDR rats. However, the globus pallidus effect size indicates a larger GMC in the EEDR rats than the control rats, despite the cluster not reaching a lenient statistical significance. Moreover, clusters in the anterior cingulate, basal forebrain, striatum, and insular cortex, which were not regions with clusters in the GMV results, indicate slightly larger GMC in the EEDR rats than controls, although these clusters do not surpass the statistical threshold.

Most of the effects present at 5 months also appear at 11 months (figure 3-4E), although all but the cortical surface clusters are no longer statistically significant. The bilateral insular cortex clusters are no longer convincing, and the basal forebrain cluster has shrunk, suggesting there is now no difference between the groups in these regions. A pair of clusters in the bilateral medial prefrontal cortex shows higher GMC in the controls, although it does not reach statistical significance; this cluster is present in the first two GMV, but not the GMC from the 5-month time point.

Finally, at the last time point (figure 3-4F), which represents approximately 17 months old, localised differences in the groups' GMC have disappeared, which mirrors the GMV results at this time point. In comparison to GMV, where the weak clusters appeared to indicate higher GMV in the control rats, the weak clusters resulting from comparing the final time point's GMC between the two groups suggests that most of the GMC differences are slightly higher in the EEDR rats, although with some clusters which suggest higher GMC in controls compared to EEDR.

Tensor-based morphometry

Absolute Jacobian Determinant

The absolute Jacobian Determinant is derived from both the linear and non-linear transformations, and thus includes scaling transformations. After the implementation of EEDR, the control rats require more contraction warping overall to match the template (figure 3-5A), which is indicative that the control rat brains are larger than EEDR. Localised effects are observed across much of the cortex and thalamus, with many of these regions reaching statistical significance.

The same pattern is observed at the 11-month time point (figure 3-5B); the difference in absolute Jacobian Determinant indicates the control rats still have larger brain volumes, which is most clearly observed in differences lateral posterior cortex and the dorsal hippocampi. However, the thalamic cortical cluster no longer reaches significance, although the direction of the effect does not seem to have changed with EEDR rats still showing smaller thalamic absolute Jacobian Determinants.

However, at 17 months (figure 3-5C), there are no differences observed that are statistically significant. The magnitude of the difference still shows larger absolute Jacobian Determinants in the controls, and thus the controls have larger brain volume particularly in the cingulate cortex, but the overall effect appears to be weaker than previous time points.

Relative Jacobian Determinant

Contrary to the absolute Jacobian Determinant, the relative Jacobian Determinant is derived from only the non-linear transformations and so does not include scaling transformations and thus disregards differences in starting volume. Correspondingly, after adjustment for overall brain volume, the EEDR rats require more contraction warping to match the template (figure 3-5D), which indicates the average EEDR rat brain has a larger volume than that of the control group. Localised effects are observed across much of the cortex, particularly across the entorhinal and perirhinal

cortices, prefrontal cortex, and cingulate cortex, and subcortical regions including the basal forebrain, the striatum, and pretectum.

Many of these regional differences are observed again at 11 months (figure 3-5E), with the biggest differences in relative Jacobian Determinant occurring at the retrosplenial, sensorimotor, parietal and prefrontal cortices, and in subcortical regions such as the dorsal hippocampi, pretectum, striatum, and basal forebrain. Clusters on the insular cortex suggest that there may be an increase in relative Jacobian Determinant in the EEDR rats, but it does not reach significance.

By 17 months (figure 3-5F), the patterns of difference between the groups have changed. There are no clusters that reach statistical significance but, for the first time, it appears that some regions are larger in the control rats relative to the EEDR rats, such as the anterior cingulate cortex.

Figure 3-4: Volumetric voxelwise differences between groups at the time points following the intervention. Grey matter volume (GMV; A-C) and concentration (GMC; D-F) differences for control – EEDR. The absolute difference in grey matter volume at A) 5 months, B) 11 months and C) 17 months. Differences in grey matter volume relative to brain volume at D) 5 months, E) 11 months, F) 17 months. P-values are uncorrected for multiple comparisons and are represented by the opacity of the overlay, with contours denoting significance criteria being met at $p < 0.001$. The colour of the overlay represents the effect size of the contrast.

Figure 3-5: Voxelwise differences for tensor-based morphometry between groups at each time point following the intervention. Differences in absolute (A-C) and relative (D-F) Jacobian for control – EEDR. Absolute differences at A) 5 months, B) 11 months and C) 17 months. Differences relative to brain size at D) 5 months, E) 11 months, F) 17 months. P-values are uncorrected for multiple comparisons and are represented by the opacity of the overlay, with contours denoting significance criteria being met at $p < 0.001$. The colour of the overlay represents the effect size of the contrast.

Voxelwise relaxometry measures

The factors that drive volumetric changes can be ambiguous, and thus relaxometry was utilised in order to probe the microstructural changes that underlie the observed volume changes in the previous section. For consistency with the previous section, the time point before EEDR implementation will not be discussed here, but the results are available in appendix 3-2B. Although many of these results do not surpass statistical significance, and therefore, like the other lenient-thresholded results, must be treated with restraint, many of these voxelwise results emulate known tissue boundaries despite no anatomical constraint during the statistical analysis.

T_1

In the first time point following EEDR (figure 3-6A), there are anatomically aligned clusters where T_1 times are longer in the controls, such as in the ventricular system aqueduct, and longer T_1 times in the striatum, the corpus callosum, and deep cerebral white matter regions for the EEDR rats. At the second time point after intervention (figure 3-6B), this pattern is observed again, but with larger effect sizes than the previous time point. The control group continue to show longer T_1 values in the ventricular system and areas surrounding the brain, while the T_1 is longer in the deep white matter, the corpus callosum and much of the dorsal cortex for the EEDR rats compared to the controls. By the final time point, the difference in T_1 time has some small clusters that surpass statistical significance. More importantly, the effect size shows longer T_1 time is widespread in the EEDR rats' grey matter relative to the controls (figure 3-6C).

T_2

While no clusters surpass the statistical threshold for the difference in T_2 time between the groups, and the maps are messy, the overall impression is that T_2 value is slightly longer in the controls at the first time point after EEDR intervention (figure 3-6D), particularly in regions associated with the ventricular system or at the edge of the brain. At 11 months (figure 3-6E), the pattern is very similar to the 5-

month time point with T_2 relaxation time longer around the brain in the controls, but with additional clusters showing longer T_2 times in EEDR rats' deep white matter and tectospinal tract. However, even weak localised differences between the groups' T_2 relaxation times at the final time point (figure 3-6F) has completely disappeared.

T_2^*

At 5 months (figure 3-6G), there are some ventrocaudal regions where T_2^* is significantly longer in the controls than the EEDR rats, including the pontine nuclei and the subcortical grey matter, near the reticular formation, and the bilateral piriform cortices. Inversely, at 11 months, the EEDR rats show longer T_2^* relaxation times symmetrically in the corpus callosum, hippocampal fimbria, and the superior colliculi (figure 3-6H). In a similar manner to the T_2 relaxation time, there is no evidence of differences between the groups' T_2^* relaxation time present at 17 months (figure 3-6I)

Figure 3-6: Voxelwise differences for relaxometry between groups at the time points following the intervention. Differences in T_1 (A-C), T_2 (D-F), and T_2^* (G-I) times for control – EEDR. Differences in T_1 relaxation at A) 5 months, B) 11 months and C) 17 months. Differences in T_2 relaxation at D) 5 months, E) 11 months, and F) 17 months. Differences in T_2^* relaxation at G) 5 months, H) 11 months, and I) 17 months. P-values are uncorrected for multiple comparisons and are represented by the opacity of the overlay, with contours denoting significance criteria being met at $p < 0.001$. The colour of the overlay represents the effect size of the contrast.

Voxelwise diffusion measures

Another method to probe the microstructure underpinning volumetric differences uses diffusion imaging. There were no differences in the DTI metric at 3 months, so only differences in NODDI scalars are given in appendix 3-2C. Almost all diffusion clusters fail to surpass the lenient statistical significance and are therefore cautiously discussed if they emulate anatomical boundaries or show bilateral symmetry.

Tensor model metrics

Fractional Anisotropy

The control rats show higher FA in the secondary visual and entorhinal cortices at 5 months, while symmetrical bilateral clusters indicate that the EEDR rats have higher FA in the brainstem and thalamus (figure 3-7A), although these effects appear to be reasonably small. At 11 months, EEDR rats show more convincing clusters in the dorsal and ventral hippocampi, globus pallidus, and basal forebrain indicating higher FA in EEDR rats (figure 3-7B). The ventricular system in the EEDR rats also displays higher FA values than in the control rats. There is almost no difference between the groups at 17 months (figure 3-7C). There are symmetrical clusters in the sensorimotor cortices indicating an increase in FA for the EEDR rats in these regions, but the effect size of these is very weak.

Mean diffusivity

At 5 months, control rats have slightly higher MD in the ventricular system, including the caudal choroid plexus and the third ventricle, and the dorsal surface of the brain (figure 3-7D). Comparatively, EEDR rats have higher MD in the deep white matter, the reticular formation, the entorhinal cortex, the rostral choroid plexus, and external capsule. Similarly, at 11 months (figure 3-7E), the control rats have higher MD in the ventricular system, particularly the third and lateral ventricles, and on the dorsal surface of the brain. However, the regions where EEDR rats had higher MD at 5 months have largely disappeared by 11 months. Similar to FA, at 17 months old, there are almost no localised differences in MD between the groups (figure 3-7F).

NODDI model metrics

Isotropic diffusion fraction

At 5 months (figure 3-8A), V_{iso} is mostly higher in the controls than the EEDR cohort, particularly in the basal forebrain, hypothalamus, posterior thalamus, and at the boundaries of the brain and the ventricular system. At 11 months (figure 3-8B), the pattern of clusters is very similar to that of the 5-month time point, with control rats having higher V_{iso} in regions close to the ventricular system. The clusters at the basal forebrain and piriform cortex have become weaker in effect, but the difference seems to have increased in the rostral corpus callosum and the ventral hippocampus. By the final time point (figure 3-8C), many of the biologically plausible clusters disappear, leaving only symmetrical clusters in the caudal entorhinal cortex indicating increased V_{iso} in the EEDR rats.

Intracellular diffusion fraction

Control rats showed higher V_{ic} relative to their EEDR counterparts at the 5-month point (figure 3-8D) in many regions, including the dorsal cortical surface, the hypothalamus, and the ventral hippocampi. Furthermore, some of these regions surpass statistical significance. There are also a pair of bilateral clusters located on the entorhinal cortex where the EEDR rats have higher V_{ic} than control rats. At the following time points, many of the observed increases of V_{ic} in the control rats largely disappear, with increases only on the ventral brain surface and in the pineal gland for control rats relative to EEDR rats at the 11-month time point (figure 3-8E), and no observable differences between the groups at the final time point (figure 3-8F).

Orientation dispersion index

The controls show higher ODI than the EEDR rats at 5 months (figure 3-8G), which is again mostly located on the ventral brain surface with differences of a smaller magnitude in the thalamus, and superior colliculi. At 11 months, the increased ODI of the ventral brain surface in the controls is still visible but additional clusters appear in the dorsal hippocampus denoting higher ODI in the controls (figure 3-8H). In

contrast, it appears there are still differences in the superior colliculi, but they are now higher in the EEDR rats than the controls, along with new clusters in the ventral hippocampus and the presubiculum. However, at 17 months (figure 3-8I), the clusters that remain, which are located mainly on the cortex and denote higher ODI in controls relative to EEDR, have relatively large p-values and should be considered with caution.

Figure 3-7: Voxelwise differences for diffusion tensor metrics between groups at the time points following the intervention. Differences in FA (A-C), and MD (D-F) for control – EEDR. Differences in FA values at A) 5 months, B) 11 months and C) 17 months. Differences in MD values at D) 5 months, E) 11 months, and F) 17 months. P-values are uncorrected for multiple comparisons and are represented by the opacity of the overlay, with contours denoting significance criteria being met at $p < 0.001$. The colour of the overlay represents the effect size of the contrast.

Figure 3-8: Voxelwise differences for NODDI model metrics between groups at the time points following the intervention. Differences in V_{iso} (A-C), V_{ic} (D-F), and ODI (G-I) for control – EEDR. Differences in V_{iso} values at A) 5 months, B) 11 months and C) 17 months. Differences in V_{ic} values at D) 5 months, E) 11 months, and F) 17 months. Differences in ODI values at D) 5 months, E) 11 months, and F) 17 months. P-values are uncorrected for multiple comparisons and are represented by the opacity of the overlay, with contours denoting significance criteria being met at $p < 0.001$. The colour of the overlay represents the effect size of the contrast.

Graph-theoretical analyses of structural networks

Network-based statistics

Network-based statistics was used to compare the Fisher-transformed morphometric similarity graphs, between groups for each time point separately, in a mass univariate analysis that is comparable to the voxelwise analysis of the previous section. No components surpassed either initial threshold at the baseline time point or the 17-month time point. At the 5-month and 11-month time points, each time point had one network, with 3941 and 3780 edges respectively, that surpassed the initial threshold of $d = 0.2$ (figure 3-9).

To better discern the largest changes, the initial threshold was increased to $d = 0.5$, and only the 5-month time point had a network that surpassed this. In this network, 1558 edges were significantly different between the groups at 5 months (figure 3-10). The distribution of differences seems fairly equal between edges which denote higher connectivity in the EEDR and edges which denote higher connectivity in the controls, as does the distribution of differences between inter-hemispheric and intra-hemispheric edges. In particular, connections involving the parietal association, sensory, and visual cortices ostensibly have the largest differences between the groups, denoting higher regional covariance in the EEDR group compared to the controls.

Figure 3-9: Group differences in network-based statistics. Differences in edges between groups for 5 months (left) and 11 months (right) that surpassed the initial threshold, $d = 0.2$, are coloured to denote the magnitude of the difference between mean control – mean EEDR edge weight.

Graph thresholding

Figure 3-11: Structural connectome cost-efficiency and small-world propensity for a range of densities (5 % to 50 % of the strongest connections). The points are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

Although the highest small-world propensity was observed in the sparsest graphs, the same sparsity resulted in the lowest cost-efficiency and is thus unlikely to represent brain networks (figure 3-11). The peak of cost-efficiency was observed when sparsity was approximately 20-25 %, and since 20 % had the higher cost-efficiency scores of the two densities, 20 % was selected as the graph density for further analyses. The inclusion of an additional 5 % of connections did not markedly improve cost-efficiency or small-worldness, and the minimal number of edges

required to reconstruct these networks was preferable to reduce the likelihood of including spurious correlations. The full table of values is available in appendix 3-3.

Global graph measures

Linear mixed-effects models were fit for each of the global metrics (figure 3-12) using the same format as used to calculate the ratio models previously:

$$Global\ metric_{ij} = \beta_0 + \beta_1 group_{ij} + (\beta_2 + u_{2ij})age_{ij} + \beta_3 group_{ij}age_{ij} + \varepsilon_{ij}$$

where i and j refer to the i th subject in the j th group.

Global efficiency

Global efficiency was estimated as 0.392 ± 0.004 for the 3-month-old controls. Comparing the controls across time, there are no significant differences between baseline and 5 months ($\beta = 0.005 \pm 0.005$), 11 months ($\beta = 0.005 \pm 0.006$), or 17 months ($\beta = 0.005 \pm 0.008$). The groups do not significantly differ at baseline ($\beta = 0.003 \pm 0.005$), but there was a significant interaction effect between group and the 5-month time point ($\beta = -0.020 \pm 0.007$), which suggests that the difference between global efficiency measures from the baseline and the 5-month time points in the EEDR group was smaller than the same difference in the controls. This difference only appears at 5 months, as there was no interaction between group and the 11-month time point ($\beta = -0.011 \pm 0.008$) or the 17-month time point ($\beta = -0.008 \pm 0.010$), indicating that the differences between baseline and these time points are the same for both groups.

Post-hoc comparisons of the groups' EMM at each time points confirms this, with no significant difference between the groups' mean global efficiency at baseline (contrast estimate = -0.003 ± 0.005 , $p = 0.5453$), 11 months old (contrast estimate 0.008 ± 0.006 , $p = 0.1932$), or 17 months old (contrast estimate = 0.006 ± 0.009 , $p = 0.4989$). However, the mean global efficiency is significantly lower in the EEDR group at the 5-month time point (contrast estimate = 0.0172 ± 0.005 , $p = 0.0014$).

Clustering coefficient

The clustering coefficient was estimated for the control group at baseline as 0.625 ± 0.011 , and the control group clustering coefficient did not change over time, as indicated by the difference between baseline and 5 months ($\beta = -0.004 \pm 0.015$), 11 months ($\beta = 0.005 \pm 0.017$), or 17 months ($\beta = -0.006 \pm 0.022$). There was no significant difference between the groups at baseline ($\beta = -0.002 \pm 0.015$). There were no significant interaction effects, indicating that the differences in clustering coefficient over time for the EEDR group were the same size as the differences for the control group, whether that is between baseline and 5 months ($\beta = 0.036 \pm 0.022$), baseline and 11 months ($\beta = 0.023 \pm 0.023$), or baseline and 17 months ($\beta = 0.032 \pm 0.027$). Due to the lack of significant predictors, post-hoc comparisons were not pursued.

Modularity

The modularity of the control group at baseline was estimated as 0.606 ± 0.005 . In terms of comparing the modularity of the control group across time, the difference between baseline and the 17-month time points ($\beta = 0.013 \pm 0.012$) did not pass significance criteria. However, the difference between baseline and the 5-month time point ($\beta = 0.020 \pm 0.008$) and baseline and 11-month time point ($\beta = 0.029 \pm 0.009$) were statistically significant, indicating an increase in modularity from baseline to the respective time points for the control group. There was no significant difference between the groups at baseline ($\beta = 0.005 \pm 0.008$), and no significant interactions between group and time point, which indicates that the EEDR group had differences of the same magnitude as the respective control differences between the baseline and 5-month ($\beta = 0.007 \pm 0.011$), 11-month ($\beta = -0.007 \pm 0.012$), and 17-month time points ($\beta = 0.020 \pm 0.014$).

Post-hoc comparison between the group EMM show that there are no statistically significant differences between the groups. At baseline (contrast

estimate = -0.005 ± 0.008 , $p = 0.5209$), 5 months (contrast estimate = -0.012 ± 0.008 , $p = 0.1352$), or 11 months (contrast estimate = 0.002 ± 0.009 , $p = 0.8441$), the size of the differences in modularity is approximately zero. However, at 17 months old (contrast estimate = -0.025 ± 0.013 , $p = 0.0522$), the contrast estimate indicates the EEDR group have a slightly higher modularity, but these differences do not surpass significance criteria.

Small-World Propensity

Small-world propensity was estimated at 3 months for the control group as 0.894 ± 0.003 . Comparing the change in small-worldness in the controls across time, no significant difference between the control group's small-worldness at baseline and at 5 months ($\beta = 0.009 \pm 0.005$), 11 months ($\beta = 0.006 \pm 0.005$), or 17 months ($\beta = 0.008 \pm 0.007$). It is important to note that the groups did have significantly different mean small-world propensities before intervention ($\beta = 0.012 \pm 0.005$). Following EEDR intervention, there is a significant interaction effect between group and the 5-month time point ($\beta = -0.021 \pm 0.007$), indicating that, at 5 months, the difference to baseline small-worldness is less for the EEDR rats than the control rats, but this interaction is absent at 11 months ($\beta = -0.013 \pm 0.007$) or 17 months ($\beta = -0.011 \pm 0.009$).

Post-hoc comparison shows that there the EEDR rats showed more small-worldness than the controls at baseline (contrast estimate = -0.012 ± 0.005 , $p = 0.0114$). However, comparing the groups at 5 months shows that the controls have a slightly higher small-world propensity, although this difference is not statistically significant (contrast estimate = 0.008 ± 0.005 , $p = 0.0816$). There were no differences in measured small-worldness between the group means at 11 months (contrast estimate = 0.005 ± 0.006 , $p = 0.9333$) or 17 months (contrast estimate = -0.007 ± 0.007 , $p = 0.9229$).

Figure 3-12: Global graph theory measures for structural networks, including global efficiency (E_{glob}), clustering coefficient (C), modularity (Q), small-world propensity (ϕ), as derived from morphometric similarity measures for both groups across time using a sparsity of 20 %. The points are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

E_{glob} model					
	β	SE	df	t value	P value
Intercept: difference from 0					
Control at 3 months	0.3921	0.0037	130.24	106.40	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	0.0032	0.0052	130.24	0.61	0.5452
Control at 5 months	0.0055	0.0052	123.77	1.06	0.2905
Control at 11 months	0.0051	0.0060	154.81	0.85	0.3942
Control at 17 months	0.0046	0.0078	115.39	0.59	0.5596
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	-0.0204	0.0073	123.77	-2.78	0.0063
EEDR difference in 3 to 11 months	-0.0114	0.0080	152.57	-1.42	0.1566
EEDR difference in 3 to 17 months	-0.0089	0.0097	125.36	-0.92	0.3588
C model					
Intercept: difference from 0					
Control at 3 months	0.6254	0.0108	127.93	57.95	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-0.0017	0.0153	127.93	-0.11	0.9131
Control at 5 months	-0.0040	0.0152	122.34	-0.26	0.7935
Control at 11 months	0.0048	0.0173	152.39	0.28	0.7802
Control at 17 months	-0.0063	0.0220	114.86	-0.29	0.7730
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	0.0359	0.0215	122.34	1.67	0.0976
EEDR difference in 3 to 11 months	0.0226	0.0233	149.52	0.97	0.3327
EEDR difference in 3 to 17 months	0.0315	0.0274	127.93	1.15	0.2519
Q model					
Intercept: difference from 0					
Control at 3 months	0.6068	0.0054	126.79	112.10	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	0.0049	0.0077	126.79	0.644	0.5209
Control at 5 months	0.0198	0.0076	119.40	2.60	0.0106
Control at 11 months	0.0286	0.0088	154.89	3.24	0.0015
Control at 17 months	0.0126	0.0116	110.18	1.09	0.2778
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	0.0067	0.0108	119.40	0.62	0.5336
EEDR difference in 3 to 11 months	-0.0068	0.0118	152.43	-0.57	0.5679
EEDR difference in 3 to 17 months	0.0201	0.0144	120.312	1.40	0.1641
ϕ model					
Intercept: difference from 0					
Control at 3 months	0.8942	0.0034	130.66	266.52	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	0.0122	0.0047	130.66	2.57	0.0114
Control at 5 months	0.0091	0.0047	125.13	1.93	0.0559
Control at 11 months	0.0064	0.0054	153.51	1.19	0.2362
Control at 17 months	0.0079	0.0069	117.67	1.14	0.2566
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	-0.0206	0.0067	125.13	-3.07	0.0026
EEDR difference in 3 to 11 months	-0.0126	0.0073	151.03	-1.74	0.0838
EEDR difference in 3 to 17 months	-0.0115	0.0086	129.23	-1.33	0.1860

Table 3-2: Coefficients for structural network graph theory models predicting global efficiency (E_{glob}), global clustering coefficient(C), modularity (Q), and small-world propensity (ϕ) from group, age, and the interaction between group and age

Summary

This work measured a variety of structural metrics of the longitudinal cohort in the Resilient study to characterise structural changes, both global and microstructural, in the ageing rat brain. Differences between rats with standard and enriched rearing were also explored to establish which modality or combination of modalities, if any, explained inter-individual variance observed in the process of ageing. It was hypothesised that changes between the groups would likely be subtle and widely distributed across the brain, and the process of identifying the microstructural underpinning of any changes would be facilitated with multiple complementary modalities. Finally, by creating a characteristic metric derived from all the structural modalities and measuring the covariance between different brain regions, the secondary hypothesis, that combining multiple modalities would better encapsulate the complex and dynamic traits of ageing compared to the univariate variance analysis, was also tested.

Increased brain volume with age

Total intracranial volume, grey matter, and white matter volumes increased with age. These findings do not recapitulate what has been found in the human ageing process, where total brain volume begins to decline in middle age, concurrent decreases with grey matter and increases in CSF from adolescence, and increases in white matter volume until middle age (Narvacan et al., 2017). It is unclear if these differences are due to the model or the study design: it is conceivable that the trajectories of the two species are not comparable, but it is also possible that the study did not run long enough to observe a decrease. The rat lifespan does not linearly map to the human lifespan. Each life stage is proportionately different lengths in human and rat lives (Quinn, 2005; Sengupta, 2013). For instance, a female rat at 17 months, the final time point in this study, is in the window for reproductive senescence. The same stage in human ageing has been shown to include age-related structural changes, even if they are not as exaggerated as in later life (Raz et al., 2005). Given the disproportionate

lengths of relative life stages, it is possible that the study's conclusion was pre-emptive and missed the window for a structural decline.

The observed changes in brain volume were found to differ between the groups. The EEDR rats had less grey matter volume at all time points after implementation of the intervention, and less white matter volume at the first two of the three post-intervention time points. Between the penultimate and last time point, the lack of difference observed between the groups is driven by an increase of the EEDR group's mean white matter to match the control group's mean, rather than a decline of the control white matter volume. There were no differences between the groups' CSF volumes at any age.

Each tissue volume was expressed as a fraction of total intracranial volume to infer the contribution of overall brain size to the observed changes. Despite differences in absolute volume, the groups did not differ in their overall ratios of tissue composition. However, there was an effect of age for grey matter and CSF where grey matter ratio generally increased over time while CSF volume declined. The lack of significant interactions between group and age in the ratio models suggest that the observed volume differences between the groups are due to the difference in total intracranial volume, rather than a change in tissue proportions. These findings will be explored further in chapter 5 (see: Exploring heterogeneity in the Resilient cohort).

Regional volume differences

Although global volume differences were likely due to differences in total intracranial volume, regional differences between the groups were present in all volumetric measurements after the EEDR implementation. Voxel-based and tensor-based morphometry were utilised to obtain complementary measurements of grey matter composition and changes in shape, respectively. VBM and TBM results often overlap; macroscale changes in tissue composition will likely involve corresponding changes in brain shape. Additionally, both share the capacity to either account for

overall brain size or not. Therefore, there are parallels between both GMV and absolute Jacobian determinant, which do not account for overall differences in brain size, and GMC and relative Jacobian determinants, which do. Despite these similarities, VBM and TBM are fundamentally independent of each other because they measure different phenomena. For example, if a change in grey matter is particularly subtle, such as a reduction of grey matter concentration but not in overall grey matter volume, the region in question may not change in its overall shape. Furthermore, the VBM analysis was constrained *a priori* to grey matter, whereas the TBM analyses use the Jacobian determinants generated from aligning the whole brain to a reference, and thus this analysis is completely data-driven and will include additional regions where shape change is caused by variation beyond grey matter. Therefore, it is highly advantageous to compare and contrast the findings from both related, but not equivalent, analysis methods.

Many of the volumetric results are consistent across multiple levels of comparison. The VBM and TBM metrics which do not account for overall brain size, GMV and absolute Jacobian determinants, were larger in the control group, which is in line with the global volumetric findings. Moreover, they both found that the cortex and thalamus were larger in controls compared to the EEDR group at 5 months, that the cortex, thalamus, and dorsal hippocampus were larger in the controls at 11 months, and that the caudal parietal cortex was higher in the EEDR group at 11 months. Regional differences GMC and the relative Jacobian determinants, which do account for overall brain size, also corresponded in some regions. For instance, both metrics found localised increases in the cingulate, basal forebrain & striatum for the EEDR group at 5 months old, and that the posterior parietal cortex and striatum were larger in the EEDR group at 11 months old.

Interestingly, the controls had more GMV and GMC in some cortical areas and the amygdala at 5 months, and in the cortex, amygdala, and medial prefrontal cortex at 11 months, but these did not have corresponding differences in the absolute or relative Jacobian determinants. Furthermore, there were evident increases in GMV

in the posterior parietal cortex for the EEDR groups at both 5 months and 11 months, which also did not have any corresponding differences in the Jacobian determinants, which suggests GMV increases in these regions are, at least partially, based on factors other than increases in total intracranial volume. Finally, both Jacobian determinants were increased in the cingulate cortex of the controls at 17 months; a finding which was not present in either the GMV or the GMC results. These shape changes may be due to a change in the tissue distribution in this area, or a change in the grey matter that is so subtle that it did not contribute to a large shape change between groups.

Regional microstructural markers

Interpretation of the volumetric measures is enhanced by the comparison of the relaxometry and diffusion metrics as markers of underlying microstructure. For the most part, these measures did not reach statistical significance, which may be indicative of the muted influences that any one of these values imparts on volumetric changes. However, despite low significance for inter-group differences, the location of the clusters continued to reflect anatomical boundaries, despite no *a priori* constraints to respect these boundaries.

Considering the relaxometry first, the EEDR group had longer T_1 times in the cortical grey matter, particularly near white matter boundaries between 5 and 11 months old, which was widespread by 17 months old. This finding is particularly interesting because it is the only metric where the difference in effect size between the groups grew over time, as many decreased or disappeared entirely by the final time point. Canonically, T_1 grey matter increases indicate increased water content in the tissue, but this was not identified by other measures which are sensitive to water content. However, T_1 is also sensitive to myelin content, so the controls' relative decrease in T_1 may indicate increased myelin content (Stüber et al., 2014).

The control group ostensibly had longer T_2 and T_2^* relaxation times at the 5-month time point, particularly in the ventricular systems, the brainstem white matter, and

near subcortical regions. However, at 11 months, the corpus callosum and brainstem white matter of the EEDR group had longer T_2 and T_2^* relaxation times than that of the controls. Lengthening white matter T_2 and T_2^* relaxation time could be due to more fluid in the white matter tract, which is a possible indicator of decreased myelination, but it could also signal the presence of iron depositions shortening the relaxation time in the control group (Bagnato et al., 2013).

As the corpus callosum demonstrated increased T_1 and T_2^* relaxation times in the EEDR rats at 11-month time point, both these metrics might be driven by the same underlying change in the microstructure. However, neither are specific to a particular microstructural change, which is problematic for interpretation. For example, although T_2^* relaxation time is canonically a measure of iron, it can also be altered by the presence of myelin; consequently, myelin is frequently measured by both T_1 and T_2^* (Cohen-Adad, 2014).

In addition to relaxometry, the diffusion metrics, derived from two of the most popular diffusion models, can be used to augment the interpretations of the microstructural changes. Diffusion tensor metrics FA and MD are the most commonly derived metrics; the former describes the overall directionality of the diffusion on a scale from isotropic to anisotropic, and the latter summarises the magnitude of diffusion across the diffusion tensor's three eigenvalues. In comparison, NODDI fits a model to represent three distinct biological compartments: intracellular, which has restricted diffusion, extracellular, which has hindered but not restricted diffusion, and isotropic diffusion. NODDI also produces a measurement of fibre angle alignment, which describes the coherency of fibre angles more accurately than FA, particularly at crossing fibres (Churchill et al., 2019). It has been previously shown that ODI and V_{ic} both contribute to FA (Timmers et al., 2016), so while these metrics are related, the inclusion of these metrics individually captures distinct features of diffusion, which should reduce the ambiguity of relating diffusion to microstructural changes.

For the DTI metrics, the EEDR rats mostly showed higher FA, particularly at 11 months, in key subcortical regions including the hippocampus, entorhinal cortex and the basal forebrain. Conversely, MD was more mixed at the first time point following intervention, with many dorsal regions showing higher MD in controls than EEDR, but more ventral regions, including the entorhinal cortex and subcortical grey matter showing increased MD in the EEDR group. The latter effect largely disappeared at the 11-month time point, with only increases in MD for the controls remaining in the same regions as previously. The NODDI metrics were mostly higher in controls relative to EEDR, and generally, the largest effect sizes were found at 5 months, which diminished by 11 months, before disappearing at the final time point. The exception appears to be ODI, which shows increased dispersion in the control group's superior colliculi at 5 months, but at 11 months, the EEDR group show more dispersion in the same region. Moreover, there is increased orientation dispersion in the dorsal hippocampus of the control group, which showed no difference at the 5-month time point.

Analogous to the relaxometry, diffusion metrics can be interpreted together to disentangle the underlying processes. For instance, coherent results from the 5-month time point show the control group had both higher ventricular MD at 5 and 11 months, and higher V_{iso} at 5 months, which is encouraging despite the lack of statistical significance. This particular result may be driven by the larger total intracranial volume of the controls, which would mean they had proportionately larger ventricles. The difference in overall brain size may also contribute to the increased MD, V_{ic} , and ODI observed at the dorsal cortical surface in the controls, particularly in the first two time points after intervention.

Furthermore, both DTI and NODDI metrics found a difference between the groups in the entorhinal cortex at 5 months: the control rats had increased FA, but the EEDR rats had higher V_{ic} , which also denotes more restricted diffusion. High FA values denote directionality, which implies that diffusion is restricted in all but the principal direction. However, the lower FA value of the EEDR group can be

interpreted as more equally restricted in all directions, which is analogous to the interpretation of the increased V_{ic} in the EEDR group.

The ventral hippocampus had higher V_{ic} in the controls at 5 months, but at 11 months, the controls had higher V_{iso} and the EEDR group showed higher ODI in the same region. Moreover, both dorsal and ventral parts of the hippocampus had higher FA in the EEDR group at the 11-month time point. The congruent increases of FA and ODI paired with the increased isotropic diffusion in the controls, suggests that there is less restrictive membrane integrity in the ventral hippocampus of the controls at 11 months, which appears to have occurred after 5 months.

In addition to comparing measurements between modalities, certain clusters also showed coherence to anatomical boundaries, such as the increase of V_{iso} in the control group's white matter at 11 months. The microstructural analyses were not anatomically constrained, so a cluster which follows anatomical boundaries suggests that despite a lack of statistical significance, this is less likely to be spurious.

There were also regions with congruent findings across relaxometry and diffusion, including the corpus callosum, which was previously discussed for having increased T_1 and T_2^* relaxation at 11 months in the EEDR group. Also, at 5 months, the control group had reduced V_{iso} in the corpus callosum, which also indicates that there is more directional diffusion in the controls than in the EEDR. Additionally, the superior colliculi had increased ODI and T_2^* in the EEDR group at 11 months. An increase in fanning and longer T_2^* relaxation may be an indication that the EEDR group has less myelination of their superior colliculi in comparison to the control group. This difference apparently occurs sometime between the 5 and 11 months, as ODI was increased in the controls at 5 months.

Coherence of volumetric and microstructural findings

Voxelwise results were generally widely distributed throughout the brain, and the microstructural findings were not as impactful in magnitude or statistical significance

as the volumetric changes. However, by combining all of the results, there are a few regions where changes were consistently observed across modalities.

In terms of cortical regions, the dorsal cortical surface, particularly the sensorimotor and parietal cortex, was repeatedly shown to have higher V_{ic} and more grey matter, both in terms of volume and concentration in the controls at 5 months. Controls also had more grey matter in these regions at 11 months. Moreover, it had a longer T_1 relaxation time in the EEDR rats, which also suggests less grey matter volume. However, the control rats had higher MD, which would suggest that there is more diffusion despite the increase in grey matter volume.

Furthermore, the thalamus generally had larger absolute Jacobian determinants and GMV in controls at 5 months and 11 months. Additionally, ODI was higher in the controls and higher FA in EEDR at 5 months. As the orientation dispersion only describes some of the variation in FA (Timmers et al., 2016), this is not inconceivable. The dorsal hippocampus also shows increased ODI in controls, and higher FA in the EEDR group at 11-months, which complements the observed higher GMV and absolute Jacobian determinants in the dorsal hippocampus of the controls, and larger relative Jacobian determinants in the EEDR group. Finally, the basal forebrain at the 5-month time point shows increased GMC and a larger relative Jacobian in the EEDR group and shows higher isotropic diffusion in the controls. At the 11-month time point, the EEDR group have higher relative Jacobian determinants and higher FA, which suggests that there is a change in the regional white matter between the groups.

Morphometric similarity networks

Morphometric similarity networks are structural connectomes which describe regional variance, using an array of imaging metrics derived from anatomical scans to create a distinctive characterisation of each region. Due to the large volume of structural images acquired, this method was hypothesised to be better at capturing the small, subtle changes expected with ageing and create a condensed measure that

allows the additive effects of small changes, and therefore could be better at finding age-related changes than any of the measures individually. In theory, the characterisation represents cytoarchitecture of the tissue, which was successfully linked to both molecular and cognitive measures in the original study, by predicting individual differences in cognitive abilities and emulate gene-expression patterns (Seidlitz et al., 2018).

Morphometric similarity networks are inherently dependent on the assumption that measuring covariance is symbolic of an underlying structural relationship. While there is evidence that spatially distinct regions are more cytoarchitecturally similar if they are physically connected (Goulas et al., 2017), one should be restrained in the conclusions drawn from the results. A structural connectome derived from covariance is fundamentally an indirect estimation of structural connectivity, in comparison to other structural connectomes which are derived from measurements of the underlying connections' integrity, with techniques such as tractography or tract-tracing (Fornito et al., 2013).

The results from mass univariate analysis with network-based statistics and graph theory analysis provide further evidence of a difference between the groups at 5 months and 11 months, which aligns with the findings from the voxelwise analyses. However, in terms of distinguishing between the groups, network-based statistics found more group differences compared to graph theory metrics. The only significant group difference from the graph theory metrics was that global efficiency was significantly lower in the EEDR group at 5 months, which may reflect the smaller volume of white matter in the EEDR group at that time point relative to the controls. Global graph theory metrics were the least successful of all the derived measures at delineating group differences, despite the presence of underlying structural differences. Moreover, when there were statistical differences, the estimates were relatively small, and one must be conscious about interpreting statistical differences as being biologically meaningful. While this might be a shortcoming of the analysis method, it could also mean that, despite the structural

differences, the connectomes of both groups are mostly similar. There were differences in graph theory metrics with age, and since all the animals were healthy, as opposed to a comparison to a group of animals that had received surgery, one might not expect to see large differences in the organisation of the connectome.

Multi-modal imaging: benefits and costs

Multiple metrics has consequently generated an abundance of information, but congruent findings at both the 5-month and 11-month time points are the most convincing. For instance, findings arise, across modalities, in the dorsal cortex, cingulate cortex, parietal cortex, medial prefrontal cortex, cingulate cortex, corpus callosum, thalamus, amygdala, striatum, basal forebrain, and hippocampus. However, the group differences were not always consistent across time points, which illustrates how variable cross-sectional findings can be in ageing research.

Using multiple structural modalities in this longitudinal study augmented the findings beyond the simple addition of each modality. By combining the metrics, it is possible to differentiate between the possible sources responsible for the observed changes. Moreover, with more information from independent sources, it lends support to complementary results that are consistent but do not pass the statistical threshold, which could be indicative of small but additive effects.

However, an increased volume of information can be difficult to summarise, which is a strength of an integrative method, such as morphometric similarity networks, that generates a single metric from multiple sources of information. Thus, this work has used both to probe the multiple components of age-related structural changes. In contradiction of the initial hypothesis, a condensed metric did not necessarily perform better at distinguishing between the groups, which could be because the compression of information to form the single metric resulted in a loss of information that was able to differentiate the groups. Future work can look at the weighted contributions of different modalities to optimise the minimal number of imaging protocols required to get the same or better results. Collecting an extensive

imaging protocol is costly in time and money, so it is beneficial to know which metrics are contributing the most to the regional characterisation.

Heterogeneity in structural metrics

Across all modalities and analysis methods, the largest differences between the groups were almost always found at 5 months. There were fewer significant differences at 11 months, and most differences between the groups entirely disappeared by 17 months, except for the increased T_1 in EEDR rats. This lack of difference may be impacted by models generally having a higher residual error with increasing age. While it is true that a reduction in group sizes increases the error of the measurement, it is also possible that potential ageing trajectories also diverge further with increasing age, given the accrual of influential lifestyle factors with increasing time, which would cause an increase in heterogeneity of values. Are the differences implemented by EEDR so short-lived that their effects have worn off by 17 months, or is the decreasing difference between the groups caused by survival bias? Longitudinal environmental enrichment has not been implemented for longer than one year (Kempermann et al., 2002b; Hüttenrauch et al., 2016), so the trajectory of environmental enrichment beyond 12 months is not known.

Qualitatively, the graph theory measures are more varied at the final time point for the EEDR group than the control group, giving credence to the hypothesis that the EEDR are a mix of healthy agers and successful agers at 17 months, while the remaining control group consists mostly of healthy agers. Moreover, it is unclear if and how differences in brain size relate to these findings, particularly the voxelwise microstructural metrics. Although differences in TIV has hampered interpretation of many of the other results, it is undesirable to control for the effect of brain size without truly understanding its mechanism of influence, particularly in models with small differences. Consequently, the relationship between total intracranial volume and non-imaging characteristics of healthy ageing will be explored in chapter 5. Next, consideration is given to how these observed structural group differences might manifest as differences in brain function.

4. VARIANCE AND COVARIANCE OF BRAIN FUNCTION IN HEALTHY AGEING

Rationale

This chapter describes the analysis of the functional imaging data collected during the Resilient study. The work aims to derive various metrics of global brain function, including summary measures of brain network topology, covariance within an established age-related network, and regional differences in cerebral blood flow (CBF), which will be compared between control and EEDR groups at all time points. The functional metrics derived from these modalities are theoretically analogous to those collected in human ageing studies but, as discussed previously, there are crucial implications for acquisition that can manipulate physiology, particularly the use of anaesthetics.

Functional imaging relies on the premise that neuronal activity is intricately linked with oxygen metabolism; oxidative phosphorylation is the primary source of energy for excitation of neurons and conduction of nerve impulses, and so the brain constantly demands oxygen, which is supplied by blood (Clarke & Sokoloff, 1999). This coupling of neuronal energy needs, oxygen consumption and alternations in regional blood flow present an opportunity to measure markers of neuronal activity changes in ageing networks.

Oxygen is carried in the blood by binding to haemoglobin. The structure, and consequently the magnetism, of haemoglobin is different in its oxygenated state compared to its deoxygenated state: oxygenated haemoglobin is diamagnetic, so shows weak repulsion to magnetic fields, while deoxygenated haemoglobin is paramagnetic, and therefore is weakly attracted to magnetic fields (Bren et al., 2015). Blood oxygen level dependence (BOLD; Ogawa et al., 1990) measures fluctuations in signal intensity with the assumption that such fluctuations are representative of changes in local blood oxygenation, driven by neuronal activity proximal to the observed fluctuation (Belliveau et al., 1991).

BOLD fMRI is commonly used to measure the response to tasks, where the activity is convolved with a haemodynamic response function to predict the pattern of fluctuation. A general linear model is fit at every voxel to find the regions where this pattern of activity is most similar to the predicted response. However, task-based fMRI in anaesthetised rodents relies on involuntary responses which occur in unconscious animals, such as sensory stimulation (Hyder et al., 1994). Consequently, resting-state fMRI, where no task is present, is preferable. Resting-state fMRI requires the application of a reversed approach: all possible sources of non-BOLD signal (i.e., noise) are removed, so only time courses that are ostensibly related to neuronal activity remain (Murphy et al., 2013). These time courses are frequently correlated with each other under the assumption that spatially distinct regions that share spontaneous fluctuations are connected, hence deriving so-called functional networks which have been shown to vary in their connectivity with increasing age.

Of these networks, the most widely reported in studies concerning age in humans is the default mode network (DMN), which consists of the medial prefrontal, medial parietal, lateral parietal, and medial temporal cortices (Greicius et al., 2003; Damoiseaux et al., 2006; Raichle, 2015). The connectivity within the DMN has been repeatedly shown to decline with increasing age (Andrews-Hanna et al., 2007; Damoiseaux et al., 2008; Sun et al., 2016), and reduced functional connectivity is more likely in those with poorer cognition compared to people of the same age (Jones et al., 2011; Fjell et al., 2013; Sun et al., 2016; Ash et al., 2016).

Resting-state networks, including the DMN, can be derived from fMRI of other animals, including rodents (Lu et al., 2012; Gozzi & Schwarz, 2016). An in-depth examination of the rat DMN by Hsu et al. (2016) showed decreased average functional connectivity between the constituent nodes of the DMN with age, and that the degree of decline also varies between cognitive outcomes. Rats that showed spatial memory similar to young rats on a Morris Water Maze test of spatial memory had smaller decreases in average functional connectivity than rats that were impaired on the task relative to young counterparts.

For this study, it was hypothesised that there would be decreased connectivity within the DMN with increasing age, based on the findings in human and rat ageing studies. As there is a lack of consensus regarding age-related changes of other resting-state networks, graph-theoretical measures were derived from examining both brain-wide covariances in regional brain function and the covariance within the DMN nodes over time for both groups. The covariance of regional brain function was derived from voxelwise BOLD time courses using multi-echo resting-state fMRI.

Traditionally, BOLD time courses are derived from echo-planar imaging (EPI) volumes which use a single gradient to rephase the spinning nuclei before acquiring the image to derive BOLD time courses. The choice of echo time (TE), at which point the nuclei rephasing gradient is given, and repetition time (TR), which is the time between successive acquisitions, are decisive for the image's signal intensity and contrast weighting. For instance, T_1 -weighting is given at short TE and TRs, when both the signal is maximised, and T_1 relaxation is minimal. Conversely, a long TE and a long TR gives a T_2 -weighted image, where the signal is relatively lower than a T_1 -weighted image, but the differences in T_2 relaxation between tissues are emphasised. However, the time to reach T_2 relaxation can be accelerated by local differences in magnetic susceptibility, which is represented by T_2^* , and so T_2 weighted sequences often suffer from a reduced signal, particularly in regions with large differences in magnetic susceptibility. The delicate balance between contrast and signal can be observed in figure 1, which displays an image which was acquired with a sequence that uses multiple echo times following a single excitation. The shortest echo, which was acquired at a TE of 11 milliseconds, is almost T_1 -weighted as it has high signal in all regions in the field-of-view, but it is difficult to differentiate between tissues, especially in regions with the highest signal. On the other hand, the longest TE was acquired at 35 milliseconds, and the signal in regions with fast T_2 times is lost entirely because the effects of excitation have subsided before the acquisition.

Figure 4-1: Change in contrast with increasing echo time in a raw image. In the echo with 11 ms echo, the signal is highest, but there is little contrast difference between tissues. Comparatively, at 35 ms there is better contrast between tissue types, but the signal is low in many regions.

A single optimal TE is consequently elusive; the TE which maximises the available signal while distinguishing between changes in magnetic susceptibility is contingent on the local environment, and so changes on a voxelwise basis. However, by combining the information from multiple echoes, linear functions can be fit at each voxel which describe the change in local signal and T_2^* of each voxel with increasing echo time. With this information, the signal is maximised, and regions which usually suffer from drop-out benefit from more statistical power (Kundu et al., 2017).

However, BOLD is an indirect measure; the values in the time course represent the signal value, which itself is not representative of neuronal activity. Neuronal activity is coupled to changing blood oxygenation, but the BOLD signal itself is not a meaningful number because of its relative nature, i.e., the individual values are relative and therefore difficult to compare across individuals without reference.

Moreover, while the BOLD signal, in theory, is driven solely by neuronal activity, noise is widespread through the raw fMRI image (Liu, 2016) and anaesthetics can

also modulate neurovascular coupling to influence the BOLD response (Aksenov et al., 2015). Consequently, it can be ambiguous whether a finding is related to the true BOLD response, an anaesthetic-related modulation of cerebral vascularity, or noise. Multi-echo has previously been leveraged to disentangle the signals of non-BOLD and BOLD related signals in rats with multi-echo independent component analysis (Kundu et al., 2014). As the T_2^* signal is known to decrease linearly with the echo time, components that cannot be classified as being driven by non-BOLD signals, such as thermal noise, and rejected (Kundu et al., 2017). However, multi-echo denoising methods are only capable of disentangling BOLD and non-BOLD signals and are therefore incapable of disentangling anaesthetic-related modulation of cerebral vascularity. Therefore, it is still necessary to assess the quality of the data, for instance by using independent component analysis to replicate previously reported resting-state networks (Gozzi & Schwarz, 2016), as illustrated in figure 4-2.

Figure 4-2: Example component from group ICA showing DMN-like component from resting-state, despite isoflurane anaesthesia. The top three rows show the component in full, overlaid on a mean EPI volume, while the bottom row shows slices of the component overlaid on a resampled template for anatomical alignment.

In comparison, CBF is a quantitative measure which directly estimates the volume of blood that perfuses through a given region of the brain in a given time frame. While blood flow is also affected by anaesthetics, such as isoflurane which is a vasodilator (Matta et al., 1999), this quantitative measure is directly comparable between individuals, and between studies. In the Resilient study, CBF was derived from arterial spin labelling, a method which ‘tags’ the hydrogen nuclei in arterial blood with an excitation pulse, so that blood enters the brain as an endogenous marker. An image is acquired after a short interval (known as the ‘arterial transit time’), and the difference in signal intensity from a control image which received no excitation pulse is proportional to the CBF (Alsop et al., 2015).

As discussed previously (see: Age-related changes to brain function), a negative correlation between global CBF and age has been previously observed in ageing humans (Restom et al., 2007; Ances et al., 2009; Chen et al., 2011a; Steffener et al., 2013). In ageing rodents, resting CBF is less frequently quantified. There have been mixed findings between studies of ageing mice: one longitudinal and one cross-sectional study have found no effect of age on regional CBF (Maier et al., 2014; Hirschler et al., 2018), while another longitudinal study reported an increase (Wei et al., 2020). Meanwhile, studies in rats have only calculated the relative change of CBF in response to stimuli rather than quantifying resting perfusion rate. For example, aged rats have been shown to have demonstrably reduced CBF changes in reactions to hypercapnia, indicating reduced vascular reactivity with increasing age (Desjardins et al., 2014). Thus, ageing perfusion rates, as estimated with MRI, are not well characterised in rats. However, CBF has been estimated using invasive measures in ageing rats, which do not require anaesthesia, and CBF was found to have differential lifetime trajectories depending on the region. Generally, regions maintain constant CBF between 3 months and a year old, although frontal regions show increases. However, after 12 months, CBF declines in several regions, including posterior regions (Ohata et al., 1981). Moreover, this study was cross-sectional because of the invasive method required for measurement. With the use of ASL to non-invasively

measure CBF, individual differences in CBF may come to light in a longitudinal design (see: The importance of longitudinal designs for ageing studies). Thus, the final aim of this analysis was to measure CBF in a collection of cortical and subcortical regions and compare them between groups and across time. It was hypothesised that the EEDR rats would be less prone to age-related decreases in regional CBF compared to controls.

Methods

Functional atlas parcellation

The images were analysed in native space to avoid interpolation of the CBF values, which were derived from images with relatively larger voxel sizes and fewer slices than most reference atlases. A native space atlas was generated for each subject by projecting the reference atlas to each CBF map, but downsampling and transforming a granular atlas with small regions of interest (ROIs) may cause heterogeneous loss of ROIs across individuals. Thus, a functional atlas with a coarse parcellation was generated by combining functionally related ROIs to make them larger and more robust to spatial transformation. Moreover, as a consequence of each ROI containing more voxels than a more refined parcellation, more time courses are used to calculate the mean, which is also more robust than ROIs with fewer voxels.

Two popular atlases had been previously combined in-house, due to their complementary delineation of cortical and subcortical features respectively (Valdés-Hernández et al., 2011; Papp et al., 2014), and it was this combined atlas that was used to generate a new functional atlas specifically for this study. ROIs in the combined atlas were assigned as either grey matter, white matter, and CSF. Only grey matter ROIs were contributed to the functional atlas, which is visible in figure 4-3, and each ROI was assigned as either cortical or subcortical and further classified as one of 22 anatomical groups (appendix 4-1).

Figure 4-3: Coronal slices of functional atlas used for analysis

Image acquisition and pre-processing

Multi-echo echo-planar imaging

Resting-state fMRI was acquired using a 2D multi-slice multi-echo EPI sequence with TR of 2750 ms, TEs of 11, 19, 27, and 35 ms, and a 70 ° flip angle, producing an image with 40 slices. Slices were 0.5mm thick with a 0.2 mm gap, which gave a 48 x 44 matrix, with an in-plane resolution of 0.5 x 0.5 mm.

The echo volumes were split with `nii_tool` from `dicm2nii` (Li, 2020) with MATLAB (version 9.2.0, MathWorks). Variability in slice timing was corrected for each echo separately using the Analysis of Functional NeuroImages (AFNI; Cox, 1996; version 17.2.17) tool `3dTshift`. The first echo was motion-corrected to its middle volume using AFNI `3dvolreg`, and the displacement and rotation parameters for each volume were output to a text file. The remaining echoes were motion-corrected using the matrix transformation generated by the motion correction of the first echo in AFNI's `3dAllineate` tool.

The motion-corrected first echo was averaged across time and co-registered to the subject's T₁-weighted image (T₁w; see 3.2.1.1 T₁-weighted image and T₁ mapping) by implementing a rigid body registration with `antsRegistration` (Avants et al., 2011a). The inverse affine transform was applied to the mask derived from the T₁w image, thereby providing a mask for the EPI volume.

The optimal combination of the echoes was calculated using the `t2smap` tool from TE-Dependent ANAlysis (tedana; DuPre et al., 2020). A non-linear monoexponential decay function was fit at each voxel to estimate the maximal signal, which is the model's intercept, and T₂^{*}, which is derived from the slope of the model. The echoes were combined using the method recommended by Posse et al. (1999), where each echo is assigned a weight using the equation:

$$\omega_{TE} = TE \times e^{\frac{-TE}{T_2^*}}$$

The weights are normalised across echoes and a weighted average of the echoes is calculated, which benefits from the shortest echo's available signal and the longer echoes' sensitivity to magnetic susceptibility. A native-space probability map of CSF was created previously from automatic segmentation of the subject's T₁w image using ANTs Atropos (see Age-specific population templates and probability maps). The probability map was projected to the resting-state space using the same inverse transform that was used to project the mask to the EPI volume, and it was binarised above 99 % probability to create a conservative mask that only contained regions that were highly likely to represent only CSF. The mean time series from within this EPI-space CSF mask was extracted from the optimally combined volume using FSL's `fslmeans`.

The optimally combined volume was smoothed with AFNI's `3dBlurInMask` with a kernel full-width half maximum of 1 mm, which is double the in-plane resolution. Finally, the nuisance time courses, i.e., the mean CSF time course and the realignment parameters from motion correction, and their first and second-order derivatives were simultaneously regressed from the volume while applying a bandpass filter of 0.01 - 0.1 Hz using AFNI's `3dTproject`.

Continuous arterial spin labelling

CBF maps were derived from a 2D multi-slice spin-echo EPI sequence with TR of 4000 ms, TE of 13.6 ms, and a 90 ° flip angle, producing an image with 18 slices. Slices were 1mm thick with a 0.5mm gap, which gave a 96 x 96 matrix, with an in-plane resolution of 0.260 mm isotropic. EPI volumes were collected sequentially: after an arterial transit time of 3000 ms, the first post-label delay was 30 ms, and the remaining slices were collected every 4 ms after. An additional proton density-weighted image was acquired at least 10 seconds after the EPI sequence concluded with identical parameters except there was no labelling radiofrequency pulse. Finally, an image that was identical to the proton density-weighted (PDw) image, except with the phase-encoding direction reversed, was acquired.

The susceptibility distortions for the PDw image were estimated and corrected using FSL `topup` (Andersson et al., 2003) and the reverse phase-encoded image. This correction was subsequently applied to the CASL volume.

CBF maps, as pictured in figure 4-4, were generated using QUIT's `qi_as1` (Wood, 2018). Briefly, CBF is quantified from the signal intensity of the pairs of distortion-corrected label and control images, and the distortion-corrected PDw EPI image using the equation:

$$CBF = \frac{6000 \cdot \lambda \cdot (SI_{control} - SI_{label}) \cdot e^{PLD/T_{1,blood}}}{2 \cdot \alpha \cdot T_{1,blood} \cdot SI_{PD} \cdot (1 - e^{-\tau/T_{1,blood}})}$$

where 6000 is required to convert the units to $\text{ml } 100 \text{ g}^{-1} \text{ min}^{-1}$, λ is the blood-brain partition coefficient that was set to 0.9 mL/g , SI is the respective time-averaged signal intensities from control, label, and proton density images, $T_{1,blood}$ is the T_1 relaxation time of blood, which is given as 2.429 seconds at 9.4 T (Dobre et al., 2007), α is the labelling efficiency of the radiofrequency pulse which was assumed to be 0.85, and τ is the label duration, as recommended by Alsop et al. (2015).

Figure 4-4: Three coronal slices from a representative cerebral blood flow map

Image analysis

For each subject, the mean of the resting-state volume along the time dimension and the CBF map was aligned to their respective T_1w image with a rigid body co-registration. By concatenating the respective co-registration transformation matrix with those from linear and non-linear transformations executed previously (see Age-

specific population templates and probability maps), the filtered resting-state volume was projected to the template space, and the functional atlas was projected to the native space of each CBF map.

BOLD-derived global and DMN functional connectivity

The mean time course for each ROI was extracted from the filtered resting-state volume using FSL's `fslmeans`, and a correlation coefficient matrix was generated for each subject by pairwise correlation of the mean time courses. Fisher's transformation was used to normalise the distribution of the correlation coefficients.

In the previous chapter, morphometric similarity graphs were thresholded to account for spurious correlations (see: Graph thresholding), which occur as a consequence of a large number of comparisons in a covariance matrix and can impact the topology of the network. Commonly, one threshold is applied across all graphs for consistency, because the choice of threshold impacts the graph density and many derived features vary as a function of graph density due to network fragmentation (Fornito et al., 2013). However, choosing one single threshold across all time points for functional networks derived from BOLD time courses proved challenging. Thresholding sets a cost limit, below which 'connections' are more likely to be spurious. This limit can be set absolutely (e.g., connectivity $Z > 1$), but this approach changes the number of edges in a graph for each individual, which is a standard metric of normalisation in the graphs. In the previous chapter, the limit was set proportionately, (e.g., connectivity in the strongest 5 % of edges), which keeps the number of edges the same. However, proportional thresholding is unsuitable when comparing graphs across groups that are likely to vary as a function of overall connectivity (van den Heuvel et al., 2017). As an example, consider two groups that differ in their magnitude of overall connectivity (i.e., a control group and a patient group); if the overall average functional connectivity is higher in controls than in patients, then the patient group's lower boundary of connections that surpass the threshold are more comparable in strength with spurious connections than the same comparison in the control group. Thus, setting a proportional threshold could

include ‘spurious’ connections in the patient group, and their inclusion would artificially impact the network topology. Since overall functional connectivity has been shown to decrease with increasing age, in humans and rats (Andrews-Hanna et al., 2007; Hsu et al., 2016), setting a single threshold for all time points would be inappropriate.

Moreover, the metrics that were used to define the threshold in the previous chapter are known to vary with age when derived from resting-state BOLD time courses. The first metric that the threshold was contingent on was cost-efficiency, which offsets the global efficiency by using the density of the network as a metric of cost (Bassett et al., 2009). However, the relationship between efficiency and cost has been shown to uncouple in older humans, with efficiency reducing disproportionately in relation to cost (Achard & Bullmore, 2007). The second metric used to determine the threshold was small-world propensity, and although the effect of age on this metric does not appear to have been measured previously, it is also partially derived from global efficiency.

Consequently, an alternative was employed to avoid necessitating a unified threshold: integrating network features over a range of costs (Ginestet et al., 2011). Briefly, graphs were thresholded based on the strongest percentage of connections at 1 % intervals from 1 to 5 % and 5 % intervals from 10 to 50 %. The minimum threshold value was calculated for each graph sparsity interval, and values which did not meet the threshold were set to zero, resulting in weighted undirected functional covariance networks for each subject.

Previously described global graph measures, including small-world propensity (ϕ) and global efficiency (E_{glob}), were derived from the thresholded graph at each level of graph sparsity in MATLAB, using code from the Brain Connectivity Toolbox (Rubinov & Sporns, 2010) and the small-world propensity (Muldoon et al., 2016) toolboxes. Additionally, the overall functional connectivity (van den Heuvel et al., 2017), whereby the overall functional connectivity is taken as the mean of the absolute non-zero values from all elements in the matrix, and the maximum

modularity (Q) was derived from an approach implemented by Seidlitz et al. (2018). In this implementation, the Louvain algorithm (Blondel et al., 2008) was applied to the graph 1000 times to generate a stable consensus matrix, which is then decomposed using the Louvain algorithm to determine Q . The area under the curve (AUC) was calculated for each group over the range of graph costs using trapezoidal numerical integration. The difference in AUC was tested with a non-parametric permutation test, whereby the group labels at each time point are interchanged 5000 times to generate a null distribution of possible differences in AUC, against which the observed difference is compared. P-values are calculated as the count where the absolute value of the random distribution is equal to or greater to the absolute value of the observed difference, divided by the number of permutations.

Figure 4-5: Regions of the functional atlas corresponding to the default mode network

Time courses from the functional atlas regions representing nodes of the rat DMN (shown in figure 4-5), including the auditory, prefrontal, cingulate, parietal, retrosplenial, visual cortices, and the dorsal and ventral hippocampus, were correlated to make a 16 x 16 graph to examine DMN connectivity between the groups. These graphs were not thresholded, due to the smaller number of edges present in the graph and the *a priori* assumption that they are connected and would therefore not be spurious. The overall functional connectivity of the DMN was calculated for each subject, and these values were entered into a conditional growth linear mixed-effects model using `lme4` and `lmerTest`. The overall functional

connectivity of the DMN was compared across time and between groups, before *post-hoc* estimation of the marginal means and contrast of the estimates using *emmeans*. Exact p-values were adjusted for multiple comparisons using a multivariate t-distribution.

Regional cerebral blood flow

Due to a combination of noise from image distortion or field inhomogeneity and low flow, voxels with irrational values for CBF values can be present in the maps, particularly in the ventral part of the brain which is furthest from the receiver coil. As excluding these values would artificially inflate the CBF in regions with genuinely low flow, the median CBF value was calculated for each region in the functional atlas using *labelStats* in ANTsR. Three images were excluded from the analysis due to the generation of corrupted CBF maps, one from the third time point (one control), and two from the final time point (one EEDR and one control).

A conditional growth model was fit for the CBF measurements using *lme4* and *lmerTest*. The rate of perfusion across factors was compared using the EMMs, initially with a joint test and with post-hoc contrasts, generated by the R package, *emmeans*, with p-values adjusted using False Discovery Rate correction.

Results

Global metrics of functional connectivity

The AUC for each metric was derived from the group mean metric at each graph cost (Appendix 4-1). Before intervention, the group differences in AUC for small-world propensity ($\Delta\text{AUC} = -0.5208$, $p = 0.1610$) or the mean maximum modularity ($\Delta\text{AUC} = -0.8457$; $p = 0.1090$) were not significant. By chance, the control group had a higher average functional connectivity (Appendix 4-2A) and global efficiency (Appendix 4-2c), with the observed difference between the group AUC passing significance criteria (functional connectivity $\Delta\text{AUC} = 6.827$; $E_{\text{glob}} \Delta\text{AUC} = 1.478$; $p < 0.0002$ for both metrics) but the groups' measurements were continuously less than one standard deviation apart (Appendix 4-2 B & D).

At the first time point after EEDR implementation, when the rats were 5 months old, EEDR functional connectivity was higher compared to controls ($\Delta\text{AUC} = -1.0638$, $p = 0.0422$). Global efficiency was also higher in the EEDR ($\Delta\text{AUC} = -0.4206$, $p = 0.1324$), but the latter did not surpass significance criteria. These differences appear to be driven by the more leniently thresholded matrices, which contain more connections because the sparser matrices with the strongest correlations were mostly very similar (figure 4-6A and C). The sparsest graphs (i.e., those with less than 5 % density) contain a large degree of variance for both the small-world propensity (figure 4-6E) and the average maximum modularity (figure 4-6G), as indicated by the large standard deviations for these threshold densities. Average maximum modularity is generally higher in the control rats ($\Delta\text{AUC} = 1.734$; $p = 0.0014$), particularly in thresholds that include more than 5 % of the strongest connections. Except for the 25 % threshold, these values appear to be reasonably stable across graphs with more than 5 % of the strongest connections. Small-world propensity decreases with thresholds that include more than 5 % of the strongest connections and decreases appear to be reasonably well-matched between the groups with no significant difference in the AUCs ($\Delta\text{AUC} = 0.160$; $p = 0.8094$).

Figure 4-6: Difference (Control – EEDR) in the area under the curve for functional connectivity metrics at 5 months old. A) functional connectivity, C) global efficiency, E) small-world propensity, and G) maximum modularity with error bars representing one standard deviation. Null distributions of B) overall functional connectivity, D) global efficiency, F) small-worldness propensity, and H) maximum modularity with a red line at the observed difference in AUC.

At 11 months old, the controls had higher overall functional connectivity (figure 4-7A; $\Delta\text{AUC} = 1.985$; $p = 0.0124$), and higher global efficiency (figure 4-7C; $\Delta\text{AUC} = 0.772$; $p = 0.0216$). While the trend appears to be the same for global efficiency, in that the sparser graphs are more similar between the two groups with larger differences in the graphs which contain weaker connections, this is not true for functional connectivity with a difference of about the same magnitude between the controls and EEDR rats across all graph densities. Small-worldness ($\Delta\text{AUC} = 0.8299$; $p = 0.6008$) appears to be largely matched in both groups across the range of thresholds, as both show a decline in small-world propensity with increasing graph density (figure 4-7e). Graph modularity (figure 4-7G; $\Delta\text{AUC} = -0.4197$; $p = 0.5390$) appears to be particularly stable across graphs that contain at least 10 % of the strongest connections, with control rats showing higher modularity in middle-density graphs (5-20 % sparsity), while the EEDR rats have higher modularity when more weak connections are included (25-50 % sparsity).

When the rats were 17 months old, the controls once again showed higher overall functional connectivity ($\Delta\text{AUC} = 5.3610$; $p < 0.0002$) and average global efficiency ($\Delta\text{AUC} = 2.3653$; $p < 0.0002$). The magnitude of difference between the curves in the groups seems roughly the same across all densities for overall functional connectivity (figure 4-8A). However, the difference in global efficiency between the groups grows larger with increasing graph density (figure 4-8C). In comparison to the previous time point, figure 4-8G shows that while control rats have a greater AUC overall, the EEDR rats appear to have slightly higher modularity in graphs with densities between 5 and 20 %, while this is reversed for the graphs with 20-50 % density. Thus, the difference is not statistically significant ($\Delta\text{AUC} = 1.0312$; $p = 0.1412$). Finally, the largest difference between the groups' average small-world propensity occurs at 10 % sparsity (figure 4-8e), but each mean is still within one standard deviation of the other group, suggesting that the overall small differences in small-worldness do not differ between groups ($\Delta\text{AUC} = 0.1885$; $p = 0.8160$).

Figure 4-7: Difference (Control – EEDR) in the area under the curve for functional connectivity metrics at 11 months old. A) functional connectivity, C) global efficiency, E) small-world propensity, and G) maximum modularity with error bars representing one standard deviation. Null distributions of B) overall functional connectivity, D) global efficiency, F) small-worldness propensity, and H) maximum modularity with a red line at the observed difference in AUC.

Figure 4-8: Difference (Control – EEDR) in the area under the curve for functional connectivity metrics at 17 months old. A) functional connectivity, C) global efficiency, E) small-world propensity, and G) maximum modularity with error bars representing one standard deviation. Null distributions of B) overall functional connectivity, D) global efficiency, F) small-worldness propensity, and H) maximum modularity with a red line at the observed difference in AUC.

Overall functional connectivity of DMN

Figure 4-9: Default mode functional connectivity adjacency matrices. A) 3 months, B) 5 months, C) 11 months and D) 17 months. The upper triangle of each matrix represents the control group mean functional connectivity for each edge, while the lower triangle represents the EEDR means.

As the functional connectivity of the DMN has been repeatedly reported to decline with increasing age, the ROIs that comprise the default mode were isolated from the overall graph for further inspection. In figure 4-9, each graph represents the covariance of the BOLD time course at each time point, with the upper triangle representing the control group and the lower triangle representing the EEDR group. While there are some ostensible differences between groups which require further attention, the most apparent difference is the decline in the DMN covariance with

increasing age. Additionally, the ventral hippocampus becomes more anti-correlated with the posterior cortical regions by the 17-month time point across groups.

To examine the difference in overall functional connectivity within the DMN, the subject-wise DMN overall functional connectivity, which has been plot across time in figure 4-10, was fit in a linear mixed-effects model with fixed effect predictors including group and age, and the interaction between them. Contrary to previous models, this was a conditional growth model with a fixed slope, due to the small number of observations and the otherwise high number of random effects. Thus, the model took the form:

$$\text{Functional connectivity}_{ij} = (\beta_0 + u_{0ij}) + \beta_1 \text{group}_{ij} + \beta_2 \text{age}_{ij} + \beta_3 \text{group}_{ij} \text{age}_{ij} + \varepsilon_{ij}$$

where with i and j refer to the i th subject in the j th group.

The DMN functional connectivity for the control group at baseline was estimated as $Z = 0.746 \pm 0.038$. The estimate decreased significantly by 0.235 ± 0.05 between baseline and 5 months, 0.304 ± 0.05 between baseline and 11 months, and 0.329 ± 0.067 between baseline and 17 months. The EEDR group's functional connectivity at baseline was estimated as 0.108 ± 0.053 lower than the control rats. There was a significant interaction between the groups at the 5-month time point ($\beta = 0.146 \pm 0.034$) indicating that the EEDR group had a smaller difference between their baseline and 5-month estimates relative to the same difference in the control rats. The other interactions were not significant; thus, the EEDR group had decreases of a similar magnitude between baseline and 11-months, and between baseline and 17-months as the decreases seen in the controls.

Consecutive contrasts of the estimated marginal means show DMN overall functional connectivity declines differentially between the groups. In the control group, the contrast between 5 and 3 months is statistically significant (contrast estimate = -0.235 ± 0.048 , $p < 0.0001$), but the estimates between 11 and 5 months (contrast estimate = -0.069 ± 0.054 , $p = 0.4664$), or 17 and 11 months (contrast estimate = -0.025 ± 0.070 , $p = 0.9736$) are not. Conversely, the difference between

5 and 3 months (contrast estimate = -0.090 ± 0.048 , $p = 0.1601$), and between 17- and 11-months (contrast estimate = -0.044 ± 0.050 , $p = 0.7139$) is not statistically significant in the EEDR group, but it is significantly different between 11 and 5 months (contrast estimate = -0.135 ± 0.048 , $p = 0.0165$). Therefore, the largest decline in DMN connectivity is after 5 months in the EEDR group, but after 3 months in the control group.

Figure 4-10: Overall functional connectivity of the default mode network. The group trajectories for control (green) and EEDR rats (purple) are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

	β	Std. Error	df	T value	P-value
Intercept: difference from 0					
Control at 3 months	0.74632	0.03788	145.445	19.704	<0.0001
Main effects: difference from controls at 3 months					
EEDR at 3 months	-0.10796	0.05303	144.808	-2.036	0.0436
Control at 5 months	-0.23515	0.04818	111.917	-4.880	<0.0001
Control at 11 months	-0.30377	0.05461	123.458	-5.562	<0.0001
Control at 17 months	-0.32855	0.06675	129.740	-4.922	<0.0001
Interaction effects: difference from control difference between 3 months and time point					
EEDR difference in 3 to 5 months	0.14568	0.06773	111.334	2.151	0.0336
EEDR difference in 3 to 11 months	0.07897	0.07283	118.410	1.084	0.2804
EEDR difference in 3 to 17 months	0.05950	0.08309	124.157	0.716	0.4753

Table 4-1: Estimates of the DMN connectivity model from group and age

Cerebral blood flow

Figure 4-11: Median regional perfusion rates of cerebral blood flow. A) median perfusion rate of cerebral blood flow (measured in $\text{ml } 100\text{g of tissue}^{-1} \text{ minute}^{-1}$) for each region in the atlas. Each column represents a different time point (3, 5, 11, and 17 months), and sub-columns for experimental group. B) Variance of median CBF values in all rats for the regions that constitute the DMN. Plots are given per time point with the median value delineated by a black solid line.

As seen in figure 4-11A, median CBF is variable across regions and time, while the variability within default mode network regions across time is shown in figure 4-11B. To assess these apparent differences, a mixed-effects model, with random slopes to account for subject-wise differences in trajectories, were fit to predict median CBF using age, ROI, and experimental group as fixed effects predictors, along with the interactions between the three predictors (see Table 4-2 and Appendix 4-2). The model had an intercept of $94.021 \pm 4.928 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$ which represents the median CBF in the visual cortex at the 3-month baseline scan for the control rats.

Model term	Joint test		By group: EEDR		By group: Control	
	<i>F ratio</i>	<i>P-value</i>	<i>F ratio</i>	<i>P-value</i>	<i>F ratio</i>	<i>P-value</i>
Main effects						
ROI	382.306	< 0.0001	145.042	< 0.0001	145.042	< 0.0001
Group	0.431	0.5114				
Age	5.556	0.0008	0.415	0.7422	6.118	0.004
Two-way interactions						
ROI and group	6.577	< 0.001				
ROI and age	49.640	< 0.0001	31.265	< 0.0001	23.857	< 0.0001
Group and age	2.661	0.0464				
Three-way interaction						
ROI, group, age	3.135	< 0.0001				

Table 4-2: CBF conditional growth model contrasts of factors and interactions

As illustrated by the values in Table 4-2, ROI was the factor which explained the most variance in median CBF. Moreover, as seen in appendix 4-2, the perfusion of the visual cortex was significantly different from multiple other ROIs at baseline: notably, the rostral cortex (cingulate, frontal, auditory, piriform, prefrontal, olfactory and insular cortices), and some subcortical regions including the ventral hippocampus, hypothalamus, globus pallidus, substantia nigra and the colliculi. It was not significantly different from the caudal cortices (the parietal, retrosplenial, motor, sensory, or rhinal cortices). In terms of subcortical regions, the median perfusion rate did not differ from that of the basal forebrain, dorsal hippocampus, thalamus, or striatum.

Although there was no significant main effect of group in the model, the ROI by group interaction was significant. The handful of areas that showed statistically different median perfusion rates between the groups at baseline were the frontal cortex and the olfactory bulb, which had lower perfusion in the EEDR group, and the hypothalamus and the colliculi, which had higher perfusion rates in the EEDR group.

There was a significant main effect of age showing that median CBF in the control group changes over time; there was a decrease in the control rats at the 5-month scan, but it was not statistically significant. In comparison, the median rate of blood

flow increased relative to baseline at the 11-month time point and at the 17-month time point.

There was a significant interaction effect between age and ROI, indicating that the difference between time points was not homogenous across ROIs in the control rats. Most ROIs had at least one significant difference over time, but the difference from baseline was not the same as that of the corresponding difference of the visual cortex at any subsequent time point for the piriform and prefrontal cortices, the globus pallidus, the substantia nigra or the colliculi. Conversely, the difference in perfusion from the baseline of the ventral hippocampus was not significantly different from that of the visual cortex at any time point.

The interaction between group and age was significant, showing that perfusion rate differences from baseline were distinct between the groups. Using a post-hoc comparison of groups by age shows that the control group has a higher perfusion rate across all ROIs, but only at 17 months ($\Delta\text{CBF} = 19.41 \pm 8.05 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$, $p = 0.0159$). At the 3-month baseline ($\Delta\text{CBF} = -3.76 \pm 6.66 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$, $p = 0.5730$), 5-month ($\Delta\text{CBF} = -7.22 \pm 7.67 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$, $p = 0.3337$), and 11-month ($\Delta\text{CBF} = 4.01 \pm 7.67 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$, $p = 0.6013$) time points, there is no difference between the groups when averaging perfusion across ROIs. The interaction between age and group is also observed when separating the factor contrasts by groups, as in Table 4-2. The main effect of age is not significant in the EEDR group, but it is for the control group.

Finally, the interaction between the three factors was also significant. To test which time points were different between the groups across all ROIs, estimated marginal mean contrasts were set up comparing groups pairwise and age consecutively by each ROI. The ROIs that had significant results are available in Table 4-3 (with all other ROIs available in appendix 4-2). Except for the substantia nigra, the regional mean difference in perfusion rate between the penultimate and final time points was between 20 and 30 $\text{ml } 100\text{g}^{-1} \text{ min}^{-1}$ higher in the control group than the same

difference for the EEDR group for these particular regions. The substantia nigra, in comparison, had a $24.76 \pm 9.63 \text{ ml } 100\text{g}^{-1} \text{ min}^{-1}$ increase in perfusion between 5 months and 11 months in the controls compared to the EEDR.

ROI	Contrast (control – EEDR)	β	Std. Error	P-value
substantia nigra	5 months - 3 months	-14.3388	10.07	0.2318
	11 months - 5 months	24.7639	9.63	0.0303
	17 months - 11 months	8.3825	8.37	0.3163
thalamus	5 months - 3 months	-2.2196	10.07	0.8256
	11 months - 5 months	7.3618	9.63	0.6668
	17 months - 11 months	20.8191	8.37	0.0385
dorsal hippocampus	5 months - 3 months	-2.9771	10.07	0.7675
	11 months - 5 months	10.7894	9.63	0.3937
	17 months - 11 months	20.2244	8.37	0.0469
ventral hippocampus	5 months - 3 months	-7.9708	10.07	0.4287
	11 months - 5 months	13.5141	9.63	0.2407
	17 months - 11 months	21.2517	8.37	0.0332
prefrontal	5 months - 3 months	-1.6757	10.07	0.8678
	11 months - 5 months	8.0293	9.63	0.6065
	17 months - 11 months	20.1106	8.37	0.0486
auditory	5 months - 3 months	-2.6517	10.07	0.7923
	11 months - 5 months	7.5163	9.63	0.6525
	17 months - 11 months	23.8852	8.37	0.0129
frontal	5 months - 3 months	-8.2272	10.07	0.4684
	11 months - 5 months	6.9807	9.63	0.4684
	17 months - 11 months	30.2925	8.37	0.0009
motor	5 months - 3 months	-1.7938	10.07	0.8586
	11 months - 5 months	7.6601	9.63	0.6394
	17 months - 11 months	22.3641	8.37	0.0225
cingulate	5 months - 3 months	1.9184	10.07	0.8489
	11 months - 5 months	10.3756	9.63	0.4218
	17 months - 11 months	20.0787	8.37	0.0492
retrosplenial	5 months - 3 months	1.675	10.07	0.8679
	11 months - 5 months	11.0539	9.63	0.3764
	17 months - 11 months	21.103	8.37	0.0349

Table 4-3: Contrasting estimated marginal means of median cerebral blood flow. This table includes ROIs that had significant differences when comparing group (pairwise, control – EEDR) and time (consecutively) across ROIs

Summary

Group differences in select global functional metrics

Functional connectivity was assessed using several global metrics to assess the ongoing changes in the functional brain networks. Overall functional connectivity is a measurement of the covariance of time courses in the brain. Global efficiency is a measurement of the ease that information can be transferred between different regions based on the apparent strength of communication between two regions. There were various group differences in functional connectivity, including before the EEDR intervention. The control rats had higher global efficiency and overall functional connectivity at the penultimate and final time points. However, the lack of difference between the groups at the next time point is due to a larger decrease in the control group's global efficiency and functional connectivity, resulting in no difference at 5 months, suggesting that the EEDR intervention is slowing the rate of decline in global efficiency and overall functional connectivity at this early stage. However, the effect appears to be short-lived as the decline appears to be of a similar magnitude between 5 and 11 months, and the EEDR decline further between 11 and 17 months while the controls' decline plateaus (appendix 4-1).

Ageing human brain networks have been observed to decrease in modularity (Cao et al., 2014), and small-worldness with ageing (Onoda & Yamaguchi, 2013). However, neither were found to decrease in the Resilient study. A node's modularity describes the ratio of communication between its neighbour nodes and others. The control rats had increased modularity at the 5-month time point relative to the EEDR group, which is not driven by segregation of EEDR brain networks; in fact, the EEDR group also had an increase in modularity, but the modularity in the controls was larger still. Increased modularity suggests that there is more intra-network communication than inter-network. As most nodes have a small number of degrees, and these are likely to be short-distance connections, an increase in modularity may represent a decreased covariance in time courses between hub nodes which have a high number of connections, particularly between modules (Power et al., 2013).

Thus, reduced connectivity between hub nodes would decrease global efficiency as a longer path is required to traverse a graph.

The small-world propensity measures the degree to which the networks display characteristics of a small-world network, which is neither random nor regular but instead is a trade-off. Small-worldness was relatively constant in both groups (appendix 4-1), and thus there were no significant differences between the groups concerning the small-worldness of the networks. As small-world propensity takes into account both global efficiency and clustering coefficient, and there were differences between the groups in global efficiency, it is implied that the clustering coefficient must also be changing to maintain the same small-world propensity. As global efficiency was found to decrease with age, to keep small-world propensity constant would suggest an increase in clustering coefficient, which is in agreement with the modularity results because a high clustering coefficient would suggest a modular network (Fornito et al., 2016).

However, the values for maximum modularity and small-worldness propensity are very noisy for sparse graphs, particularly those for less than 5 % density. The extreme sparsity of the graph is likely to be too stringent, which consequently disconnects critical hub regions and changes topological metrics. While densities of 5 % density and fewer have been used to view metrics over a range of graph densities in the literature before (Achard & Bullmore, 2007), the threshold likely had an insufficient number of edges for calculating the global metrics. For instance, structural brain network graphs show small-world characteristics in graphs that have been a graph cost between 8-27 % (Sun et al., 2012), so future work should repeat these measurements with a more fine-tuned density range to explore if the sensitivity of the measurements can be improved.

A limitation of these metrics is that it is difficult to derive the underlying biological changes that drive the observed differences. Functional connectivity is related to white matter connections, but regions that are functionally connected are not necessarily physically connected by a single white matter pathway, and instead can

be mediated by one or more hub regions (Honey et al., 2009). Moreover, the difference between structural and functional connectivity has been shown to increase with age, particularly in those with age-related pathologies (Reijmer et al., 2015). Thus, without pathology or a given task, global measures are blunt instruments for detecting the finite changes that occur, even on a regional level. As the default mode network has repeatedly been shown to decline with age, it was chosen for further investigation.

DMN covariance declined with age

These findings have successfully replicated the decrease in overall functional connectivity within the DMN with increasing age, which has been observed in humans (for review, see: Sala-Llloch et al., 2015) and rodents (Hsu et al., 2016; Ash et al., 2016). This decrease was further explored between groups since the effect of a ‘healthy’ lifestyle on the DMN is not known in rodents. Regardless of the group, the default mode network declined in its overall functional connectivity with age, with both groups decreasing by $Z \approx 0.33$ between 3 months and 17 months.

However, there was a distinct effect of group on DMN decline: the EEDR group declined at a slower rate, with the largest difference between time points occurring between 5-months and 11-months, while the control group declined rapidly between 3-months and 5-months before stabilising. Neither group had large differences in overall functional connectivity from the 11-month and 17-month time points, suggesting that the mechanism for DMN decline may be occurring at middle age or earlier. These findings are further evidence that there should be a larger emphasis on human ageing studies that are designed to be longitudinal. Many ageing studies compare ‘young’ with ‘old’, instead of viewing ageing as a continuous process. Instead, recruitment should include participants from a range of ages, and age should be considered as a continuous variable instead of a group. By doing so, more meaningful age-related changes, such as the onset of decline, can be explored.

Cerebral blood flow increased with age

Contrary to our hypothesis and the existing human literature, the Resilient study found that CBF largely increased with age, rather than decreased. However, as expected, regional differences were the largest driver of variance in CBF, suggesting regional trajectories must be taken into account when considering the change in perfusion rates. Regional differences in CBF have been observed in humans and rats (Chen et al., 2011a; Ohata et al., 1981; Wei et al., 2020) with increasing age, although admittedly in opposite directions, so the regional differences in blood flow are unsurprising.

Moreover, particular regions are more susceptible to age-related changes than others, including several cortical regions, the hippocampus, and the thalamus, as shown by the interaction between region and age, which marks demonstrable changes in regional perfusion with age. The largest changes between the groups in blood flow occurred between 11 and 17 months, particularly in regions that were particularly susceptible to age-related changes.

Although the main effect of group was not significant, there were significant interactions which suggest that there were regional effects of EEDR. Furthermore, contrast testing by group showed the groups had a differential outcome with age: the main effect of age is not significant in the EEDR group, but it is for the control group. That is likely because key regions had an increase in the control group's median CBF between 11 and 17 months compared to the EEDR group. The exception to this is the substantia nigra, which had an increase in blood flow at 5 months in the EEDR group. Moreover, the regions that were found to have different trajectories between the groups included 7 out of 8 regions from the default mode network analysis. Given that the largest difference in blood flow changes occurred between 11 months and 17 months, and the largest difference in DMN connectivity occurred before 11 months, functional reorganisation appears to precede changes in perfusion in this cohort.

Despite the tight coupling between oxygen transportation and energy requirements of neurons, the observed changes in CBF do not necessitate age-related changes in metabolism. Regional blood flow and local glucose utilisation do not follow identical ageing trajectories across the lifespan of the rat brain (Ohata et al., 1981), particularly between the ages of 3 and 12 months old, so the two may become uncoupled to different degrees throughout the lifespan.

Moreover, changes in blood flow may not be directly related to metabolism, but instead could be related to an age-related effect to anaesthesia, for example. Aged rats are more sensitive to general anaesthetics; they took longer to emerge, required a smaller dose to maintain loss-of-righting reflex, and showed decreased EEG spectral power compared to young rats (Chemali et al., 2015).

High variance in group-level measures

Many of these results displayed high degrees of variability at a subject level. Therefore, in contrast to our initial hypothesis that EEDR implementation would model a healthy lifestyle, it seems that both groups are more heterogeneous than initially expected, with a great deal of overlap between the groups. Therefore, EEDR implementation may have exaggerated the tails of the distribution of individual differences in ageing, which would explain the means of the global metrics being within one standard deviation of each other, despite the magnitude of difference between many of the global metric curves indicating that they are statistically significant. These findings suggest that the heterogeneity of the group is greater than intended, and therefore there may be ‘good agers’ and ‘bad agers’ in each group, inducing a large degree of individual variation that impedes observation of age-related changes.

5. EXPLORING HETEROGENEITY IN THE RESILIENT COHORT

Rationale

The previous analyses have tested the hypothesis that implementing an enriched environment and diet restriction on a group of rats would make them become ‘healthy agers’ relative to rats that have lived in standard conditions, and therefore various metrics have been compared between the two groups. However, fewer group differences were exposed than expected, likely due to large within-group variance leading to considerable overlap in the measured values. These findings suggest that the EEDR implementation is not sufficient to emulate healthy ageing. Instead, EEDR ostensibly generated an entire spectrum of ageing, with the groups overlapping in the middle of the spectrum. The resulting cohort is more similar to human ageing than initially intended, particularly because of the large degree of individual difference, even in shared environments.

Moreover, while the EEDR group have consistently been hypothesised to be the ‘healthy’ group, the experimental design is likely to be influenced by survivor bias, particularly at the final time point. Survivor bias is a phenomenon concerning the impact of a latent influence on the subjects of a study, whereby the subjects observed are not impacted by the latent influence, while those that were not observed are. Consequently, the former are not representative of the entire population. For instance, in a study of treatments for cats that fall from buildings, cats that fell more than six stories had fewer injuries than those that fell from a lower height (Whitney & Mehlhaff, 1987). One might assume that cats are more resilient to falls of greater heights while overlooking the premise that if most cats that fall from greater than six stories do not survive, treatment will not be sought. Thus, the degree of injury resulting from lower falls will be more variable than that of high falls. Analogously, studies of healthy ageing in humans are likely to be vulnerable to survival bias because the individuals who are healthy enough to return for multiple visits are also more

likely to continue with the study, while those that have poorer health outcomes are more likely to drop out.

Therefore, a new hypothesis is suggested: the ten subjects of the control group that were scanned at the final time point are actually exemplary of successful ageing, while the 20 subjects in the EEDR group will be a mixture of both healthy ageing, i.e., those that were always healthy and at no risk of age-related decline, and successful ageing, such that they were perhaps prone to poor ageing, but the implementation of EEDR ameliorated their outcomes. From the 14 rats in the control group that were excluded before the end of the study, six were subject to poor health and succumbed to age-related illness or decline, and eight were excluded for other reasons, thus leaving ten rats that have been successful in ageing despite their lack of a ‘healthy’ lifestyle. This discrepancy between the groups is illustrated in figure 5-1.

Figure 5-1: Representation of overlapping outcomes between the groups

In this chapter, additional non-scanning metrics are explored to gain a better characterisation of the cohort, comprising the assessment of physical and behavioural outcomes. Following this, relationships and patterns between the metrics are explored, with the prospect of finding an alternative classification method that might better characterise ageing outcomes within the cohort. Animals were consequently classified based on behaviour outcomes and study outcomes, and these classifications were used to compare the other behavioural and physical metrics, based on the hypothesis that group assignment does not provide a robust delineation of the final ageing outcomes.

Finally, to address the large variance in most of the measures, the degree of individualisation in the present findings is explored. Biological outcomes are believed to be a consequence of genetics, environment or a combination of the two. In this cohort, the genetics are more controlled than in a comparable human study, although Sprague-Dawley rats are outbred, and therefore it is not expected that the cohort is genetically very similar. Moreover, their environment has been mostly controlled, with standardised enriched or traditional cage environments, in the same room, and subject to the same experiences. However, there is still a large degree of non-shared environment that has not been controlled, and this includes individual drives for exploration in the environment, personal experiences, and level of activity (Kempermann, 2019). The final section of this chapter addresses the primary age-related findings of the previous chapters and tests how much variance can be explained by the environment (group), genetics (weight), and age, or any combination of the three factors. The degree of individual variance is considered as the residual variance that is not explained by the model, and which may be indicative that the ageing trajectory has substantial influence by non-shared environmental factors.

Methods

Lifespan outcomes

As attrition is a common consideration for longitudinal studies, subject numbers were expected to diminish in this study. Consequently, there are fewer scans available for analysis at the penultimate and final time points compared to the first two scanning time points. However, this was not entirely due to death, as some rats were too large to scan safely, or some were excluded from the study as a consequence of the death of their cage mate despite potentially no problems in their own health.

Animals were classified as one of three outcomes: ‘end of study’ denotes animals that were scanned at the final time point; ‘age-related outcomes’ denoted animals that were excluded from the study because of age-related outcomes including arthritis, cancer, abscesses, and diabetes; ‘other’ outcomes included all other animals who could not be included in the study but were not excluded explicitly for health-related reasons. Consequently, it is unclear whether the ‘other’ group would have been good or bad agers and can be considered as a combination of the other two groups.

Behavioural and physical outcomes

In addition to the scanning protocol, various other descriptive metrics, which are listed in Table 5-1 by scanning time point, were collected from the Resilient cohort throughout the study.

Age	Total scanned	EEDR	Control	Behaviour testing
3 months	48	24	24	Open Field Tests Grip strength
5 months	48	24	24	Open Field Tests
11 months	40	23	17	Open Field Tests Frailty index
17 months	30	20	10	Attentional set shifting Frailty index Grip strength

Table 5-1: Number of animals per group at each scanning time point and associated measures

In addition to these measures, all rats had their weight recorded weekly and before each scan. Weight is a metric that is frequently synonymous with age in rats (Quinn, 2005), but it can be highly variable between individuals, sexes, strains, and suppliers of experimental animals (McCutcheon & Marinelli, 2009). Further physical metrics included measurement of frailty and grip strength. Frailty is a condition that is characterised by the increased risk of adverse health outcomes, and the symptoms associated with frailty become more common with age. A frailty index (see Clinical Frailty Index) was used to measure the manifestation of conditions associated with being frail in all rats at 11 months and 17 months. The outcome assessment was supplemented by grip strength, which was assessed as described previously (see Grip Strength test) at the first time point for all rats, and at the final time point for those rats that were scanned at the last time point. Grip strength is a potential biomarker for healthy ageing in humans (Syddall et al., 2003; Bohannon, 2019) and rats have been shown to demonstrate age-related declines in grip strength (Gallagher & Burwell, 1989).

Open-field behavioural tests included locomotor activity and novel object recognition (see Locomotor activity and Novel Object Paradigms), which were tested at the first three scanning time points. Locomotor activity tests measure voluntary movements, such as time spent moving and the distance travelled, and can also gauge anxiety by measuring the time spent in the exposed centre of the arena relative to the sheltered space beside the arena walls. It was hypothesised that the variance of voluntary movement could be an indicator of heterogeneity within the

cohort, as individual differences in exploration preferences have been shown in genetically identical mice (Freund et al., 2013). Roaming entropy was calculated as a measurement of motivation to explore an enriched home cage, and individual differences in roaming entropy positively correlated with adult neurogenesis in the hippocampus, the rate of which is generally reduced with age (Galvan & Jin, 2007). Novel object recognition is a prevalent test for hippocampal-dependent recognition memory in rodents (Broadbent et al., 2010), which has been shown to incur age-related deficits in performance if the delay between familiarisation and test sessions are more than 15 minutes apart (Burke et al., 2012).

At the 17-month time point, a subset of nine EEDR and nine control rats were tested on the attentional set-shifting task (see Attentional Set-Shifting). This test relies on the medial frontal cortex (Birrell & Brown, 2000) and has been shown to have age-related deficits in aged rats (Beas et al., 2013). The aged rats' performance was compared to a naïve cohort of ten male Sprague-Dawley rats that were approximately 3 months old at the time of testing.

Statistics

Behaviour tests and physical outcomes were analysed using GraphPad Prism 8. Normality was assessed with a D'Agostino and Pearson normality test; if the data were normally distributed, the mean $\pm$ SEM is presented in text, with the mean and standard deviation given in the graphs, and otherwise, the median is reported accompanied by the range of scores. Mixed-effects models were used for longitudinal analyses, in place of two-way repeated-measures ANOVAs, to control for missing values in later time points.

MRI metrics were derived from analyses described in previous chapters except for the global median CBF and the default mode network group means for each outcome group. Global median CBF was derived as described before (see Regional cerebral blood flow), using the template mask projected into the native space of the CBF map. Mean Fisher transformed correlation matrices of the DMN regions were

created for the end of study and poor health outcome groups with custom MATLAB scripts, using the Brain Connectivity Toolbox (Rubinov & Sporns, 2010) and the code for small-world propensity (Muldoon et al., 2016) as previously described (see BOLD-derived global and DMN functional connectivity).

Longitudinal models combining scanning and non-scanning metrics were fit using `lme4` and `lmerTest` in R, as described previously. Linear models were also fit, using R's `lm` command, when including weight in the model, due to the random variability largely being controlled by the weight factor. *Post-hoc* estimation of the marginal means (EMM) or the estimated marginal trends, for categorical and continuous models, respectively, were estimated and contrasted using `emmeans`. Exact p-values were adjusted for multiple comparisons using a multivariate t-distribution.

Results

Description of non-scanning metrics

Final study outcomes

As the study duration was established *a priori*, all rats were incorporated by one of three outcome categories: study end, poor health, or other (figure 5-2A). It is assumed that, concerning the assignments listed in figure 5-1, the study end group will mostly consist of a mixture of ‘successful’ and ‘good ageing’ outcomes, while the poor health group will represent those that have had ‘poor ageing’ outcomes. As the ‘other’ group were excluded for various reasons that were not related to age, it is also thought that they will have a mixture of all outcomes, since they were excluded before their outcome could be conclusively decided.

The most frequent outcome was reaching the study end, with 32 rats in total, split 21 and 11 between EEDR and control groups. The next outcome was fatality before the study end, particularly due to health-related incidents. Examples of outcomes which were classified as poor health include sudden death, rapid weight loss, arthritis, and presence of growths such as tumours or abscesses. Seven rats were in this category, six of which were in the control group. Finally, other outcomes included being excluded due to isolation, either by the death of a cage mate or aggression, or death from injuries that were not age-related. Of the nine rats in this category, seven were control rats.

All animals were present for the first two scan points, so all deaths occurred between the second and final time point (figure 5-2b). Before the third time point, five rats had been removed from the study, four of which were controls. The area under the curve (AUC) for all animals in the study is 63.81, while the EEDR and control AUCs were 67.37 and 60.25 respectively.

Figure 5-2: Study outcomes of the Resilient cohort. A) Frequency of study outcomes separated by group. B) Survival plot denoting the fraction of rats remaining in the study over time. The black line denotes all rats combined, while the EEDR group is denoted by the purple line and the control by the green line. The black dotted lines mark the start of a scan time point.

Weight

Rat weight was first recorded at 9 weeks old and was recorded weekly until the final scan (figure 5-3A). A mixed-effects model predicting weight from age, intervention, and the interaction between the fixed effects indicates that the main effect of age ($F(2.055, 82.68) = 799.4$, $p < 0.0001$) and the main effect of the intervention ($F(1, 46) = 54.27$, $p < 0.0001$) both had significant impacts on the model. Moreover, there was a significant interaction effect ($F(633, 2521) = 32.07$, $p < 0.0001$), which indicates that the two groups had different weight trajectories. Post-hoc comparisons of the groups indicate that the mean weight per group was not different at the first time point (estimated difference 8.250 g, 95 % CI of difference = -13.57 to 30.07 g, $p > 0.999$), and this remains true until the rats were approximately 4 months old (estimated difference = 44.67 g, 95 % CI of difference = 8.712 to 80.62 g, $p = 0.0037$), which was three weeks after the implementation of EEDR. These differences increased until the 11-month time point (estimated difference = 147.8 g, 95 % CI of difference = 76.55 to 219.0 g, $p < 0.0001$), at which point the mean weight of the control rats was stabilising, while the EEDR rats continue to increase. The mean weight of the groups is no longer different at week 68 (estimated

difference = 101.6 g, 95 % CI of difference = -4.399 to 207.6 g, $p = 0.0736$) until the final scan point.

A restricted cubic smoothing spline was fit to each group's mean weight across time, and the first derivative of the spline was calculated to estimate the mean rate of weight gain per week (figure 5-3B). The rate of weight gain is highest when the rat is young but quickly declines from about 14 weeks. However, at 38 weeks, the weight gain of the EEDR rats stabilises and is maintained at an increase between 3 and 4 grams per week until the final time point, while the control rats' mean weight gain is between 4 and 6 grams per week from 32 weeks until approximately 12 months. At this point, the mean weight gain continues to decrease week on week until 60 weeks old, when the mean weight is decreasing every week. Although dietary restriction for the EEDR animals was ended before the final time point to facilitate the training of the attentional set-shifting test (see Environmental Enrichment Dietary Restriction model), the rate of weight gain in the EEDR rats remains stable after the termination of dietary restriction at week 59.

The distribution of the weights at each of the time points was calculated using 10g bins for both groups. As expected, the groups' distribution of counts mostly overlaps before the EEDR intervention, although it is clear that the heaviest rats were not distributed between both groups, and instead were only in the control group (figure 5-3D). Moreover, the control group had a larger standard deviation ($SD_{\text{control}} = 31.06$) compared to the EEDR group ($SD_{\text{EEDR}} = 21.87$). At the time point following EEDR intervention, the variance of each group has increased ($SD_{\text{EEDR}} = 31.36$; $SD_{\text{control}} = 51.14$), as illustrated by the increased number of bins and the decreased frequency in any one bin (figure 5-3E). However, there is still a considerable overlap in many of the rat's weights between the groups, specifically with rats weighing between 510 and 560 grams. At 11 months, the groups' standard deviations continue to increase ($SD_{\text{EEDR}} = 53.49$; $SD_{\text{control}} = 72.81$), but an overlap persists with rats from both groups weighing between 680 – 720g). At the final time point, there is a smaller increase in standard deviation for both groups ($SD_{\text{EEDR}} = 59.93$, $SD_{\text{control}} = 78.80$).

Additionally, there is still an overlap, albeit less pronounced, with the heaviest EEDR rats counted in the 760 g bin, and the lightest control rats being counted in the 710 g bin.

Figure 5-3: Rat weights across time per group. A) mean weight per group, lines denote 95 % confidence interval of the mean. B) The mean rate of weight gain in a given week, given by the first derivative of the spline fit to mean weight of each group at each week. A black dotted line indicates the start of a scan time point, whereas a purple dotted line indicates a change to the EEDR group's living conditions. C-F) Count of rats with weight within 10 g bins for each group at C) 3 months, D) 5 months, E) 11 months, and F) 17 months.

Frailty index

A frailty score was assigned to each rat using a modified version of the frailty index questionnaire devised by Yorke et al. (2017), which characterises the presence of symptoms associated with increased mortality. As described previously (see Clinical Frailty Index), a score of 0 was given if a symptom was not present, a score of 0.5 if a symptom was present but mild, and a score of 1 if the presentation of the symptom was acute, and thus an increased score is associated with poor health. As the battery was scored on 24 criteria, the maximum score was 24, but the largest score observed was 5.5. The small maximum score is not surprising, as the scores were calculated at ages 11 and 17, which are representative of middle-aged and ‘young elderly’ rats. Thus, larger scores may have been observed in even older rats. Moreover, an extremely frail animal would probably be excluded from the study before assessment as it would be unethical to extend its suffering.

At 11 months, the most common signs of frailty were coat condition, piloerection, hearing loss, body condition, hunched posture, weight more than one standard deviation from the median, alopecia, and evidence of porphyrin secretion (figure 5-4A). Most of these symptoms were mild in presence, gaining a score of 0.5, with the most common severe symptoms, which gain a score of 1, included piloerection and coat condition.

At 17 months (figure 5-4B), the hearing loss, coat condition, and body condition are the most pervasive symptoms of frailty, although there is a decline in coat condition from 27 rats to 20. Similarly, the presence of dermatitis, tumours, and diarrhoea is the same as at the 11-month time point. Extreme variance in body weight from the cohort median, alopecia, hunched body position, presence of porphyrin, and piloerection all decreased in frequency at the 17-month time point, in comparison to 11 months, while gait disorders increased in frequency (figure 5-4C). Of these conditions, the most frequently severe was body condition, followed by coat condition. However, the frequency of severe scores for coat condition decreased

from three to two, and there were no severe scores for the most frequent at the previous time point, piloerection.

A linear mixed-effects model was used to evaluate the difference in frailty between the two groups at both time points. The main effect of age was not significant ($F(1,30) = 0.7149$, $p = 0.4045$), but the main effect of group was ($F(1,37) = 18.15$, $p = 0.0001$). The interaction between the main terms was not significant ($F(1,30) = 0.06706$, $p = 0.7974$). Post-hoc comparisons show that the controls have a higher frailty score at both 11 months (95 % CI of difference: 0.4829 to 2.057, $p = 0.0009$) and 17 months (95 % CI of difference: 0.4873 to 2.253, $p = 0.0014$) relative to the EEDR rats.

Spearman's r was calculated for each rat's pair of frailty scores (figure 5-4D) to assess the relationship between frailty at the two time points. The resulting moderate, positive correlation ($r = 0.4583$, $p = 0.0083$) indicates that a rat with a low frailty score at 11 months was likely to also have a low frailty score at the 17-month time point. Multiple linear regression was employed to test if frailty at the 17-month score could be predicted from the score at 11 months, the experimental group, and the possible interaction between the two factors, as such:

$$frailty_{17i} = \beta_0 + \beta_1 group_i + \beta_2 Frailty_{11i} + \beta_3 group_i Frailty_{11i} + \varepsilon_i$$

where i refers to the i th subject. The model explained 36 % of the variance in the 17-month frailty scores (adjusted $R^2 = 0.3638$). It estimated the intercept as 0.4496 ± 0.7413 ($p = 0.5491$) and found no significant effect of group ($\beta = 0.3769 \pm 0.8465$, $p = 0.6596$), but there was a significant effect of the 11-month frailty score ($\beta = 0.8453 \pm 0.2955$, $p = 0.0079$). The interaction effect was also not significant ($\beta = -0.5619 \pm 0.3827$, $p = 0.1532$).

Figure 5-4E shows the scores for each rat at both time points, highlighting group differences in how frailty indices changed over time. The five rats with the highest frailty scores at the first time point were control rats; two of which also have the highest scores at 17 months, and the other three do not have a 17-month frailty score

due to attrition. All rats that scored 1.5 or lower at 11 months were present at the last time point and scored 3.5 or less. However, scoring between 2 and 3.5 brought heterogeneous outcomes for both groups, with attrition, improvements, and declines occurring. Finally, figure 5-3F shows the change in frailty score from the first to the second time point; a positive score indicates an increase in the frailty index. The mean of the EEDR group, -0.2143, was lower than that of the controls, 0.0909. However, an unpaired t-test indicates that the means are not significantly different (estimated difference = 0.3052 ± 0.3886 , $p = 0.4383$), which is not surprising given the small difference in individual scores over time.

Figure 5-4: Frailty index between groups. Counts of rats that displayed each measure of the frailty index at A) 11 months or B) 17 months. Black bars indicate the number of rats which scored a half point on the metric, indicating the symptom was present but not severe, and pink bars indicate the number of rats that scored a full point on the metric, indicating the extensive presence of the condition. C) Heat map of condition incidence across time points. The colour indicates the number of rats with any degree of condition manifestation D) The correlation of frailty scores at 11 and 17 months. E) Frailty scores for each rat at 11 and 17 months. F) The difference in frailty score at 17 months from 11 months, broken down by experimental group. The error bars denote one standard deviation, while the line indicates the group mean.

Grip strength

Grip strength is known to be an indicator of physical fitness in ageing humans, and consequently, it was explored as an indicator of physical fitness in this study. The mean of three grip strength trials was calculated at 3 months and 17 months. To test if these two measures were correlated, Pearson's R was calculated for all available pairs of grip strengths. The two measurements display a moderate, positive correlation ($r = 0.3948$, $p = 0.0253$, figure 5-5A), but there is still a large degree of variance unexplained ($R^2 = 0.1559$). A mixed-effects model was fit to determine if the main factors of age or group, or an interaction between them, influences grip strength. There was a significant main effect of age ($F(1,30) = 4.366$, $p = 0.0453$), but the effects of group ($F(1,46) = 1.779$, $p = 0.1888$) and the interaction effect ($F(1,30) = 0.9283$, $p = 0.3430$) were not significant. The model predicted a difference of 0.4824 ± 0.2309 newtons between 17 months and 3 months. Post hoc comparison of the model estimates shows that neither the EEDR (95 % CI: -0.4290 to 0.9491, $p = 0.6174$) or the control group (95 % CI of difference: -0.1359 to 1.546, $p = 0.1121$) showed a significant mean difference between 17 months and 3 months. Furthermore, there is no discernible pattern between group and time point means (figure 5-5B) or in the individual differences between time points (figure 5-5C).

Figure 5-5: Measures of grip strength. A) correlation of average grip strength at 3 months with grip strength at 17 months. B) average grip strength for each subject at both time points. The line denotes the mean, and the error bars represent one standard deviation. C) Subject-wise change in average grip strength between 3 months and 17 months.

Locomotor activity

The open-field test is a well-known test, which can be used to gauge two facets of rodent behaviour. The first is to quantify spontaneous activity and is measured using time spent moving and distance travelled. It is also a popular test for evaluating anxiety, on the premise that rodents are averse to exploring unknown open spaces for fear of predators, and therefore will spend less time in exposed areas, such as the most central part of a square arena.

Mixed-effects models were used to assess the effect of age and group, and their interaction, on the distance travelled and the time spent in the central zone in the locomotor activity test. Considering the distance travelled first (figure 5-6A), the main effect of age was statistically significant ($F(1.587, 69.04) = 47.83, p < 0.0001$). The main effect of group was not significant ($F(1, 46) = 0.06718, p = 0.7966$), and neither was the interaction between group and age ($F(2, 87) = 0.9302, p = 0.3984$). Post-hoc comparison of the main effect of age indicated that the rats travelled 988 ± 135.7 cm further at 3 months than at 5 months ($p < 0.0001$), 1518 ± 199.8 cm further at 3 months than 11 months ($p < 0.0001$), and 529.7 ± 139.3 cm further at 5 months than 11 months ($p = 0.0013$).

Considering the time spent in the central zone of the arena next (figure 5-6B), the main effect of age was found to be statistically significant ($F(1.813, 78.86) = 5.898, p = 0.0053$), but the main effect of group ($F(1,46) = 1.432, p = 0.2376$) and the interaction between the two terms ($F(2,87) = 2.510, p = 0.0872$) were not. Post-hoc comparisons of the main effect of age show that the rats spend -0.0481 ± 0.3989 seconds less in the central zone at 5 months than at 3 months ($p = 0.9991$), 1.113 ± 0.3145 seconds less in the central zone at 11 months compared to the 3 months ($p=0.0030$), and 1.161 ± 0.4280 seconds less time in the central zone at 11 months compared to the 5 months ($p = 0.0286$). It is clear that the animals travel less distance (figure 5-6A) and spend less time in the central zone (figure 5-6B) with increasing age.

Figure 5-6: Locomotor activity results. A) centimetres travelled for each animal across all zones and B) seconds spent in the central zone during the observation period. The mean is shown as a line with the error bars indicating one standard deviation.

Novel Object Recognition

Novel object recognition is a standard memory test for recognition memory, which infers that animals will explore an unfamiliar object more than a familiar one and therefore requires recognition of the familiar object. Given the lack of exploration of the open field in the locomotor activity test, the total time spent moving, and total exploration time (i.e., the total time spent sniffing either object) were assessed across the three time points, in addition to the more customary discrimination index. Time spent moving in the arena (figure 5-7A) appears reasonably constant across time for both groups. However, the total exploration time is reduced at time points following the baseline (figure 5-7B). A mixed-effects model was fit to assess the effect of group and time point, and their interaction, on either the time spent moving or the total time spent exploring the objects.

In the model for movement time, the main effect of age was not statistically significant ($F(1.823, 79.29) = 1.271, p = 0.2844$), but the effect of group ($F(1,46) = 6.199, p = 0.0165$) and the interaction between group and age ($F(2,87) = 4.093, p = 0.0200$) were significant. Post-hoc comparison of the effect of

age between the groups shows that the EEDR group moved 15.93 ± 4.429 seconds less at 11 months than at 3 months ($p = 0.0048$). Differences between 3 months and 5 months ($\beta = -1.171 \pm 7.019$ seconds, $p = 0.9977$), and 5 months and 11 months ($\beta = 17.10 \pm 6.974$, $p = 0.0664$) were not statistically significant. In comparison, the difference in movement for the control group was not significantly different between 3 and 5 months ($\beta = 4.751 \pm 5.949$, $p = 0.8173$), 3 and 11 months ($\beta = -2.121 \pm 5.794$, $p = 0.9777$), or 5 and 11 months ($\beta = -6.872 \pm 5.326$, $p = 0.5115$).

Given that the EEDR group decreased their movement after 3 months, the time spent specifically exploring the objects was considered. The effect of age was statistically significant ($F(1.8949, 80.45) = 18.64$, $p < 0.0001$), but the main effect of group ($F(1,46) = 2.236$, $p = 0.1417$) nor the interaction effect ($F(2,87) = 2.740$, $p = 0.0702$) was not. Post-hoc comparison of the main effect of age showed that the rats explored the objects 7.932 ± 1.711 seconds more in the baseline test than at 5 months ($p < 0.0001$), and 9.939 ± 1.897 seconds more in the baseline than at 11 months ($p < 0.0001$). The difference between the 5-month and 11-month exploration times was not significantly different ($\beta = 2.007 \pm 1.561$ seconds, $p = 0.4987$).

The median time spent exploring the objects for all animals at 5 months and 11 months was 16.094 and 14.072 seconds respectively, in a test window of 5 minutes. Given the median exploration time was equivalent to $\sim 5\%$ of the window, the justification for concluding object preference from such a short window is dubious. Consequently, only the baseline time point will be used for further analysis. A one-sample t-test was used to test whether the mean of discrimination indices at 3 months was different from zero, which would represent equal time spent exploring both objects. The observed mean was 0.1062, which is significantly different from 0 ($t = 2.370$, $p = 0.0222$, 95 % CI: 0.01589 to 0.1966), although the biological relevance of such a small difference from the outcome expected by chance is uncertain. An

unpaired t-test indicated that the discrimination index did not significantly differ between the groups at baseline ($t = 0.4296$, $p = 0.6696$, 95 % CI: -0.2214 to 0.1436), as illustrated in figure 5-7c.

Figure 5-7: Novel object recognition task results. A) time moving in seconds and B) time spent exploring in seconds, for both groups across time. C) Discrimination index per group for the 3-month test. The lines denote the mean, and the error bars represent one standard deviation.

Attentional set-shifting

Attentional set-shifting is a task that has been derived for assessing executive function. The task requires learning a rule for discriminating between two choices, only one of which will be rewarded. Attention is considered successfully applied when the rule is applied correctly in order to receive the reward, and this is assessed using various measurements including trials to criteria, time to criteria, and the number of errors. Successful completion was defined as six consecutive correct choices, with one error resetting the progress and thus increasing the number of trials. The task increases in complexity with every round, starting with a simple discrimination of only the relevant ‘set’ being present, before moving to a compound discrimination that includes the relevant set and an extraneous set that is not associated with the reward. Further trials include the reversal of the rule within the relevant dimension, intra-dimensional shifts, when the same rule applies to an unseen set, and extra-dimensional shifts, when the previously extraneous set is now the set which determines the reward in a third unseen set. By assessing the perseverance of

errors across the various trial types, different facets of executive function can be dissociated, thus making it an incredibly informative behavioural test for ageing studies (Bizon et al., 2012).

The median score for the time taken to reach criteria, trials to criteria, and the number of errors for each group at each trial type is available in table 5-3. The young rats' performance on the simple, compound, intra-dimensional, and extra-dimensional trials was, on average, in the minimum number of trials and with no errors. Reversal trials required more trials to complete than the relative original trial, and most rats made at least one mistake in the reversal trials. There was also more variability between subjects in reversal trials, compared to the original trials. Despite increases in errors and trials, the time to completion stayed mostly consistent across trial types with the median time lasting between 10 and 16 minutes across the trials, although there was a large degree of variability in the simple discrimination and the first reversal trials.

The EEDR rats required more trials for the first and third reversal sets on average but, in contrast to the young rats, completed the intra-dimensional reversal in the minimal number of trials on average, with only two rats scoring more than one error. Moreover, the median time to criteria better reflects the difficulty of each trial: the reversal trials take longer than their original counterpart. However, the most straightforward trial, the simple discrimination, required the same length of time on average as the extra-dimensional reversal.

Finally, unlike the number of trials taken by the EEDR and the young rats, the median number of trials for the control rats was mostly unchanged across the set difficulties. However, the presence of at least one error on the compound and intra-dimensional reversal trials does suggest these trials were more difficult than the others. The time required to complete each set was similar to that observed in the EEDR rats: the longest trials were the first trial, the simple discrimination, and the reversal trials.

The metrics are related, as the number of trials increases only in the event of errors being made, although they are not interchangeable because a late error (i.e., one made after 4 or 5 correct choices) will increase the number of trials more than an early error. While one would expect that time would be related to the number of trials, and therefore the number of errors, in so much that more trials require a longer time to complete, it appears that this is less frequently the case. For instance, many of the aged rats that made no errors completed the set in more time than young rats that made errors. Moreover, the young group was less variable in the time taken to complete than the aged rats, despite the former having more error variability than the latter across all trials. The correlations of time, trials, and errors across all trials for each group are displayed in figure 5-8D-F.

Only time was considered for further analysis, due to the notable floor effects of the number of trials and total errors, and thus lack of variability in measurements except for time to complete the task. A two-way repeated-measures ANOVA was used to assess the effects of the experimental group and discrimination set, and their interaction, on log-transformed time scores. The set type explained 7.525 % of the total variation, which was found to be statistically significant ($F(12,144) = 1.590$, $p = 0.0056$). However, the main effect of group, which accounted for 6.324 % of the total variance, was not found to be statistically significant ($F(2,24) = 2.086$, $p = 0.1461$). The interaction between group and discrimination set explained 5.822 % of the total variation but was not found to be statistically significant ($F(12,144) = 1.590$, $p = 0.1007$). However, subject was found to explain 36.38 % of the total variability ($F(24,144) = 4.966$, $p < 0.0001$). As time differences between the discrimination sets were expected, and the main effect of group was not statistically significant, no post-hoc comparisons were pursued.

Test set	Minutes (lower limit, upper limit)			Trials (lower limit, upper limit)			Errors (lower limit, upper limit)		
	Young	EEDR	Control	Young	EEDR	Control	Young	EEDR	Control
Simple	12 (9, 44)	18 (15, 35)	17 (10, 43)	6 (6, 17)	6 (6, 10)	6 (6, 20)	0 (0, 5)	0 (0, 6)	0 (0, 5)
Compound	11 (9, 18)	13 (7, 17)	11 (7, 36)	6 (6, 12)	6 (6, 12)	6 (6, 9)	0 (0, 4)	0 (0, 2)	0 (0, 1)
Reversal 1	15 (9, 47)	24 (12, 54)	14 (6, 65)	9 (6, 12)	9 (6, 25)	7 (6, 14)	1 (0, 6)	1 (0, 10)	1 (0, 3)
Intra-dimensional	16 (9, 23)	14 (10, 19)	13 (8, 34)	6 (6, 17)	6 (6, 6)	6 (6, 15)	0 (0, 3)	0 (0, 0)	0 (0, 2)
Reversal 2	12 (9, 19)	20 (9, 40)	15 (9, 50)	11 (6, 25)	6 (6, 18)	7 (6, 19)	2 (0, 8)	0 (0, 3)	1 (0, 6)
Extra-dimensional	10 (8, 16)	16 (9, 26)	12 (8, 40)	6 (6, 19)	6 (6, 15)	6 (6, 8)	0 (0, 4)	0 (0, 4)	0 (0, 1)
Reversal 3	10 (7, 20)	18 (9, 64)	13(6, 41)	11 (6, 22)	11 (6, 25)	6 (6, 15)	2 (0, 8)	2 (0, 23)	0 (0, 2)

Table 5-2: Attentional set-shifting task results. Subjects were required to achieve criteria of 6 consecutive correct choices, and results are reported by group and set type. All scores are presented as median values with the lower and upper limit in the group.

Figure 5-8: Attentional set-shifting task results. A) The time in minutes taken to achieve criteria of 6 consecutive correct choices; B) The number of trials required to achieve criteria; C) The number of errors made before achieving criteria. Purple circles represent EEDR rats, green circles represent control rats, and grey triangles represent young rats. The error bars represent the interquartile range, and the line represents the median correlation of D) trials to criteria and number of errors made, E) trials to criteria and time to criteria, and F) the number of errors made and time to criteria, for each group across all trials.

Association of non-scanning metrics

Are cognitive outcomes from the final cohort consistent with early life cognition scores?

To better characterise the animals that were included in the final time point, the subjects were separated into groups based on their time to complete the attentional set-shifting task relative to the young rats. Young rats required a mean time of 97 minutes to complete the entire test, with a standard deviation of 26.55 minutes. Thus, the nine aged rats that completed all trials in under 124 minutes were designated unimpaired, and the other nine were considered age impaired. Of the impaired rats, six were in the EEDR group, while three were in the control group. All rats except one, an unimpaired control rat which suddenly died before scanning took place, survived to the end of the study. The exception was excluded from further analyses. The previously described physical and behavioural metrics were compared between the aged impaired and unimpaired groups (appendix 5-1). However, there was no indication on any of the metrics that separating the end-of-study group by the time to complete the attentional set-shifting task was indicative of sub-populations which could not be accounted for by the experimental groups. Therefore, no statistical tests were carried out following this observation.

Are non-scanning metrics characteristic of life outcome?

To further explore the hypothesis that life outcome is reflected in the previously collected non-scanning outcomes, the animals were divided into new groups based on their study outcome: 'end of study', 'poor health', and 'other'.

Mid-life frailty

Frailty at 11 months was compared between groups defined by life outcome (figure 5-9A), but since only two animals from the 'other' outcomes group were included at the 11-month time point, the median frailty at 11 months was tested between the end of study group and the poor health group. A Mann-Whitney test showed that

the median frailty was higher in the poor health group compared to the study end group ($U = 21.50$, $p = 0.0058$).

Physical and behavioural outcomes before intervention

Weight at 3 months was explored using a Welch's one-way ANOVA between the outcome groups (figure 5-9B), which found a significant effect of outcome on weight ($W(2, 12.68) = 9.326$, $p = 0.0049$). The rats that reached the end of the study were lighter than those of the poor health group by 33.18 ± 9.951 g (95 % CI: - 66.27 to - 0.09712 g, $p = 0.0246$), and lighter than those of the 'other outcomes' group by 29.14 ± 9.352 g (95 % CI: -58.39 to 0.1218, $p = 0.0256$). There was not a significant difference in weight between those of poor health and the other category (95 % CI: -33.91 to 42.01, $p = 0.9821$). Grip strength at 3 months was similarly compared between the three outcome groups (figure 5-9C) with a Welch's one-way ANOVA. However, the main effect of group was not significant ($W(2, 10.13) = 0.9094$, $p = 0.4333$), so no further comparison between the groups was sought.

Behavioural metrics at 3 months included the distance travelled (figure 5-9D) and time spent moving (figure 5-9E) in the locomotor activity test and the discrimination index from novel object recognition (figure 5-9F). Welch's one-way ANOVAs found that there were no significant effects of the grouping on distance travelled ($W(2, 10.94) = 0.5192$, $p = 0.6089$), time spent moving ($W(2, 11.01) = 0.8173$, $p = 0.4667$), or discrimination index ($W(2.00, 15.11) = 0.6655$, $p = 0.5284$). Therefore, no further comparison between groups were explored.

Figure 5-9: The relationship of life outcome to other non-scanning metrics. A-C denote physical outcomes, while D-F denotes behaviour outcomes. A) The frailty index measured at 11 months; B) Weight in grams at the 3-month scanning time point; C) Average grip strength at 3 months old in newtons; D) distance travelled in the open field during the locomotor activity task; E) time spent moving in the open field during the locomotor activity task; F) discrimination index during the novel object recognition task.

How do brain network metrics differ with life outcome?

As the poor outcome group were not present at the final time point by definition, models were created excluding the final time point *a priori*. Moreover, as the ‘other’ category was expected to be a mix of animals that would have had good health beyond the final time point and those that would not, it was concluded that they were representative of neither and were also excluded *a priori*.

Life outcome and morphometric similarity networks

Conditional growth models with random intercepts were fit to the four graph-theoretical metrics described in chapter 3 (see Graph-theoretical analyses), using outcome in place of group (figure 5-10).

The global efficiency estimate for the end of study group at 3 months was 0.3920 ± 0.003582 . The estimate for the poor health group at 3 months was not significantly different from that of the end of study group ($\beta = 0.0004775 \pm 0.008454, p = 0.955$), and the estimates for the end of study group at 5 months ($\beta = -0.006482 \pm 0.004899, p = 0.190$) and 11 months ($\beta = -0.0004039 \pm 0.004984, p = 0.936$) were not different from those observed at baseline. The interaction effects were also not significant, denoting that the poor health group had similar differences between baseline and 5 months ($\beta = 0.01058 \pm 0.01156, p = 0.363$) and baseline and 11 months ($\beta = -0.006610 \pm 0.01201, p = 0.584$) as the difference in the corresponding time points in the end of study group.

The end of study group at 3 months had an average clustering coefficient of 0.627620 ± 0.010221 . The difference in the estimate for the poor health group was not significantly different from this estimate ($\beta = 0.001195 \pm 0.024126, p = 0.961$). Moreover, the estimates for the end of study group at 5 months ($\beta = 0.016509 \pm 0.013059, p = 0.210$) and 11 months ($\beta = 0.016712 \pm 0.013308, p = 0.213$) were not significantly different from the baseline estimate. There was also no significant interaction effect, which indicates that the estimates for the poor health group at 5 months ($\beta = -0.023090 \pm 0.0030825, p = 0.456$), and 11 months ($\beta = 0.007790 \pm 0.032122, p = 0.809$) were not significantly different from the corresponding contrast of the end of study group to baseline.

The estimate for modularity in the end of study group at baseline was 0.610049 ± 0.005174 , and the corresponding estimate for the poor health group was not significantly different from that of the end of study group ($\beta = -0.014036 \pm 0.012213, p = 0.2530$). However, the estimate at 5 months ($\beta = 0.023133 \pm 0.007313, p = 0.0020$) and 11 months ($\beta = 0.021639 \pm 0.007438, p = 0.0044$) show increases in the modularity in the end of study group compared to baseline. The interaction effects were not significant, indicating that the poor health

group showed similar increases relative to baseline at 5 months ($\beta = 0.017711 \pm 0.017272$, $p = 0.30745$) and 11 months ($\beta = 0.022685 \pm 0.017902$, $p = 0.2078$). This is confirmed by *post-hoc* contrasts of the estimated marginal means for consecutive time points by outcome group. The study end group had higher modularity at 5 months compared to 3 months ($\beta = 0.02313 \pm 0.00732$, $p = 0.0045$), which was also true for the poor health group ($\beta = 0.04084 \pm 0.01565$, $p = 0.0209$). Neither the study end group ($\beta = -0.00149 \pm 0.00744$, $p = 0.9710$) or the poor health group ($\beta = 0.00348 \pm 0.01631$, $p = 0.9676$) displayed significant differences in modularity between the 5- and 11-month time points.

Finally, the end of study group had an estimated small-world propensity of 0.901939 ± 0.003058 at the 3-month time point. The estimate for the poor health group at baseline was lower, but the difference did not surpass significance criteria ($\beta = -0.014280 \pm 0.007218$, $p = 0.0504$). The estimates for the end of study group did not differ from baseline at 5 months ($\beta = -0.014280 \pm 0.007218$, $p = 0.3374$) or at 11 months ($\beta = -0.001462 \pm 0.004396$, $p = 0.7401$). However, there was a significant interaction effect, which indicated that the poor health group estimate at 5 months was larger than the difference from baseline in the end of study group ($\beta = 0.026672 \pm 0.010207$, $p = 0.0103$), but the increase from baseline at 11 months was not significantly different from that observed from the end of study group ($\beta = 0.010998 \pm 0.010580$, $p = 0.3009$). These interactions were investigated further using post-hoc marginal means comparisons, where the study end group did not show any significant differences between either 3 and 5 months ($\beta = -0.00417 \pm 0.00432$, $p = 0.5303$) or 5 and 11 months ($\beta = 0.00270 \pm 0.00440$, $p = 0.7648$). While the poor health group did not display any changes in small-worldness between 5 months and 11 months ($\beta = -0.01297 \pm 0.00964$, $p = 0.3081$), there was a significant increase observed from 3 months to 5 months ($\beta = 0.02250 \pm 0.00925$, $p = 0.0328$).

Figure 5-10: Differences in structural network graph theory metrics between life outcome groups. Measures include global efficiency (E_{glob}), global clustering coefficient (C), modularity (Q), and small-world propensity (ϕ) over time between animals that were included at the study end and those that were excluded for poor health. The points are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

Life outcome and overall functional connectivity of the DMN

A conditional growth model with a random intercept was fit to predict overall functional connectivity from outcome and time point (figure 5-14). The intercept, which is representative of the end of study group at time point 3, was estimated as 0.7066 ± 0.03424 . The estimate for poor health group was not significantly different from the end of study group ($\beta = -0.08095 \pm 0.08566$, $p = 0.3469$), but the estimates for the end of study group did decrease significantly at 5 months ($\beta = -0.16524 \pm 0.04217$, $p = 0.0002$) and 11 months ($\beta = -0.28946 \pm 0.04300$, $p < 0.0001$) relative to the 3-month time point. The poor health group decreased their functional connectivity between baseline and the 5-month time point with a magnitude that was not distinct from the respective difference in the end of study group ($\beta = 0.06368 \pm 0.10352$, $p = 0.540391$). The decrease between baseline and 11 months was smaller in the poor health group, but the magnitude of the difference did not significantly differ from the end of study group ($\beta = 0.20043 \pm 0.10793$, $p = 0.067228$). Pairwise comparison of the estimated marginal means of consecutive time points by outcome, indicates that in the study end group, there was an average decrease of 0.1652 ± 0.0422 between 3 and 5 months ($p = 0.0004$) and an average decrease of 0.1242 ± 0.0430 between 5 and 11 months ($p = 0.099$) in the study end group, but the poor health group stayed more stable between 3 and 5 months ($\beta = -0.1015 \pm 0.0947$, $p = 0.4652$), and between 5 and 11 months ($\beta = 0.0125 \pm 0.0947$, $p = 0.9876$).

Figure 5-11: Difference in default mode network overall connectivity by life outcome groups. Animals that reached the end of the study are compared to those that were prematurely excluded for poor age-related health outcomes. The points are fitted with an interpolated locally estimated scatterplot smoothing (LOESS) curve.

Figure 5-12: Difference in DMN connectivity between life outcome groups. Mean Fisher transformed correlations for the end of study group are compared to the poor health group

As overall functional connectivity is averaged across all regions of the default mode network, an exploratory analysis compared the mean group graphs (included in appendix 5-2) for each time point, and the difference was calculated by subtracting the poor health group mean from the end of study group mean (figure 5-12). The overall trend is similar to that observed in overall functional connectivity: there are largely stronger correlations in the study end group in the earliest time point, mixed differences, if any, between the groups in the 5-month time point, and lower correlations in the end of study group compared to the poor health group at the 11-month time point. However, the retrosplenial cortex ostensibly does not follow this trend, which appears to be slightly higher in the poor health group although, due to the small sample size of this group, future confirmatory analysis is necessary.

.

The relative contributions of genetics and shared environment on age-related outcomes

The most prominent age-related changes found in this cohort have been the increase in grey and white matter volume, the decrease in overall DMN connectivity, and the increase of average cerebral blood flow. However, when modelling these effects previously, a large degree of residual variability was persistently observed, despite the inclusion of experimental groups in the model, which represent the environment of a given animal. It is assumed that individual differences are a result of either genetics, environment, or a combination of the two. As the experimental grouping describes the prescription to one of two standardised environments, group was expected to account most of the variance, especially given the genetic similarity in Sprague-Dawley strain rats. However, the experimental condition does not model all aspects of environment; namely, the non-shared environment, which includes how individuals interact with their environment, which may be represented in the residual variance after modelling for genetics and shared environment. Given the variability present before the intervention, weight was taken as a measure that is partially representative of a genetic component for each individual, and the experimental group was considered representative of the shared environment. Linear models were fit to grey matter volume, white matter volume, median cerebral blood flow, and DMN overall connectivity, using group, weight, and their interaction as predictors.

How does body weight relate to tissue volume?

The relationship between weight and brain size has so far been overlooked, despite the substantial difference in weight between the groups; the control rats were heavier and had larger intracranial volumes than the EEDR rats for nearly all of the study. Consequently, the controls had higher volumes of grey and white matter, but their relative proportions were similar to the EEDR group (see: Global volumetric measures). Weight and intracranial volume showed a moderate positive correlation ($R = 0.5843$, $p < 0.0001$, 95 % CI: 0.4730 to 0.6772), as seen in figure 5-13.

Figure 5-13: Correlation between weight at scanning time points and total intracranial volume. Colour denotes the scan time point, while the shape represents the group: a circle, for controls, or triangle for EEDR. The line represents the linear best fit as provided by a linear model where total intracranial volume is predicted by weight, and the error margin denotes the standard error.

To further assess the relationship between weight and brain structure, a linear model was fit to predict total intracranial volume from group and weight, and the interaction between them. The model indicated that there was a significant effect of weight ($\beta = 0.55055 \pm 0.08990$, $F = 87.6607$, $p < 0.0001$), and a significant effect of group ($\beta_{\text{EEDR}} = -103.27338 \pm 79.84118$, $F = 9.7206$, $p = 0.0021$), but the interaction between them was not significant ($\beta = 0.08979 \pm 0.13866$, $F = 0.4194$, $p = 0.5182$). To establish if this change in brain size with increasing weight was related to a particular tissue type, or if all tissue types were increased equally, a second linear model was used to predict the ratio of grey matter, white matter, or CSF from weight, group, tissue type, and the interaction between them (figure 5-14). The main effects of weight ($\beta = 0.0001912 \pm 0.00001572$, $F = 102.3751$, $p < 0.0001$), group ($\beta_{\text{EEDR}} = -0.0004732 \pm 0.01396$, $F = 14.4511$, $p = 0.0002$) and tissue type ($\beta_{\text{WM}} = -0.2262 \pm 0.01347$, $\beta_{\text{CSF}} = -0.2460 \pm 0.01415$, $F = 17637.6720$, $p < 0.0001$)

were statistically significant. There were significant interactions between weight and tissue type ($F = 82.4009$, $p < 0.0001$), and between group and tissue type ($F = 14.2741$, $p < 0.0001$). However, neither the interaction between weight and group ($F = 2.3497$, $p = 0.1260$) or the three-way interaction between weight, group, and tissue type ($F = 0.1568$, $p = 0.8549$) were significant, thus indicating that the ratio of brain volumes was relatively consistent between groups, but that the relative ratio of tissues did change with increasing weight.

Figure 5-14: Ratio of tissue volume normalised for total intracranial volume with increasing weight. The bars denote the linear model line of best fit for each group with the standard error.

Grey matter volume

The adjusted R^2 value of the model, and therefore the proportion of variance accounted for by the variables, is 0.5317. The main effects of weight ($F = 155.0492$, $p < 0.0001$) and group ($F = 15.8940$, $p = 0.0001$) were significant. The main effect of age when controlling for weight and group did not surpass significance criteria ($F = 3.1497$, $p = 0.0779$). There was a significant interaction between group and age ($F = 11.4115$, $p = 0.0009$), but the interactions of weight and group ($F = 0.4461$, $p = 0.5052$), weight and age ($F = 2.7050$, $p = 0.1021$) and weight, group, and age ($F = 2.2424$, $p = 0.1363$) were not significant. Post-hoc comparison of the estimated marginal trend (figure 5-16a) of weight by group shows that the control group has a positive correlation between weight and grey matter volume at 3 months ($\beta_m = 0.761 \pm 0.110$, 95 % CI: 0.54292 to 0.979), 5 months ($\beta_m = 0.786 \pm 0.119$, 95 % CI: 0.55174 to 1.021), 11 months ($\beta_m = 0.863 \pm 0.247$, 95 % CI: 0.37444 to 1.351), and 17 months ($\beta_m = 0.939 \pm 0.412$, 95 % CI: 0.12486 to 1.753). However, this positive correlation does not change significantly between 3 and 5 months ($\Delta\beta_m = 0.0255 \pm 0.0586$, $p = 0.6646$), 5 and 11 months ($\Delta\beta_m = 0.0764 \pm 0.1758$, $p = 0.6646$), or between 11 and 17 months ($\Delta\beta_m = 0.0764 \pm 0.1758$, $p = 0.6646$). However, while the EEDR group trend is also positive, its strength is weakest at 3 months ($\beta_m = 0.291 \pm 0.149$, 95 % CI: - 0.00298 to 0.584), but grows at 5 months ($\beta_m = 0.420 \pm 0.134$, 95 % CI: 0.15598 to 0.684), 11 months ($\beta_m = 0.807 \pm 0.146$, 95 % CI: 0.51806 to 1.097), and 17 months ($\beta_m = 1.195 \pm 0.223$, 95 % CI: 0.75491 to 1.635).

White matter volume

The adjusted R^2 value of the model is 0.3518, and the model's main effects of weight ($F = 71.6666$, $p < 0.0001$) and group ($F = 6.2074$, $p = 0.0138$) are significant. The main effect of age does not surpass significance criteria ($F = 0.7532$, $p = 0.3868$). Echoing the findings in the grey matter volume model, there was a significant

interaction effect was found between group and age ($F = 9.9807$, $p = 0.001902$). Moreover, the interactions of weight and group ($F = 0.5607$, $p = 0.4551$), weight and age ($F = 2.3046$, $p = 0.1310$), and weight, group, and age ($F = 3.4492$, $p = 0.0652$) are not statistically significant. The estimated marginal trend of weight by group (figure 5-16b) reveals a positive relationship between white matter volume and weight in the control group at 3 months ($\beta_m = 0.2297 \pm 0.0563$, 95 % CI: 0.1184 to 0.3410), 5 months ($\beta_m = 0.2301 \pm 0.0606$, 95 % CI: 0.1103 to 0.3499), 11 months ($\beta_m = 0.2310 \pm 0.1262$, 95 % CI: -0.0183 to 0.4803), and 17 months ($\beta_m = 0.2320 \pm 0.2105$, 95 % CI: -0.1837 to 0.6478). In comparison, the same trend in the EEDR group shows a marginally negative relationship at 3 months ($\beta_m = -0.0745 \pm 0.0759$, 95 % CI: -0.2245 to 0.0754) and 5 months ($\beta_m = -0.0085 \pm 0.0682$, 95 % CI: -0.1433 to 0.1262), but a weak positive relationship between weight and white matter volume at 11 months ($\beta_m = 0.1895 \pm 0.0748$, 95 % CI: 0.0418 to 0.3373), and a moderate positive relationship at 17 months ($\beta_m = 0.3876 \pm 0.1137$, 95 % CI: 0.1629 to 0.6123).

Figure 5-15: The change in the mean trend between weight and tissue volume for A) grey matter volume; and B) white matter volume. The error bars denote the standard error of the mean, and the contour denotes the 95 % confidence interval of the trend estimate.

Default mode network overall functional connectivity

The adjusted R^2 for the model was 0.2827, indicating that less than a third of the variance was explained by weight and shared environment. The main effects of weight ($F = 40.6545$, $p < 0.0001$), group ($F = 11.8599$, $p = 0.0007$), and age ($F = 5.7456$, $p = 0.0177$) had significant effects on overall default mode functional connectivity. There was a significant interaction between weight and age ($F = 11.1923$, $p = 0.0010$), but the interactions between weight and group ($F = 0.4913$, $p = 0.4844$), group and age ($F = 0.1427$, $p = 0.7062$), and weight, group,

and age ($F = 1.1401$, $p = 0.2873$) were not significant. Post-hoc comparison of the estimated marginal trends shows that weight and connectivity are negatively correlated at 3 months ($\beta_m = -0.000615 \pm 0.000242$, 95 % CI: -0.001092 to -0.000137), positively correlated at 17 months ($\beta_m = 0.001405 \pm 0.000611$, 95 % CI: 0.000197 to 0.002612). Although the trend estimates at 5 ($\beta_m = -0.000326 \pm 0.000234$, 95 % CI: -0.000788 to 0.000136) and 11 months ($\beta_m = 0.000539 \pm 0.000375$, 95 % CI: -0.000202 to 0.001280) suggest negative and positive correlations with weight respectively, the confidence intervals for both contain zero and are therefore not considered significant.

Figure 5-16: The change in the mean trend between weight and default mode network functional connectivity. The error bars denote the standard error of the mean, and the contour denotes the 95 % confidence interval of the trend estimate.

Global cerebral blood flow

The adjusted R^2 for the model was 0.03271, indicating that approximately 97 % of the variance could not be explained by weight, environment, or age. Moreover, the main effects of weight ($F = 2.7738$, $p = 0.0979$), group ($F = 0.1690$, $p = 0.6816$), and

age ($F = 2.3998$, $p = 0.1234$) were not significant. There were also no significant interactions for weight and group ($F = 1.8856$, $p = 0.1717$), weight and age ($F = 3.0583$, $p = 0.0823$), group and age ($F = 0.3745$, $p = 0.5415$), or weight, group, and age ($F = 1.7830$, $p = 0.18375$). The estimated trend slope for the relationship between weight and global CBF are shown in figure 5-17, confirming the relationship has a slope $m \approx 0$ across groups and age.

Figure 5-17: The change in the mean trend between weight and median global cerebral blood flow. The error bars denote the standard error of the mean, and the contour denotes the 95 % confidence interval of the trend estimate.

Summary

Sources of noise in behavioural tests

Due to the variability observed within the groups in the MRI metrics, the behaviour results were anticipated to be a tool to delineate good agers from the bad. As cognitive function in ageing is ultimately the most important outcome, being able to classify the subjects as cognitively impaired and unimpaired would have simplified the process of identifying markers associated with those outcomes. However, it was difficult to classify animals as ‘good’, ‘successful’ or ‘bad’ based on the behaviour alone, because there was little consistency between the behaviour measures, and repeated measures were subject to habituation effects.

For instance, the decreased time spent moving in the open-field tests during subsequent time points of the open-field tests is evidence that the animals habituated to the environment faster than intended. Moreover, the animals can be observed grooming themselves on the video playback, which is further evidence that the animals were not displaying the usual tendency to explore the arena. Also, although the objects were counterbalanced across time points, and each rat was given a previously unseen set of objects in each iteration of the novel object recognition test, the subjects were not motivated to explore them, despite their novelty, and the median exploration time was $\sim 5\%$ of the window of testing after the first time point. Open-field tests are beneficial because they leverage the proclivity for rodents to explore novel surroundings after the initial habituation, and therefore do not require any training or rewards to complete the task (Seibenhener & Wooten, 2015). However, measuring the time spent in the middle as a gauge of anxiety requires an assumption that an animal is not moving because it is anxious, which was not true here. Thus, the measurement of time spent in the central zone during the locomotor activity cannot be interpreted as usual.

Although the protocol was tested before the longitudinal study, using Sprague Dawley and Lister-Hooded rats to test for strain differences (see: Choice of rat

breed), and has been used as a protocol successfully in unpublished data in our group (Cash et al., King's College London, unpublished), the protocol may have required further optimisation for this cohort of rats. Although the effect of repetition appears to have influenced the result, as demonstrated by the reduced distance travelled and time spent either in the central zone or engaging with the objects with increasing age, the result at baseline was not particularly convincing with a discrimination index of ~ 0.1 , which was tested as being significantly different from zero, but which is arguably psychologically irrelevant.

Although the discrimination index was invented in order to compare across studies, the protocols involved with the test vary widely (Antunes & Biala, 2012). The possible degrees of freedom associated with this task include, though are not limited to, the objects used, the length of time for the familiarisation session, the length of time of the test session, the length of time between the familiarisation and the test sessions, and how the discrimination index is calculated. For instance, either the absolute or relative discrimination index can be calculated; the former disregards time spent at the familiar object, while the latter uses the difference in time between both objects. Moreover, a discrimination index may not be calculated at all, instead replaced with the more relative measure of time spent exploring the novel object, without normalising the exploration time of the objects.

Furthermore, although mean values of discrimination index are reported in the values, these results are often illustrated with summary bars rather than including individual points and use standard error bars rather than standard deviation (Broadbent et al., 2010; Melani et al., 2017; Gámiz & Gallo, 2012; Ennaceur et al., 1997). Behaviour testing is well known for being subject to noise from environmental variables, such as home cage hierarchy, cues from the researcher, previous experiences, and lighting and noise of the testing environment, which can obscure reproducibility of findings (Gulinello et al., 2019). Therefore, the combination of differential effects of environmental noise, variable test protocols, and reporting of

data averages but not variance can obscure data, making it difficult to discern the quality of published findings and compare behaviour between studies.

EEDR improves frailty and survival outcomes relative to standard environments

There was an abundance of evidence that the implementation of the enriched environment and dietary restriction was beneficial for ageing outcomes, even if the group differences in other metrics were not as apparent. Only one EEDR rat was in the poor health outcome; thus, the majority of rats with poor health outcomes were from the control group, and while only 42 % (10) of the control rats survived to the final time point, 83 % (20) of the EEDR rats were scanned at the final time point. Moreover, the controls had a higher frailty score at each time point relative to the EEDR group. The difference in the groups' mean weights was statistically significant after three weeks of EEDR, which continued until week 68, which is likely because the mean weight of the control rats was brought down by exclusion of the rats with the poorest outcomes. These are in agreement with previous findings which suggest that diet restriction is beneficial for extending the lifespan of rodents, which is likely a consequence of providing neuroprotection to metabolic insults (Martin et al., 2006). Furthermore, enriched environments ostensibly augment age-related decline of memory in rodents (Nilsson et al., 1999), likely through the preservation of adult neurogenesis in the hippocampus (Speisman et al., 2013). Although there were no significant differences in the group's mean weight until after intervention, the heaviest rats were not distributed between both groups, and instead were only in the control group. The rats were randomly assigned to a group but stratified randomisation could have been employed to minimise the group differences observed before intervention, especially given the importance of variable relationships of weight and brain structure and function (see: Body weight contributes to brain metrics with different degrees in ageing).

Alternative classification did not improve on experimental groups

Despite the presence of evidence that the intervention was successfully implemented, it did not incur large differences in many group metrics as discussed in previous chapters, which was likely a consequence of the large variance within groups and individuals. To this end, alternatives to the group dichotomy were explored, such as by apparent impairment as classified in comparison to an independent young cohort, or by categorising the subjects by study outcome.

No relationship was found between cognitive outcomes in the final time point and cognition scores acquired earlier in the study. While this could be evidence of non-linearity in individual cognitive outcomes, it is more likely that the behavioural measurements are subject to noise and that obscured the underlying relationships between cognition throughout the lifespan. Rodent behavioural testing is reputedly variable, with evidence that the animals' behaviour is subject to cues from the investigator, and the particulars of the room, including odours, lighting, and temperature (Gulinello et al., 2019). Moreover, while previous studies have successfully used trials to criteria to dichotomise a group to impaired and unimpaired (Beas et al., 2013), this study found a strong floor effect in trials to criteria and therefore chose to use the time to complete the experiment, which showed more individual variability. However, this approach is arguably crude, due to the loss of information given the group variation as the trials became increased in difficulty. Therefore, more work is required to find an appropriate metric, perhaps by creating a weighted time to complete that places more emphasis on the difficulty of the trial. The open-field tasks are widely used because they are based on the premise that, once habituated to an arena, rodents have an innate preference for novelty, and therefore, it does not require additional rewards to prompt the subjects to complete the task, as the attentional set-shifting does, for example. However, this innate motivation is subject to the testing environment described above and is therefore highly sensitive. The subjects of this study showed decreasing motivation for open-

field tests in the repetitions at 5 and 11 months. Unfortunately, the results were discarded due to lack of interaction with the task. The degree that behaviour was affected by changes in motivation will likely vary between individuals, as genetically similar mice that were left to explore the same environment showed increasing group variance in exploratory motivation with increasing time (Freund et al., 2015).

An alternative approach was tested to counter the inherent variability in measures that require the co-operation of the rats, which used each rat's study outcome as the basis of categorisation. Frailty at 11 months was lower in animals that were included in the final study than those that were in the poor health group, which is further evidence of the survivorship bias hypothesis, where those animals in the control group that reached the final time point were less frail at 11 months than those that did not. Moreover, animals in the end of study group had a lower body weight at 3 months than those of the poor health group. However, there was no relationship with the other behavioural and physical metrics compared, which were subject to the motivation of the rats and are probably noisy as a result.

Differences between previously described elements of the outcome groups' structural and functional connectome were tested. There was no difference between the global efficiency, clustering coefficient, or the modularity of structural networks between the two groups, but there was an increase in small-worldness propensity of the poor health group between 3 months and 5 months. As small-worldness is a composite measure, which refers to the relative difference in clustering coefficient of a lattice network and the characteristic path length of a random network (Muldoon et al., 2016), and no differences were found in the clustering coefficients between the groups, it can be deduced that there was a change in the characteristic path length. As an increase in small-world propensity would result from a decrease in the difference from a random characteristic path length, it is hypothesised that the poor health group characteristic path length also decreased, which could be confirmed with further analysis. A decrease in characteristic path length denotes the increased ease of information transfer between nodes and, given the increase in modularity

observed in the cohort, could indicate greater covariance between hub regions. However, the characteristic path length is the inverse of global efficiency (Rubinov & Sporns, 2010), which also did not differ between the groups, so it is possible that this finding is the result of random sampling or noise.

Overall functional connectivity of the DMN was also assessed between the two outcome groups, which uncovered differences in the trends of decline between the two groups. The end of study group showed declining correlation within the DMN, between 3 and 5 months. The poor health group were comparatively stable, and the mean overall connectivity within the DMN was higher at 11 months in the poor health group than the end of study group. The stability of the poor health group DMN connectivity is unexpected, as declining DMN is observed in healthy ageing as well as pathological ageing (Sala-Llonch et al., 2015). The higher connectivity observed at 11 months in the poor health group may be the consequence of compensatory reorganisation required given the low level of overall functional connectivity at 3 months. However, as some of the end of study group had lower overall functional connectivity than any member of the poor health group, so it would be interesting to investigate whether these subjects were noise or perhaps EEDR animals whose DMN connectivity was improved after the intervention.

While the higher DMN connectivity in the poor health group relative to the end of study group is primarily reflected in the edgewise comparisons at 11 months, the retrosplenial and auditory cortices displayed higher inter-hemispheric correlations in the poor health group. The retrosplenial cortex is a known hub of the default mode network, and the covariance of the retrosplenial cortex has been shown to decrease in rats with age-related cognitive deficits (Ash et al., 2016). Therefore, an increase in inter-hemispheric covariance between the retrosplenial and other posterior cortical regions may be indicative of a compensatory mechanism to improve integration of information. However, these findings, and those of the difference in small-world propensity in the structural connectome, are speculative, not least because of the small number of animals in the poor health outcome group. While final health

outcome has the potential to be a pragmatic grouping for pursuing retrospective healthy ageing biomarkers, it would require future studies which are not restrained to a predetermined duration. However, such studies are rarely uninhibited by time due to the substantial cost associated with them.

Body weight contributes to brain metrics with different degrees in ageing

As the environmental grouping and alternatives did not provide sufficient explanation for the individual variation in many of the metrics investigated, the hypothesis that individual variation arising from non-shared environmental factors is considered. First, the contribution of body weight, which is driven by genetics, particularly at the 3-month time point before the intervention, was explored.

Unsurprisingly, body weight and total intracranial volume were positively correlated. Thus, larger rats had bigger brain volumes, and there was a difference in the groups' total intracranial volumes, but these effects were additive as there was not a significant interaction. Moreover, the composition of the total intracranial volume varied with weight, so a heavier rat had relatively more grey and white matter and less CSF compared to those with lighter body weights, and there was no significant difference in weight trends between the experimental groups.

Thus, metrics which were previously found to be highly influenced by age were re-examined by exploring the effect of weight at each time point and between groups. Grey matter volume was positively correlated with weight, but the slope was relatively stable across all time points for the control group, while in the EEDR group, it increased with increasing age. A similar trend was observed for white matter volume, where control rats showed no difference in the magnitude of their positive trend over time, but the EEDR rats show a positive weight trend at 11 and 17 months, but not earlier. Thus, for EEDR rats at 3 months, weight and tissue volume seem unrelated. However, at 11 and 17 months, the heaviest EEDR rats had the largest grey matter and white matter volumes. Conversely, for control rats, there was

always a relationship between weight and tissue volume, and this strength of this relationship does not seem to change. Thus, the heaviest control rats always had the most grey and white matter.

The comparative energy requirements for grey matter and white matter growth are likely related to the degree of variance explained by the trend slope. The primary source of energy in the brain is oxidative glucose metabolism, which requires both glucose and oxygen, and is mostly used for synaptic transmission. However, glucose metabolism occurs at a rate higher than that of oxygen metabolism, suggesting that the brain also uses glucose without oxidation as an energy source, which is known as aerobic glycolysis and is associated with synapse formation and growth (Goyal et al., 2014). Aerobic glycolysis is potentiated in the developing brain and in adult brain regions that are characterised as having developmental features, and thus in these critical points in the developmental – ageing continuum, increased availability of glucose from diet, such as in the control group, will potentiate aerobic glycolysis for increased rate of synapse formation. However, the EEDR group are undertaking a form of intermittent fasting, and therefore have a higher ratio of ketosis as an energy source, which has been shown to reduce glucose metabolism in the cortex and cerebellum in rats (Zhang et al., 2013).

In comparison, white matter is less associated with weight at early time points, but even when this relationship increases in later life, the slope of the trend is less than that of grey matter. The difference in strength for the relationship with weight between the two tissue types may be because the white matter trend is representative of the overall ratio of brain volume, which is likely driven by the increasing intracranial volume, which itself is driven by the growth of grey matter. The fraction of white matter volume normalised for total intracranial volume was found to be consistent across weights, and there is either no weight-related trend or it is negative when the rats are young adults, which is also the time when aerobic glycolysis is likely to be most abundant. However, with increasing age, the slope of the trend increases, which may be because larger rats have a higher glucose availability for aerobic

glycolysis that drives the growth of grey matter, and therefore increasing total intracranial volume, beyond the developmental period.

While there were no group differences in the trend between weight and DMN functional connectivity, the relationship did vary with age. Heavier body weight at 3 months was related with lower DMN functional connectivity, but heavier body weight at 17 months was related to higher DMN functional connectivity. Although the effect size of this trend appears very small, with an estimated slope increase of 0.001405 to the Z score for every gram of weight, for a difference in 100 grams, which is the approximate difference between EEDR and control mean body weight at week 68 when the groups are no longer significantly different, this would result in a change of ~ 0.1 in DMN connectivity Z score. However, the variance in body weight is much smaller at 3 months, and so its associated negative slope estimate suggests that a more likely increase of 50 grams in body weight results in a decrease of approximately only ~ 0.03 in the DMN connectivity Z score. Moreover, due to the proximity of the 3-month estimate to zero, further confirmatory work could explore potential underpinning mechanisms, including whether a higher weight before intervention is indicative of a metabolic vulnerability, which might be reflected in the observed increase in body weight at 3 months in the poor health outcome group. Meanwhile, at 17 months, larger bodyweight could be related to the additional energy reserves available to promote plasticity via the formation of new synapses to withstand metabolic insults. The highest dependence on aerobic glycolysis is in the default mode network, which is hypothesised to use high levels of aerobic glycolysis for new synaptic growth as a consequence of its hub-like role in integrating information from various sources (Glasser et al., 2014). Moreover, the default mode and other association networks are less myelinated compared to evolutionarily old regions, and light myelination is characteristic of increased aerobic glycolysis.

However, there was no relationship between CBF and weight at any age, which may be because CBF is a measurement of brain oxygenation and, if the slope of the weight

trend is representative of glucose metabolism above the rate of oxygen metabolism, then the measurement of oxygen metabolism may not be relevant. Alternatively, the quantitative rate of perfusion as calculated from the ASL scan is normalised for tissue volume, which is itself correlated to weight.

Residual variance persists after controlling for environment, weight, and age

When fitting these models, environment, weight, and age explain half of grey matter variance, a third of white matter variance, over a quarter of DMN connectivity variance, but only 3 % of CBF variance. Given that there are genetic and shared environmental predictors included in these models, there is substantial unaccounted variance. Future work could pursue the hypothesis that the remaining variance, in particular aerobic glycolysis metabolism and CBF, is influenced by the non-shared environment, by more using in-cage monitoring to assess hierarchies in feeding, and the resilience to external stressors. For instance, social competition for food among adolescent rats was found to reduce hippocampal cell proliferation in subordinates relative to dominant cage mates (Hoshaw et al., 2006). Chronic stress is associated with overproduction of proinflammatory cytokine, interleukin 6, which is related to age-related diseases, including cardiovascular disease and diabetes. While age-related changes in circulatory levels are contested (Maggio et al., 2006), it has been shown that its presence is predictive of successful ageing (Puzianowska-Kuźnicka et al., 2016), and IL-6 has been shown to increase in the presence of chronic environmental stressors (Kiecolt-Glaser et al., 2003). Circulating levels of these markers can be assessed in the blood plasma that was collected at each scan point for the entire cohort (see: Scanning preparation), in addition to regional evidence which can be derived from the fixation of each rat's brain tissue following the final scan session. Thus, these markers can be assessed in later work, but inflammatory markers and their relationship to the metrics described here is beyond the scope of this work.

6. CONCLUSIONS OF THE RESILIENT STUDY

The primary aims of the Resilient study were twofold: to characterise longitudinal changes in brain structure and function associated with ageing and to identify changes associated with the long-term implementation of environmental enrichment and diet restriction. Multi-modal imaging methods were employed that provide complementary information about several facets of the ageing brain to obtain a variety of metrics to identify changes in structural and functional brain networks. Structural modalities were combined to derive networks based on morphometric similarity, which relies on the finding that brain regions that are comparable in terms of their structural characterisation metrics are more likely to be physically connected. Functional changes were measured from both a network perspective, using regional BOLD covariance to determine functional connectivity graphs, and by comparing quantitative maps of cerebral blood flow. Multiple approaches and modalities were fundamental to comparing the model of healthy ageing, as any group differences were likely to be subtle because the majority of rats were healthy for the entire study. However, individual differences were not sufficiently explained by environmental condition, so alternative explanatory predictors were sought from the behaviour and physiological outcomes of the animals. The alternatives did not surpass the experimental group dichotomy at explaining variance, and thus it is suggested that an individual's non-shared environment is a fundamental driver of the non-linear ageing trajectory.

Age-related changes in brain structure and function

Structural changes

The Resilient cohort had an increased brain volume with increasing age, which was driven by increases in grey matter and white matter. However, when comparing the relative contributions of each tissue volume as a ratio to the total intracranial volume, the ratio of white matter volume stayed mostly constant over time, while grey matter ratio increased over time with a corresponding decrease in CSF ratio. This finding is

contrary to what has been found in human studies, where after young adulthood, grey matter volume has been repeatedly shown to decrease with increasing age, white matter is generally found to decline after middle age, and CSF compartments increase in volume. Hamezah et al. (2017) found decreases in the medial prefrontal cortex, striatum, and hippocampus and increases in ventricular compartments between 14-month and 27-month rats, so it is possible that our study ended before grey matter volume decreased and CSF volume increased. The duration of this study was designed to model lifestyle factors up to the onset of senescence. However, as rodent lifespans do not linearly map onto human ones (Quinn, 2005), the adulthood period for this cohort may have lasted longer than the study had planned, subsequently truncating the observations before the age-related decline.

While many of the structural network metrics did not change with age, it is interesting that the modularity of the network followed an inverted U trajectory, where modularity was lowest at 3 months and 17 months with increases at 5 and 11 months. A graph's modularity describes how easily a graph can be segregated into non-overlapping segments, so a given node's modularity describes the proportion of information transfer from that node to those within its segment versus those that are further in the graph, thus comparing within and between module communication. Low modularity indicates that there is less segregation between nodes of different modules, which could be a larger number of nodes being classified as 'rich club' nodes with many connections, or it could be that there is more robust covariance between existing hub regions. Human studies have not yet compared modularity longitudinally from structural networks, but modularity that was derived from cortical thickness was found to be lower in adults over 60 years old relative to a group of subjects in their twenties (Chen et al., 2011b), thereby missing out on the ages in between. As human structural network modularity has not previously been explored across the lifespan, it is not clear if this inverted U shape in rodents is comparable to humans.

Functional changes

A key age-related finding in human ageing literature is the decreased covariance between BOLD signal time courses of DMN regions with increasing age (Sala-Llonch et al., 2015), which was replicated here. This decline was ubiquitous, although the poor health group ostensibly moderated the decline relative to the end of study group. However, given the small number of subjects in the poor health group, it is not clear whether this is just a result of chance sampling. As a consequence of being unable to find group differences in the magnitude of the decline, it is unclear whether the observed decline in DMN covariance is a compensatory or a strategy to prevent de-differentiation.

Both global and many regional measurements of cerebral blood flow increased with age. Increased CBF is contrary to what has been observed in humans (Tarumi & Zhang, 2018), but age-related increases in CBF were observed in mice (Wei et al., 2020). As the authors of the latter paper suggest, this could be due to species differences, but it could also be the difference in relative stages of life not being equivalent lengths in humans and rodents. Thus, a longer study that has no prescribed end-date might see CBF in rodents begin to decrease.

Other changes

The Resilient cohort increased in weight with age, with the most rapid weight gain occurring from the start of observations until approximately 4 months old. The weight gain follows previous studies that have repeated bodyweight measurements of Sprague-Dawley rats sourced from Charles River, but the weight gain was found to be faster than the same strain from a different supplier (Brower et al., 2015). Age was not a significant predictor of frailty, but this could be because the animals that had the highest frailty at 11 months were subsequently excluded as part of the poor health group.

Old rats were found to make fewer errors than young rats in the attentional set-shifting task, but they took more time to reach criteria. Tests of attention can be

dissociated into distinct components; for instance, the accuracy of responses is used to quantify attention, but latency to respond can provide an insight into decision making (Muir et al., 1996). For instance, a meta-analysis into differences in decision making-strategies with age showed that when a decision is made through experience, such as in the attentional set-shifting task, rather than description, an example of which is a gambling task, that decision making shows the largest age-related difference. Moreover, the age-related difference is likely to be a consequence of individual differences in learning performance, where learning of the rule is impaired, leading to risk-averse behaviour compared to young subjects (Mata et al., 2011). Therefore, it is suggested that the old rats that took the longest to reach criteria did so because they had not learned the rule and there was a high risk of a wrong choice that would not provide a reward. In comparison, young rats that had not learned the rule made many mistakes but completed the trials quickly, suggesting that they were more accepting of the risk that a wrong choice would not reward them.

Finally, the relationship of body weight and tissue volume, and body weight and strength of DMN covariance changed over the lifespan: the strength of the relationships with weight increased with age, meaning that every additional gram of weight is associated with a larger brain volume and stronger DMN covariance. An increased relationship with bodyweight may be indicative of the availability of glucose for aerobic glycolysis, which is highest in association networks that are constantly integrating information from various sources and therefore have high energy demands (Vaishnavi et al., 2010).

Changes associated with EEDR implementation

Environmental enrichment and dietary restriction have each been shown to enhance life outcomes in rodents, but they have not been studied together in a longitudinal study. By combining environmental enrichment and dietary restriction, it was hypothesised that the group that received this intervention would display better ageing outcomes, and these outcomes would be associated with particular changes

in brain structure and function. While full divergence of the groups did not materialise, there were still differences between the two experimental groups.

Structural changes

Differences were most frequently observed at 5 months, which is unsurprising given that this is the first time point after implementation of EEDR. Moreover, given the first time point was approximately 8 weeks after the start of the intervention, it is within a window that is widely reported to involve cognitive and biological changes (Simpson & Kelly, 2011).

There are fewer studies that have investigated the long-term effects of environmental enrichment, particularly in a longitudinal cohort, and no studies that have examined the longitudinal trajectories of the combined effects of environmental enrichment and dietary restricted. Almost all structural changes disappeared by 17 months, possibly caused by a slower brain growth trajectory in the EEDR groups. Thus, the EEDR group have caught up to the stabilised control group by 17 months. It could also reflect the exclusion of control rats with the poorest health outcomes before the time point, leaving only the “successful agers”, making the groups more similar. The pattern in group differences was also reflected in the network-based statistics, which showed group differences at 5 and 11 months, but not 17 months. Group differences in microstructure were frequently not statistically significant, which is not entirely surprising, as the animals did not model overt pathology.

Graph metrics, derived from morphometric similarity networks of the structural connectome, showed little change between the groups. Global efficiency was lower in the EEDR group at 5 months, which may reflect the relatively smaller volume of white matter in the EEDR group at that time point compared to the controls. Again, this difference lessened with time, suggesting that the EEDR rats’ comparative reduction in global efficiency was ameliorated by 11 months.

Functional changes

The largest group differences of the functional connectivity graph theory analysis also occurred before the final time point. The control rats showed increased modularity at the 5-month time point relative to the EEDR group, and control rats had higher global efficiency and overall functional connectivity at the penultimate and final time points. Thus, increased modularity in the control group is observed in both structural and functional networks at 5 months. Modular networks minimise cost by reducing connections between modules, and instead, the largest proportion of connections are between neighbouring nodes within a module. Thus, modular networks are more adaptable and are robust to small changes in network topology, but modular networks are consequently less efficient at transferring information across the network (Bullmore & Sporns, 2012). Conversely, networks with high global efficiency have more integration of nodes, which are higher cost and more susceptible to small perturbations, but information transfer is more direct than in segregated networks. As the brain seems to deliver a compromise between features of integrated and segregated topologies, the control group having first increased modularity and then increased global efficiency relative to the EEDR group indicates differential trajectories in global network topology between the two groups over time.

Furthermore, although default mode network covariance declined in both groups, there were differences in the timing of the decline between the groups: the largest decline in the control rats was between 3 and 5 months, while in the EEDR group it was between 5 and 11 months. This difference provides further evidence that EEDR age slower than the control group, as the groups are not different by the final time point, and this could be the EEDR rats ‘catching up’ to the controls. As this is the first study exploring longitudinal differences in default mode network connectivity using a combined enriched environment and dietary restriction model of ageing, it appears that the EEDR intervention delays the declining covariance observed within the ageing default mode network.

Other changes

The EEDR intervention was successful at improving ageing outcomes in several ways. The EEDR group had only one rat characterised as having ‘poor health’, and 83 % of the group survived to the final time point, while only 46 % of the control group were included at the final time point. The EEDR group also had a lower mean frailty index at both measured time points. The EEDR rats were also lighter in body weight in comparison to the controls, which was true from three weeks after EEDR implementation until approximately one month before the study concluded. The rats with the heaviest body weights probably drove this finding, as they were excluded from the study for health reasons, which reduced the mean weight of the group.

There were also differences in time to criteria in the attentional set-shifting test, with the EEDR rats taking longer to complete the test than the control rats. While this could be evidence that the control rats at this point are mostly ‘successful’ agers and that the EEDR rats are a mix of ‘successful’ agers and ‘healthy’ agers, it is also likely that the food reward given with the task would motivate the two groups differently. All rats were fasted in advance of the study to generate the motivation to seek the food reward, as is standard in the test protocol. As the EEDR rats had previously experienced fasting in their life, it is possible that they were less anxious about the lack of food and therefore, less motivated to seek the reward. The controls, on the other hand, had never experienced a shortage of food, and therefore could be more motivated to find the reward.

Finally, there were also differential trends in the relationship between weight and brain tissue volume: the strength of the relationship grew with increasing age in the EEDR group, whereas it did not change with age in the control rats. Thus, for EEDR rats, the relationship between body weight and tissue volume was negligible in earlier time points, but old, heavy EEDR had increased grey matter volume compared to old, lighter EEDR rats.

Increased grey matter is an example of increased brain reserve, which can be considered as the available physical resources that an individual has to overcome age-related physical insults (Stern et al., 2020). While this is an important facet of preventing age-related cognitive decline, brain reserve is not necessarily sufficient. Cognitive reserve, which is used to explain individual differences in susceptibility to age-related cognitive, encompasses more than the physical resources available. For instance, it also includes the apparent functional reorganisation supported by ‘scaffolding’ from various sources (Park & Reuter-Lorenz, 2009; Reuter-Lorenz & Park, 2014), which consequently explains the Posterior-Anterior Shift in Ageing (Davis et al., 2008), and the increased symmetry in bilateral recruitment (Cabeza, 2002). The distinction between brain and cognitive reserve is exemplified by the poor health group having a higher mean weight at 3 months than the end of study group: the poor health group, which was mostly control rats, would likely have had high volumes of grey matter at 3 months, but this was insufficient to prevent them from succumbing to poor health outcomes. Thus, while having more grey matter should generally be considered beneficial, their ageing outcomes cannot be reduced to the idea that, for physical resources, ‘more is more’.

The role of diet and enrichment in ageing

The evidence evokes two possible hypotheses: the first supposes that the structural and functional differences from EEDR are short-lived and would require further novelty to prevent from the normalisation of the enriched environment. The second is that diet restriction is a known experimental model for extending lifespan (Martin et al., 2006), and so the development-ageing continuum of the EEDR rats is slower. There is certainly evidence that the average trajectories often occur at a later time in the EEDR for several metrics, including the DMN connectivity, and tissue volume. Moreover, the EEDR rats were continuing to gain ~5 grams in body weight per week, which stabilised before (and was seemingly unaffected) by the end of the diet restriction, while the mean rate of weight gain in the control group had started to decline.

The previously discussed ABC model (Salthouse, 2011) considers the possible relationships between age, brain structure, and cognition (see: Heterogeneity of cognitive ageing). From these data, it is hypothesised that a fourth factor drives the relationship between brain structure and cognition with age: metabolism. Aerobic glycolysis has been shown to manipulate brain structure by driving grey matter growth, particularly when there are changes in the availability of its substrate which alter the relative contribution to overall energy production compared to oxidative metabolism. Moreover, the impact of aerobic glycolysis mostly occurs in integrative regions that are important for higher cognitive functioning (Goyal et al., 2014). These metabolic changes could be manipulated by both the environmental enrichment and diet. Environmental enrichment has previously been found to reduce the expression of proteins from the glycolytic metabolic pathways that are upregulated in stressful situations (Kang et al., 2016). Dietary restriction has also been shown to reduce the cerebral metabolic rate of glucose because ketones can be used as an alternative energy substrate for oxidative metabolism, which can account for most of the cerebral metabolism (Lin et al., 2015).

For instance, as control rats were heavier in terms of body weight, they had greater availability of glucose as an energy source. The control rats also had larger total intracranial volume with more grey matter. Most energy costs are associated with synaptic transmission (Raichle & Gusnard, 2002), but increased grey matter compels growth of white matter to support the growing need for efficient information transfer across the enlarged brains (Bullmore & Sporns, 2012). White matter maintenance requires four times more energy than white matter signalling processes (Yu et al., 2018), which is likely required to reset the resting membrane potential that is essential for signal conduction (Harris & Attwell, 2012), so larger brain volumes will have higher energy requirements to maintain the existing resources, which would itself require more oxidative phosphorylation, and thus increased CBF. Rats that had rapid growth of grey matter, and subsequently more white matter to support it, in adulthood possibly had metabolic costs that were difficult to sustain in later life.

Although glucose-derived energy drives grey matter growth, grey matter maintenance is supported by oxidative phosphorylation. Oxidative phosphorylation produces reactive oxygen species (ROS) that are ostensibly protective at low concentrations but are related to cellular senescence at high concentrations (Taufenberger et al., 2019). Moreover, ROS can inhibit glycolysis at several stages (Mullarky & Cantley, 2015), which is a potential mechanism for the hypometabolism of glucose in ageing brains (Mosconi, 2013).

Functional adaptations to high-cost structural networks would likely follow, which themselves would require increased metabolic requirements. For example, the availability of glucose for aerobic glycolysis has been shown to manipulate the function of integrative brain regions, including the default mode network (Vaishnavi et al., 2010). The relationship between brain metabolism, structure, and function is likely to be a self-perpetuating loop, where changes in brain structure require further metabolic changes in a loop that further manipulates age-related changes, but essentially triggered by the availability of metabolic substrates after adolescent development.

Limitations and challenges

Study design and methods

Unfortunately, the study design had several limitations. The first of which is that the cohort comprised of only male rats, but there are sex differences in age-related declines, both in clinically normal populations (McCarrey et al., 2016) and in age-related diseases like Alzheimer's disease (Pike, 2017). An unexpected side-effect of an entirely male cohort was the rapid weight gain which had impacts on housing. Male rats grow to be heavier than female rats (Slob & van der Werff Ten Bosch, 1975), and due to the size limitations of standard animal caging, the rate of weight gain in the control rats necessitated all rats to be housed in pairs earlier than expected, after initially being housed in fours, and consequently, the rats had limited social interaction. A crucial part of many enriched environments is the opportunity for

rodents to socialise (Kempermann, 2019), and being housed in pairs restricted the possibilities for that interaction. Moreover, there are likely cage-effects, where social hierarchies and other cohabitating arrangements are different between cages, which adds to the non-shared environment that an individual rat might experience. Ideally, the experiment would have taken place in a single cage that is large enough to accommodate all the animals, as in the study by Freund et al. (2013) but the available facilities could not have handled such a cage. Moreover, in an attempt to mitigate attrition effects, occasionally on the exclusion of two animals from different cages, their cage mates would be paired together. However, a cage change brings new experiences which is a further factor to consider in the non-shared environment. These pairings can be stressful, particularly for male rats, as a hierarchy has to be established. Finally, since this did not happen equally to all rats, is a potential confound.

Another issue with implementing the model is the potential confound that food restriction had on the groups beyond the observed outcomes. For instance, when possible, rats were scanned after a period of food access, and not at the end of a fasting period. However, food restriction was undoubtedly a confound in behaviour tests, as it is easy to conceive that motivation to engage with an open-field arena or to find a food reward may have been perturbed or encouraged by the presence of food. BOLD signals are generally not found to be impacted by fasting, except for a noticeable decline in attention networks has been observed following cognitive effort (Orfanos et al., 2018). While this is reassuring for the differences in fasting state for the MRI scanning, which was captured under anaesthesia and would therefore be less likely to measure an attention network, it does serve as evidence that the open-field tests may be subject to differences in circulating glucose levels.

Despite their benefits, longitudinal designs also have limitations. The obvious example is the repetition effects that were observed in the open-field tests, where the animals lost motivation to explore the arena, and the results from those late time points had to be excluded. However, there are also potential effects in the EEDR

cohort, who may have become normalised to the enriched environment and dietary restriction. Finally, there were also physiological impacts of the study on the rats. The ageing brain is known to be particularly susceptible to repeated exposure of general anaesthesia (Wu et al., 2019). Moreover, loss of hearing is a risk factor for dementia (Livingston et al., 2017), which is likely a result of the subsequent social isolation in old age. In the Resilient study, the cohort was subjected to repeated general anaesthesia and scanning without ear protection, and this is likely to have impacted their ageing outcomes.

As discussed previously, behavioural findings are extremely susceptible to environmental variables, so the degree of noise in the collected data was disappointing but not entirely surprising. Behaviour tests are an excellent method of characterisation, but they are susceptible to small details which can derail findings, including the noise from researcher handling, light, and many other factors as described previously. Unfortunately, these problems are inherent in behavioural testing, but the effect of some of these trials can be minimised. For instance, acclimatising the animals to handling by the researcher before the experiment and minimising the researchers conducting the tests can reduce stress. Ideally, behaviour experiments will take place in a dedicated behaviour facility that has control over lighting and is soundproof to minimise noise from adjacent rooms, but realistically this is not always possible. Therefore, experiments could be planned for times when animal facilities are less likely to be busy (e.g., weekends) to reduce the likelihood of external influence.

Unlike cognitive batteries in humans, communicating the test to the subject is not possible, and it, therefore, requires a level of interpretation that the animal is acting in a given way for a particular reason. For example, in the attentional set-shifting trial, the EEDR group took longer than the control group to complete the trial. While this can be interpreted as an impairment in attention, or perhaps memory interference that obstructs recollection of the rule, it can also be a lack of motivation, heightened anxiety or stress. Generally, physical metrics were less noisy than

behavioural metrics, except for grip strength. The grip strength test was implemented using a bar that the subject would hold on to, before being pulled horizontally away from the apparatus. However, this required cooperation on the part of the subject, and that was not guaranteed. Adaption of the grip strength test by rotating the apparatus vertically was found to improve results in mice (Takeshita et al., 2017), but due to the size difference, this has not yet been implemented successfully in rats. Unlike the environmental factors, these differences are less easy to control and must be accepted, unless more specific tests can be conceived and developed.

Analysis

The impact of the so-called reproducibility crisis in the field of neuroscience, and neuroimaging more specifically, has been a driving force for reassessing the preprocessing methods and data analysis techniques that are applied to MRI data (Carp, 2012; Poldrack et al., 2017). Recently, the impact of researcher-specific choices in these methodological choices has been brought to light, with differences in processing pipelines between laboratories finding different results in the same dataset (Botvinik-Nezer et al., 2019). These problems are further exaggerated for preclinical MRI, where researchers build upon a wealth of software tools created for the processing and analysis of human MRI, but as these tools expect human MRI data as input, they are incompatible out-of-the-box with rodent data. Instead, preclinical researchers adjust existing tools or their data to align their methods with those used in human imaging which reduces transparency and the ability to generalise between preclinical studies. Care was taken in this pipeline to use data-agnostic tools and to minimise the requirement of adapting the data or the pipeline. *fMRIPrep* (Esteban et al., 2019) is a successful software tool which standardises the image preprocessing pipeline for humans and reduces incidences of parameter tuning by creating an automated pipeline based on the information provided by leveraging the brain imaging data structure compatible input files (Gorgolewski et al., 2016). A rodent MRI analogue, which explicitly addresses idiosyncrasies inherent to preclinical imaging, is currently in development and aims to improve reproducibility between

rodent studies while facilitating translation between preclinical and clinical MRI. However, more standardised preprocessing would not augment the issues in data analysis. Linear mixed-effects models were utilised in this study because of the large attrition of subject numbers in the later time points. However, this analysis has not yet been widely adopted by the neuroimaging community, which still instead implements repeated measures ANOVAs for mass univariate analysis, and consequently, voxelwise analysis across time was not completed for this work. Linear mixed-effects models have been previously implemented in a longitudinal study between groups (Qiu et al., 2018), by calculating the comparative fit of two models to assess the relative contributions of factors, rather than the *t* or *F*-values calculated in repeated measures ANOVAs. However, the statistics were done using RMINC package, which uses MINC files that are rich in meta-data but are less widely supported than NIfTI files (Vincent et al., 2016; Gorgolewski et al., 2016). Furthermore, it requires accurate normalisation of all time points to a single common space. While Qiu and colleagues achieved this with a sensible approach, it remains that each transformation that is concatenated further manipulates the image further from its original orientation, requiring interpolation of values, and adds error to the transformation. Therefore, it is not trivial to achieve a satisfactory registration to ensure adequate alignment of anatomical boundaries.

Finally, a commonly cited problem related to the reproducibility crisis is the focus on statistical *p*-values, and the exclusion of effect sizes (Sullivan & Feinn, 2012). In particular, reporting statistically significant values, but not necessarily biologically meaningful has obscured progress by focusing on findings that are essentially statistical red herrings. However, with that in mind, it is not clear how the graph theory metrics translate to the underlying biology. For example, while effect sizes can be quantified using statistics, it is unclear what a change in 0.1 increase in global efficiency means on a biological level.

Future directions

Possible alternative biomarkers

Future analyses for this dataset could apply a voxelwise mixed-effects regression approach as detailed therein, but further group comparisons could also replicate the same ‘life outcomes’ approach described previously (see: Are non-scanning metrics characteristic of life outcome?). Briefly, voxelwise comparisons can be made between the end of study group and the poor health group, but this would still be subject to the survivor bias that precluded the inclusion of the poor health group at the final time point. A third alternative analysis includes calculating the distribution of voxelwise quantitative metrics, as was leveraged in a recent study of microstructural integrity (Wearn et al., 2020). The distribution width of an individual’s T_2 values in the hippocampus and thalamus was a predictor of cognitive decline, but the median T_2 value was not. Moreover, multi-modal imaging was utilised here to extract the most information from the cohort as possible, since longitudinal studies are not common and expensive to run. However, it is likely that not all measures equally contribute to the variance underpinning age-related changes, and if this was to be translated to the clinic, an MRI protocol of this length is unlikely to be used beyond research. Therefore, future work could spend time determining which measures provide the most information. For instance, what are the minimum modalities required to delineate a person’s trajectory accurately? If these modalities, which are likely to incorporate elements of both brain structure and function, are collected, it is likely that a multi-modal imaging protocol of this length would be unnecessary.

In addition to the MRI data, blood samples, faecal samples, and brain tissue samples were collected from the Resilient cohort, and thus are imminent avenues for future work. The possibility of assessing peripheral inflammation markers in blood samples (Lim et al., 2013), or the histological verification of the microstructural changes observed in relaxometry are obvious next steps to take for the Resilient cohort. For example, histological staining for markers of myelin (e.g., Luxol Fast Blue or myelin basic protein) staining could assess whether the widespread increase in T_1 time in the

EEDR rats at 17 months was due to a relative increase of myelin, as the increased water content in the grey matter was ruled out by other metrics quantifying water, such as the diffusion scalars.

Another avenue for future work is another longitudinal study, but one that does not prescribe set end time points. The benefits to this would be twofold: first, it can confirm or reject the exploratory findings discussed in this work. Second, it would allow a characterisation based on the full lifespan of each animal in the cohort. It is unclear which of the animals in the ‘end of study’ cohort were truly super-agers, and which were declining, and determining the natural length of the lifespan would be beneficial for classifying the groups. Once these classifications are complete, animals can be classified by a continuum score which is included in the analysis as a covariate of interest.

The use of machine learning, a branch of artificial intelligence which allows computers to solve complex problems through repeated experience, has rapidly increased in the last two decades (Jordan & Mitchell, 2015) and their application in neuroscience is no exception. Brain age models are one avenue that is of particular interest: a model is developed by feeding in brains of different chronological ages, and the algorithm ‘learns’ what an old brain looks like (Franke & Gaser, 2019). When applying the trained model to new, unseen data, a predicted brain age is estimated, and, crucially, the difference between the predicted brain age and the chronological age is predictive of mortality in a longitudinal study of Scottish adults born in 1931 (Cole et al., 2018). Although this method has not been widely used in rodents, except for Franke et al., (2016), its applications on animal models of healthy and pathological ageing would provide a useful biomarker for drug development. One such study is currently taking place in our laboratory, using the Resilient data and an associated cohort.

Finally, as alluded to previously (see: The role of diet and enrichment in ageing), the potential impact of metabolism should be explored in future work. Aerobic glycolysis is traditionally measured using gold-standard methods, such as administering a

radioactive glucose analogue and measuring with Positron Emission Tomography (Vaishnavi et al., 2010), but this is an invasive and costly technique. However, future methods may leverage the innovative, non-invasive methods for quantifying metabolism with MRI, such as measuring the rate of oxygen metabolism using susceptibility-weighted MRI (Rodgers et al., 2016) or with asymmetric spin echo (Stone & Blockley, 2017), which has been extended for use in rodents by our group (Wood et al, KCL, in preparation). Oxygen metabolism could therefore be compared with glucose metabolism, measured using chemical exchange saturation transfer (CEST; Wang et al., 2016).

Large longitudinal studies are necessary to account for individual variance

Individual differences are ubiquitous. As they can be driven by genetics, environment, or an interaction between the two, individual ageing trajectories are likely to be incredibly idiosyncratic. Moreover, even in a study with rats, which are more similar to each other than an analogous human study's participants, which took place in an experimentally controlled home environment, there was still a substantial impact of non-shared environmental factors. Examples of such factors that have previously been described in rats include the individual variation in weight associated with litter size (Wurtman & Miller, 1976) and the differences in maternal care which is predictive of exploratory behaviour preferences (van Hasselt et al., 2012). These examples are far from exhaustive, and therefore countless possible sources drive individual differences and cannot reasonably be controlled. Therefore, if accounting for genetics and shared environment does not explain the majority of the individual variation in ageing trajectories of rats, this is evidence of how difficult it is to achieve a similar human study, given the complicated and complex lives of humans, which cannot realistically control for every environmental and genetic factor. While longitudinal studies are susceptible to learning effects in behaviour testing, for measurements like resting-state fMRI, dense subject sampling has shown that information is lost when using group averages (Poldrack, 2017; Gordon et al., 2017),

which could be critical to understanding individual ageing trajectories. Although structural imaging gains some benefits from the use of group averages, it is more likely that the individual variation from the average (e.g., brain age difference) is predictive of ageing outcomes, rather than the age or group mean.

A potential experiment to examine the effect of non-shared environment could leverage the approach used to model resting-state MRI: by modelling the variance of all known sources of variation, such as using individualistic, longitudinal studies which can reduce between-subject variance, using shared enriched homes to control for environmental effects, and using genetically similar subjects, only the variance caused by non-shared environment should remain. In this idealised experiment, the relative contributions of aerobic glycolysis and oxidative phosphorylation can be assessed on the growth of brain areas using CEST and susceptibility MRI. Furthermore, the hypothesis that ketosis and environmental enrichment in early adulthood ameliorate rapid brain growth relative to glucose and a sedentary lifestyle can be tested. Glucose and oxygen metabolism can be measured over time to assess the sustainability of the energy requirements, particularly in functionally integrative regions, such as the default mode network, which have high energy requirements. Finally, evidence of a compromise between myelin maintenance and neuroplasticity as a consequence of age-related hypometabolism can be tested with histology.

For this shift to idiosyncratic individual trajectories to be adopted, it also requires a change in perspective for ageing research: the most meaningful results are not likely to come from studies which compare ‘young’ versus ‘old’, but from the exploration of the non-linear, life-long trajectories within-subjects, which sees development and ageing as a continuous process and ageing outcomes as a continuum. These studies are not easier to run or cheaper to fund than their cross-sectional counterparts, but by controlling for individual differences, these methods only can expose the information necessary to disentangle the continuum of cognitive ageing.

REFERENCES

- Achard, S. & Bullmore, E. (2007) Efficiency and cost of economical brain functional networks. *PLoS Computational Biology*. 3 (2), 0174–0183.
- Aksenov, D. P., Li, L., Miller, M. J., Iordanescu, G. & Wyrwicz, A. M. (2015) Effects of anesthesia on BOLD signal and neuronal activity in the somatosensory cortex. *Journal of Cerebral Blood Flow & Metabolism*. 351819–1826.
- Albrecht, M., Henke, J., Tacke, S., Markert, M. & Guth, B. (2014) Influence of repeated anaesthesia on physiological parameters in male Wistar rats: A telemetric study about isoflurane, ketamine-xylazine and a combination of medetomidine, midazolam and fentanyl. *BMC Veterinary Research*. 10 (1), 310.
- Alexander-Bloch, A., Giedd, J. N. & Bullmore, E. (2013) Imaging structural covariance between human brain regions. *Nature Reviews Neuroscience* 14 (5) p.322–336.
- Alexander, A. G., Marfil, V. & Li, C. (2014) Use of *Caenorhabditis elegans* as a model to study Alzheimer’s disease and other neurodegenerative diseases. *Frontiers in Genetics*. 5 (JUL), 279.
- Alexander, A. L., Lee, J. E., Lazar, M. & Field, A. S. (2007) Diffusion Tensor Imaging of the Brain. *Neurotherapeutics*. 4 (3), 316–329.
- Allen, E. A., Erhardt, E. B. & Calhoun, V. D. (2012) Data visualization in the neurosciences: overcoming the curse of dimensionality. *Neuron*. 74 (4), 603–608.
- Alsop, D. C., Detre, J. A., Golay, X., Günther, M., Hendrikse, J., Hernandez-Garcia, L., Lu, H., Macintosh, B. J., Parkes, L. M., Smits, M., Van Osch, M. J. P., Wang, D. J. J., Wong, E. C. & Zaharchuk, G. (2015) Recommended implementation of arterial spin-labeled Perfusion mri for clinical applications: A consensus of the ISMRM Perfusion Study group and the European consortium for ASL in dementia. *Magnetic Resonance in Medicine*. 73 (1), 102–116.
- Ances, B. M., Liang, C. L., Leontiev, O., Perthen, J. E., Fleisher, A. S., Lansing, A.

- E. & Buxton, R. B. (2009) Effects of aging on cerebral blood flow, oxygen metabolism, and blood oxygenation level dependent responses to visual stimulation. *Human Brain Mapping*. 30 (4), 1120–1132.
- Anderson, S. W., Damasio, H., Jones, R. D. & Tranel, D. (1991) Wisconsin card sorting test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology*. 13 (6), 909–922.
- Andersson, J. L. R., Skare, S. & Ashburner, J. (2003) How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*. 20 (2), 870–888.
- 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*. 1251063–1078.
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R. & Buckner, R. L. (2010) Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*. 65 (4), 550–562.
- Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E. & Buckner, R. L. (2007) Disruption of Large-Scale Brain Systems in Advanced Aging. *Neuron*. 56 (5), 924–935.
- Antunes, M. & Biala, G. (2012) The novel object recognition memory: neurobiology, test procedure, and its modifications. *Cognitive Processing*. 13 (2), 93–110.
- Arnsten, A. F. T., Datta, D., Leslie, S., Yang, S. T., Wang, M. & Nairn, A. C. (2019) Alzheimer's-like pathology in aging rhesus macaques: Unique opportunity to study the etiology and treatment of Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America*. 116 (52), 26230–26238.
- Ash, J. A., Lu, H., Taxier, L. R., Long, J. M., Yang, Y., Stein, E. A. & Rapp, P. R. (2016) Functional connectivity with the retrosplenial cortex predicts cognitive aging in rats. *Proceedings of the National Academy of Sciences of the United States of*

- America*. 113 (43), 12286–12291.
- Ashburner, J. & Friston, K. (2003) ‘Spatial Normalisation Using Basis Functions’, in *Human Brain Function: Second Edition*. pp. 655–672.
- Ashburner, J. & Friston, K. J. (2001) Why Voxel-based morphometry should be used. *NeuroImage* 14 (6) p.1238–1243.
- Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. (2008) Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*. 12 (1), 26–41.
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A. & Gee, J. C. (2011a) A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*. 54 (3), 2033–2044.
- Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A. & Gee, J. C. (2011b) An open source multivariate framework for N-tissue segmentation with evaluation on public data. *Neuroinformatics*. 9 (4), 381–400.
- Avants, B. B., Yushkevich, P., Pluta, J., Minkoff, D., Korczykowski, M., Detre, J. & Gee, J. C. (2010) The optimal template effect in hippocampus studies of diseased populations. *NeuroImage*. 49 (3), 2457–2466.
- Bagnato, F., Hametner, S. & Welch, E. B. (2013) Visualizing iron in multiple sclerosis. *Magnetic Resonance Imaging* 31 (3) p.376–384.
- Barnes, C. A., Nadel, L. & Honig, W. K. (1980) Spatial memory deficit in senescent rats. *Canadian journal of psychology*. 34 (1), 29–39.
- Bartzokis, G., Cummings, J. L., Sultzer, D., Henderson, V. W., Nuechterlein, K. H. & Mintz, J. (2003) White matter structural integrity in healthy aging adults and patients with Alzheimer disease: A magnetic resonance imaging study. *Archives of Neurology*. 60 (3), 393–398.
- Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J. & Aldroubi, A. (2000) In vivo fiber

- tractography using DT-MRI data. *Magnetic Resonance in Medicine*. 44 (4), 625–632.
- Bassett, D. S. & Bullmore, E. (2006) Small-World Brain Networks. *The Neuroscientist*. 12 (6), 512–523.
- Bassett, D. S. & Bullmore, E. T. (2017) Small-World Brain Networks Revisited. *The Neuroscientist*. 23 (5), 499–516.
- Bassett, D. S., Bullmore, E. T., Meyer-Lindenberg, A., Apud, J. A., Weinberger, D. R. & Coppola, R. (2009) Cognitive fitness of cost-efficient brain functional networks. PNAS July 14.
- Bassett, D. S. & Sporns, O. (2017) Network neuroscience. *Nature Neuroscience*. 20 (3), 353–364.
- Baxter, M. G. (2009) ‘Age-Related Effects on Prefrontal Cortical Systems: Translating Between Rodents, Nonhuman Primates, and Humans’, in *Animal Models of Human Cognitive Aging*. Totowa, NJ: Humana Press. pp. 1–14.
- Baxter, M. G., Patrick, R. H. & Charles, V. M. (2001) ‘Cognitive Aging in Nonhuman Primates’, in *Functional Neurobiology of Aging*. pp. 407–419.
- Beard, J. R., Officer, A., De Carvalho, I. A., Sadana, R., Pot, A. M., Michel, J. P., Lloyd-Sherlock, P., Epping-Jordan, J. E., Peeters, G. M. E. E., Mahanani, W. R., Thiyagarajan, J. A. & Chatterji, S. (2016) The World report on ageing and health: A policy framework for healthy ageing. *The Lancet*. 387 (10033), 2145–2154.
- Beas, B. S., Setlow, B. & Bizon, J. L. (2013) Distinct manifestations of executive dysfunction in aged rats. *Neurobiology of Aging*. 34 (9), 2164–2174.
- Beason-Held, L. L., Kraut, M. A. & Resnick, S. M. (2009) Stability of Default-Mode Network Activity in the Aging Brain. *Brain Imaging and Behavior*. 3 (2), 123–131.
- Becq, G. J. P. C., Barbier, E. L. & Achard, S. (2020) Brain networks of rats under anesthesia using resting-state fMRI: Comparison with dead rats, random noise and generative models of networks. *Journal of Neural Engineering*. 17 (4), .
- Belliveau, J. W., Kennedy, D. N., McKinstry, R. C., Buchbinder, B. R., Weisskoff, R.

- M., Cohen, M. S., Vevea, J. M., Brady, T. J. & Rosen, B. R. (1991) Functional mapping of the human visual cortex by magnetic resonance imaging. *Science (New York, N.Y.)*. 254 (5032), 716–719.
- Berg, E. A. (1948) A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*. 39 (1), 15–22.
- Bilkei-Gorzo, A. (2014) Genetic mouse models of brain ageing and Alzheimer's disease. *Pharmacology & Therapeutics*. 142 (2), 244–257.
- Birrell, J. M. & Brown, V. J. (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. *Journal of Neuroscience*. 20 (11), 4320–4324.
- Bizon, J. L., Foster, T. C., Alexander, G. E. & Glisky, E. L. (2012) Characterizing cognitive aging of working memory and executive function in animal models. *Frontiers in Aging Neuroscience*. 4 (SEP), 1–14.
- Bloch, F. (1946) Nuclear Induction. *Physical Review*. 70 (7–8), 460–474.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*. 2008 (10), P10008.
- Bohannon, R. W. (2019) Grip Strength: An Indispensable Biomarker For Older Adults. *Clinical interventions in aging*. 141681–1691.
- Botvinik-Nezer, R., Iwanir, R., Holzmeister, F., Huber, J., Johannesson, M., Kirchler, M., Dreber, A., Camerer, C. F., Poldrack, R. A. & Schonberg, T. (2019) fMRI data of mixed gambles from the Neuroimaging Analysis Replication and Prediction Study. *Scientific Data*. 6 (1), 1–9.
- Brant-Zawadzki, M., Fein, G., Van Dyke, C., Kiernan, R., Davenport, L. & de Groot, J. (1985) MR imaging of the aging brain: patchy white-matter lesions and dementia. *AJNR. American journal of neuroradiology*. 6 (5), 675–682.
- Bren, K. L., Eisenberg, R. & Gray, H. B. (2015) Discovery of the magnetic behavior of hemoglobin: A beginning of bioinorganic chemistry. *Proceedings of the National*

- Academy of Sciences of the United States of America*. 112 (43), 13123–13127.
- Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y. Y., Weber, J. & Kober, H. (2011) Meditation experience is associated with differences in default mode network activity and connectivity. *Proceedings of the National Academy of Sciences of the United States of America*. 108 (50), 20254–20259.
- Brioni, J. D., Varughese, S., Ahmed, R. & Bein, B. (2017) A clinical review of inhalation anesthesia with sevoflurane: from early research to emerging topics. *Journal of Anesthesia*. 31 (5), 764–778.
- Broadbent, N. J., Gaskin, S., Squire, L. R. & Clark, R. E. (2010) Object recognition memory and the rodent hippocampus. *Learning & Memory*. 17 (1), 5–11.
- Brower, M., Grace, M., Kotz, C. M. & Koya, V. (2015) Comparative analysis of growth characteristics of Sprague Dawley rats obtained from different sources. *Laboratory animal research*. 31 (4), 166–173.
- Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. (2008) The Brain's Default Network. *Annals of the New York Academy of Sciences*. 1124 (1), 1–38.
- Bullmore, E. & Sporns, O. (2012) The economy of brain network organization. *Nature Reviews Neuroscience*. 13 (5), 336–349.
- Bullmore, E. T. & Sporns, O. (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*. 10 (3), 186–198.
- Burke, S. N. & Barnes, C. A. (2006) Neural plasticity in the ageing brain. *Nature Reviews Neuroscience*. 7 (1), 30–40.
- Burke, S. N., Ryan, L. & Barnes, C. A. (2012) Characterizing cognitive aging of recognition memory and related processes in animal models and in humans. *Frontiers in Aging Neuroscience*. 4 (SEP), 15.
- Burzynska, A. Z., Preuschhof, C., Bäckman, L., Nyberg, L., Li, S. C., Lindenberger, U. & Heekeren, H. R. (2010) Age-related differences in white matter

- microstructure: Region-specific patterns of diffusivity. *NeuroImage*. 49 (3), 2104–2112.
- Cabeza, R. (2002) Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*. 17 (1), 85–100.
- Cabeza, R., Anderson, N. D., Locantore, J. K. & McIntosh, A. R. (2002) Aging Gracefully: Compensatory Brain Activity in High-Performing Older Adults. *NeuroImage*. 17 (3), 1394–1402.
- Cacioppo, S., Capitanio, J. P. & Cacioppo, J. T. (2014) Toward a neurology of loneliness. *Psychological Bulletin*. 140 (6), 1464–1504.
- Calabrese, E., Badea, A., Watson, C. & Johnson, G. A. (2013) A quantitative magnetic resonance histology atlas of postnatal rat brain development with regional estimates of growth and variability. *NeuroImage*. 71 196–206.
- Cao, M., Wang, J. H., Dai, Z. J., Cao, X. Y., Jiang, L. L., Fan, F. M., Song, X. W., Xia, M. R., Shu, N., Dong, Q., Milham, M. P., Castellanos, F. X., Zuo, X. N. & He, Y. (2014) Topological organization of the human brain functional connectome across the lifespan. *Developmental Cognitive Neuroscience*. 7 76–93.
- Carlson, A. J. & Hoelzel, F. (1946) Apparent prolongation of the life span of rats by intermittent fasting. *The Journal of nutrition*. 31 363–375.
- Carp, J. (2012) The secret lives of experiments: Methods reporting in the fMRI literature. *NeuroImage*. 63 (1), 289–300.
- Carter, C. S., Richardson, A., Huffman, D. M. & Austad, S. (2020) Bring Back the Rat! Rozalyn Anderson (ed.). *The Journals of Gerontology: Series A*. 75 (3), 405–415.
- Chan, M. Y., Park, D. C., Savalia, N. K., Petersen, S. E. & Wig, G. S. (2014) Decreased segregation of brain systems across the healthy adult lifespan. *Proceedings of the National Academy of Sciences*. 111 (46), E4997–E5006.
- Chang, P.-C., Procissi, D., Bao, Q., Centeno, M. V., Baria, A. & Apkarian, A. V. (2016) Novel method for functional brain imaging in awake minimally restrained

- rats. *Journal of Neurophysiology*. 116 (1), 61–80.
- Chemali, J. J., Kenny, J. D., Olutola, O., Taylor, N. E., Kimchi, E. Y., Purdon, P. L., Brown, E. N. & Solt, K. (2015) Ageing delays emergence from general anaesthesia in rats by increasing anaesthetic sensitivity in the brain. *British Journal of Anaesthesia*. 115 (Suppl 1), i58–i65.
- Chen, J. J., Rosas, H. D. & Salat, D. H. (2011a) Age-associated reductions in cerebral blood flow are independent from regional atrophy. *NeuroImage*. 55 (2), 468–478.
- Chen, Z. J., He, Y., Rosa-Neto, P., Gong, G. & Evans, A. C. (2011b) Age-related alterations in the modular organization of structural cortical network by using cortical thickness from MRI. *NeuroImage*. 56 (1), 235–245.
- Chou, N., Wu, J., Bai Bingren, J., Qiu, A. & Chuang, K. H. (2011) Robust automatic rodent brain extraction using 3-D pulse-coupled neural networks (PCNN). *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*. 20 (9), 2554–2564.
- Chung, M. K., Worsley, K. J., Paus, T., Cherif, C., Collins, D. L., Giedd, J. N., Rapoport, J. L. & Evans, A. C. (2001) A unified statistical approach to deformation-based morphometry. *NeuroImage*. 14 (3), 595–606.
- Churchill, N. W., Caverzasi, E., Graham, S. J., Hutchison, M. G. & Schweizer, T. A. (2019) White matter during concussion recovery: Comparing diffusion tensor imaging (DTI) and neurite orientation dispersion and density imaging (NODDI). *Human Brain Mapping*. 40 (6), 1908–1918.
- Clarke, D. D. & Sokoloff, L. (1999) ‘Regulation of Cerebral Metabolic Rate.’, in George J Siegel, Bernard W Agranoff, R Wayne Albers, Stephen K Fisher, & Michael D Uhler (eds.) *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*. 6th edition Philadelphia: Lippincott-Raven. p.
- Cohen-Adad, J. (2014) What can we learn from T2* maps of the cortex? *NeuroImage* 93 p.189–200.

- Cohen, S. J. & Stackman Jr., R. W. (2015) Assessing rodent hippocampal involvement in the novel object recognition task. A review. *Behavioural Brain Research*. 285105–117.
- Cole, J. H., Ritchie, S. J., Bastin, M. E., Valdés Hernández, M. C., Muñoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S. E., Zhang, Q., Wray, N. R., Redmond, P., Marioni, R. E., Starr, J. M., Cox, S. R., Wardlaw, J. M., Sharp, D. J. & Deary, I. J. (2018) Brain age predicts mortality. *Molecular Psychiatry*. 23 (5), 1385–1392.
- Collard, R. M., Boter, H., Schoevers, R. A. & Oude Voshaar, R. C. (2012) Prevalence of frailty in community-dwelling older persons: A systematic review. *Journal of the American Geriatrics Society*. 60 (8), 1487–1492.
- Cox, R. W. (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 29 (3), 162–173.
- Csardi, G. & Nepusz, T. (2006) The igraph software package for complex network research. *InterJournal Complex Systems*. Complex Sy (1695), 1695.
- Damoiseaux, J. S. (2017) Effects of aging on functional and structural brain connectivity. *NeuroImage*. 16032–40.
- Damoiseaux, J. S., Beckmann, C. F., Arigita, E. J. S., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M. & Rombouts, S. A. R. B. (2008) Reduced resting-state brain activity in the ‘default network’ in normal aging. *Cerebral Cortex*. 18 (8), 1856–1864.
- Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M. & Beckmann, C. F. (2006) Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*. 103 (37), 13848–13853.
- Davis, S. W., Dennis, N. a, Daselaar, S. M., Fleck, M. S. & Cabeza, R. (2008) Que

- PASA? The Posterior-Anterior Shift in Aging. *Cerebral Cortex*. 18 (5), 1201–1209.
- Deoni, S. C. L. (2010) Quantitative Relaxometry of the Brain. *Topics in Magnetic Resonance Imaging*. 21 (2), 101–113.
- Depp, C. A. & Jeste, D. V (2006) Definitions and Predictors of Successful Aging: A Comprehensive Review of Larger Quantitative Studies. *The American Journal of Geriatric Psychiatry*. 14 (1), 6–20.
- Desjardins, M., Berti, R., Pouliot, P., Dubeau, S. & Lesage, F. (2014) Multimodal study of the hemodynamic response to hypercapnia in anesthetized aged rats. *Neuroscience Letters*. 56333–37.
- Dobre, M. C., Uğurbil, K. & Marjanska, M. (2007) Determination of blood longitudinal relaxation time (T1) at high magnetic field strengths. *Magnetic Resonance Imaging*. 25 (5), 733–735.
- DuPre, E., Salo, T., Markello, R., Kundu, P., Whitaker, K. & Handwerker, D. (2020) *ME-ICA/tedana: 0.0.9a*. [online]. Available from: <https://zenodo.org/record/3786890> (Accessed 17 August 2020).
- Edler, M. K., Sherwood, C. C., Meindl, R. S., Hopkins, W. D., Ely, J. J., Erwin, J. M., Mufson, E. J., Hof, P. R. & Raghanti, M. A. (2017) Aged chimpanzees exhibit pathologic hallmarks of Alzheimer’s disease. *Neurobiology of Aging*. 59107–120.
- Ellenbroek, B. & Youn, J. (2016) Rodent models in neuroscience research: Is it a rat race? *DMM Disease Models and Mechanisms*. 9 (10), 1079–1087.
- Ennaceur, A., Neave, N. & Aggleton, J. P. (1997) Spontaneous object recognition and object location memory in rats: The effects of lesions in the cingulate cortices, the medial prefrontal cortex, the cingulum bundle and the fornix. *Experimental Brain Research*. 113 (3), 509–519.
- Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A. & Gorgolewski, K. J. (2019) fMRIPrep: a

- robust preprocessing pipeline for functional MRI. *Nature Methods*. 16 (1), 111–116.
- Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Gunther, M., Glasser, M. F., Miller, K. L., Ugurbil, K. & Yacoub, E. (2010) Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging Pedro Antonio Valdes-Sosa (ed.). *PLoS ONE*. 5 (12), e15710.
- Feo, R. & Giove, F. (2019) Towards an efficient segmentation of small rodents brain: A short critical review. *Journal of Neuroscience Methods* 323 p.82–89.
- Fillmore, P. T., Phillips-Meek, M. C. & Richards, J. E. (2015) Age-specific MRI brain and head templates for healthy adults from 20 through 89 years of age. *Frontiers in Aging Neuroscience*. 7 (MAR), 44.
- Fjell, A. M., McEvoy, L., Holland, D., Dale, A. M. & Walhovd, K. B. (2013) Brain changes in older adults at very low risk for Alzheimer’s disease. *Journal of Neuroscience*. 33 (19), 8237–8242.
- Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., Brewer, J. B. & Dale, A. M. (2009) One-year brain atrophy evident in healthy aging. *Journal of Neuroscience*. 29 (48), 15223–15231.
- Flood, D. G. & Coleman, P. D. (1988) Neuron numbers and sizes in aging brain: Comparisons of human, monkey, and rodent data. *Neurobiology of Aging*. 9 (C), 453–463.
- Fornito, A., Zalesky, A. & Breakspear, M. (2013) Graph analysis of the human connectome: Promise, progress, and pitfalls. *NeuroImage*. 80426–444.
- Fornito, A., Zalesky, A. & Bullmore, E. T. (2016) *Fundamentals of Brain Network Analysis*. Elsevier Inc.
- Fox, J. G. & Bennett, B. T. (2015) ‘Chapter 1 - Laboratory Animal Medicine: Historical Perspectives’, in James Fox, Lynn Anderson, Glen Otto, Kathleen Pritchett-Corning, & Mark Whary (eds.) *Laboratory Animal Medicine: Third Edition*.

- 3rd edition Elsevier. pp. 1–21.
- Fox, P. T. & Friston, K. J. (2012) Distributed processing; distributed functions? *NeuroImage*. 61 (2), 407–426.
- Franceschi, C., Garagnani, P., Morsiani, C., Conte, M., Santoro, A., Grignolio, A., Monti, D., Capri, M. & Salvioli, S. (2018) The continuum of aging and age-related diseases: Common mechanisms but different rates. *Frontiers in Medicine* 5 (MAR).
- Franke, K. & Gaser, C. (2019) Ten years of brainage as a neuroimaging biomarker of brain aging: What insights have we gained? *Frontiers in Neurology* [online]. Available from: www.frontiersin.org (Accessed 14 December 2020). 10 (JUL) p.789.
- Frankea, K., Dahnke, R., Clarke, G., Kuo, A., Li, C., Nathanielsz, P., Schwab, M. & Gaser, C. (2016) MRI based biomarker for brain aging in rodents and non-human primates. *PRNI 2016 - 6th International Workshop on Pattern Recognition in Neuroimaging*. 1–4.
- Freund, J., Brandmaier, A. M., Lewejohann, L., Kirste, I., Kritzler, M., Kruger, A., Sachser, N., Lindenberger, U. & Kempermann, G. (2013) Emergence of Individuality in Genetically Identical Mice. *Science*. 340 (6133), 756–759.
- Freund, J., Brandmaier, A. M., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberger, U. & Kempermann, G. (2015) Association between exploratory activity and social individuality in genetically identical mice living in the same enriched environment. *Neuroscience*. 309140–152.
- Friston, K. J. (1994) Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*. 2 (1–2), 56–78.
- Gage, F. H., Dunnett, S. B. & Björklund, A. (1984) Spatial learning and motor deficits in aged rats. *Neurobiology of Aging*. 5 (1), 43–48.
- Gallagher, M. & Burwell, R. D. (1989) Relationship of age-related decline across

- several behavioral domains. *Neurobiology of Aging*. 10 (6), 691–708.
- Gallagher, M. & Nicolle, M. M. (1993) Animal models of normal aging: Relationship between cognitive decline and markers in hippocampal circuitry. *Behavioural Brain Research*. 57 (2), 155–162.
- Galvan, V. & Jin, K. (2007) Neurogenesis in the aging brain. *Clinical interventions in aging*. 2 (4), 605–610.
- Gámiz, F. & Gallo, M. (2012) Spontaneous object recognition memory in aged rats: Complexity versus similarity. *Learning and Memory*. 19 (10), 444–448.
- Geerligs, L., Renken, R. J., Saliassi, E., Maurits, N. M. & Lorist, M. M. (2015) A Brain-Wide Study of Age-Related Changes in Functional Connectivity. *Cerebral Cortex*. 25 (7), 1987–1999.
- Geinisman, Y. & Bondareff, W. (1976) Decrease in the number of synapses in the senescent brain: A quantitative electron microscopic analysis of the dentate gyrus molecular layer in the rat. *Mechanisms of Ageing and Development*. 5 (C), 11–23.
- Geinisman, Y., De Toledo-Morrell, L. & Morrell, F. (1986) Loss of perforated synapses in the dentate gyrus: Morphological substrate of memory deficit in aged rats. *Proceedings of the National Academy of Sciences of the United States of America*. 83 (9), 3027–3031.
- Geng, Y. Q., Guan, J. T., Xu, M. Y., Xu, X. H. & Fu, Y. C. (2007) Behavioral study of calorie-restricted rats from early old age. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. 5150312393–2395.
- Gerstorf, D., Hülür, G., Drewelies, J., Eibich, P., Duezel, S., Demuth, I., Ghisletta, P., Steinhagen-Thiessen, E., Wagner, G. G. & Lindenberger, U. (2015) Secular changes in late-life cognition and well-being: Towards a long bright future with a short brisk ending? *Psychology and Aging*. 30 (2), 301–310.
- Gill, T. J., Smith, G. J., Wissler, R. W. & Kunz, H. W. (1989) The Rat as an

- Experimental Animal. Source: Science, New Series 245 (4915).
- Ginestet, C. E., Nichols, T. E., Bullmore, E. T. & Simmons, A. (2011) Brain network analysis: Separating cost from topology using Cost-Integration. *PLoS ONE*. 6 (7), .
- Glasser, M. F., Goyal, M. S., Preuss, T. M., Raichle, M. E. & Van Essen, D. C. (2014) Trends and properties of human cerebral cortex: Correlations with cortical myelin content. *NeuroImage* 93 p.165–175.
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., Van Essen, D. C. & Jenkinson, M. (2013) The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*. 80105–124.
- Goerzen, D., Fowler, C., Devenyi, G. A., Germann, J., Madularu, D., Chakravarty, M. M. & Near, J. (2020) An MRI-Derived Neuroanatomical Atlas of the Fischer 344 Rat Brain. *Scientific Reports*. 10 (1), 6952.
- Golde, T. E., Schneider, L. S. & Koo, E. H. (2011) Anti-A β therapeutics in alzheimer's disease: The need for a paradigm shift. *Neuron*. 69 (2), 203–213.
- Gong, G., Rosa-Neto, P., Carbonell, F., Chen, Z. J., He, Y. & Evans, A. C. (2009) Age- and gender-related differences in the cortical anatomical network. *Journal of Neuroscience*. 29 (50), 15684–15693.
- Gonzalez-Lima, F., Ferchmin, P. A., Eterovic, V. A. & Gonzalez-Lima, E. M. (1994) Metabolic activation of the brain of young rats after exposure to environmental complexity. *Developmental Psychobiology*. 27 (6), 343–351.
- Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J. & Frackowiak, R. S. J. (2001) A Voxel-Based Morphometric Study of Ageing in 465 Normal Adult Human Brains. *NeuroImage*. 14 (1), 21–36.
- Gordon, E. M., Laumann, T. O., Gilmore, A. W., et al. (2017) Precision Functional Mapping of Individual Human Brains. *Neuron*. 95 (4), 791-807.e7.

- Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L. & Ghosh, S. S. (2011) Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*. 513.
- Gorgolewski, K. J., Auer, T., Calhoun, V. D., et al. (2016) The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*. 3 (1), 1–9.
- Goulas, A., Uylings, H. B. M. & Hilgetag, C. C. (2017) Principles of ipsilateral and contralateral cortico-cortical connectivity in the mouse. *Brain Structure and Function*. 222 (3), 1281–1295.
- Goyal, M. S., Hawrylycz, M., Miller, J. A., Snyder, A. Z. & Raichle, M. E. (2014) Aerobic glycolysis in the human brain is associated with development and neotenus gene expression. *Cell Metabolism*. 19 (1), 49–57.
- Gozzi, A. & Schwarz, A. J. (2016) Large-scale functional connectivity networks in the rodent brain. *NeuroImage*. 127496–509.
- Grady, C. L., McIntosh, A. R., Bookstein, F., Horwitz, B., Rapoport, S. I. & Haxby, J. V (1998) Age-Related Changes in Regional Cerebral Blood Flow during Working Memory for Faces. *NeuroImage*. 8 (4), 409–425.
- Grady, C., Sarraf, S., Saverino, C. & Campbell, K. (2016) Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks. *Neurobiology of Aging*. 41159–172.
- Grandjean, J., Schroeter, A., Batata, I. & Rudin, M. (2014) Optimization of anesthesia protocol for resting-state fMRI in mice based on differential effects of anesthetics on functional connectivity patterns. *NeuroImage*. 102 (P2), 838–847.
- Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. (2003) Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of*

- America*. 100 (1), 253–258.
- Guerin, M. V., Finisguerra, V., Van den Eynde, B. J., Bercovici, N. & Trautmann, A. (2020) Preclinical murine tumor models: A structural and functional perspective. *eLife* 9.
- Gulinello, M., Mitchell, H. A., Chang, Q., Timothy O'Brien, W., Zhou, Z., Abel, T., Wang, L., Corbin, J. G., Veeraragavan, S., Samaco, R. C., Andrews, N. A., Fagiolini, M., Cole, T. B., Burbacher, T. M. & Crawley, J. N. (2019) Rigor and reproducibility in rodent behavioral research. *Neurobiology of Learning and Memory*. 165106780.
- Gunning-Dixon, F. M. & Raz, N. (2000) The cognitive correlates of white matter abnormalities in normal aging: A quantitative review. *Neuropsychology*. 14 (2), 224–232.
- Guo, J., Bakshi, V. & Lin, A.-L. (2015) Early Shifts of Brain Metabolism by Caloric Restriction Preserve White Matter Integrity and Long-Term Memory in Aging Mice. *Frontiers in aging neuroscience*. 7 (NOV), 213.
- Gur, R. C., Gur, R. E., Obrist, W. D., Skolnick, B. E. & Reivich, M. (1987) Age and Regional Cerebral Blood Flow at Rest and During Cognitive Activity. *Archives of General Psychiatry*. 44 (7), 617–621.
- Gusnard, D. A. & Raichle, M. E. (2001) Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience*. 2 (10), 685–694.
- Haacke, E. M., Cheng, N. Y. C., House, M. J., Liu, Q., Neelavalli, J., Ogg, R. J., Khan, A., Ayaz, M., Kirsch, W. & Obenaus, A. (2005) Imaging iron stores in the brain using magnetic resonance imaging. *Magnetic Resonance Imaging*. 23 (1), 1–25.
- Hales, K. G., Korey, C. A., Larracuente, A. M. & Roberts, D. M. (2015) Genetics on the Fly: A Primer on the *Drosophila* Model System. *Genetics*. 201 (3), 815–842.
- Hall, C. & Ballachey, E. L. (1932) A study of the rat's behavior in a field. A contribution to method in comparative psychology. *University of California*

Publications in Psychology. 61–12.

- Haller, S., Zaharchuk, G., Thomas, D. L., Lovblad, K. O., Barkhof, F. & Golay, X. (2016) Arterial spin labeling perfusion of the brain: Emerging clinical applications. *Radiology* 281 (2) p.337–356.
- Hallgren, B. & Sourander, P. (1958) The effect of age on the non-haemin iron in the human brain. *Journal of Neurochemistry*. 3 (1), 41–51.
- Hamezah, H. S., Durani, L. W., Ibrahim, N. F., Yanagisawa, D., Kato, T., Shiino, A., Tanaka, S., Damanhuri, H. A., Ngah, W. Z. W. & Tooyama, I. (2017) Volumetric changes in the aging rat brain and its impact on cognitive and locomotor functions. *Experimental Gerontology*. 9969–79.
- Hao, X., Xu, D., Bansal, R., Dong, Z., Liu, J., Wang, Z., Kangarlu, A., Liu, F., Duan, Y., Shova, S., Gerber, A. J. & Peterson, B. S. (2013) Multimodal magnetic resonance imaging: The coordinated use of multiple, mutually informative probes to understand brain structure and function. *Human Brain Mapping*. 34 (2), 253–271.
- Harada, C. N., Natelson Love, M. C. & Triebel, K. L. (2013) Normal Cognitive Aging. *Clinics in Geriatric Medicine*. 29 (4), 737–752.
- Hardy, J. & Higgins, G. (1992) Alzheimer's disease: the amyloid cascade hypothesis. *Science*. 256 (5054), 184–185.
- Harris, J. J. & Attwell, D. (2012) The energetics of CNS white matter. *Journal of Neuroscience*. 32 (1), 356–371.
- van Hasselt, F. N., Tieskens, J. M., Trezza, V., Krugers, H. J., Vanderschuren, L. J. M. J. & Joëls, M. (2012) Within-litter variation in maternal care received by individual pups correlates with adolescent social play behavior in male rats. *Physiology and Behavior*. 106 (5), 701–706.
- Head, D., Kennedy, K. M., Rodrigue, K. M. & Raz, N. (2009) Age differences in perseveration: Cognitive and neuroanatomical mediators of performance on the

- Wisconsin Card Sorting Test. *Neuropsychologia*. 47 (4), 1200–1203.
- Hedden, T. & Gabrieli, J. D. E. (2004) Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience* 5 (2) p.87–96.
- Hedden, T., Oh, H., Younger, A. P. & Patel, T. A. (2013) Meta-analysis of amyloid-cognition relations in cognitively normal older adults. *Neurology*. 80 (14), 1341–1348.
- Henstridge, C. M., Jackson, R. J., Kim, J. M., Herrmann, A. G., Wright, A. K., Harris, S. E., Bastin, M. E., Starr, J. M., Wardlaw, J., Gillingwater, T. H., Smith, C., McKenzie, C.-A., Cox, S. R., Deary, I. J. & Spires-Jones, T. L. (2015) Post-mortem brain analyses of the Lothian Birth Cohort 1936: extending lifetime cognitive and brain phenotyping to the level of the synapse. *Acta neuropathologica communications*. 353.
- van den Heuvel, M. P., de Lange, S. C., Zalesky, A., Seguin, C., Yeo, B. T. T. & Schmidt, R. (2017) Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations. *NeuroImage*. 152437–449.
- Hirschler, L., Munting, L. P., Khmelinskii, A., Teeuwisse, W. M., Suidgeest, E., Warnking, J. M., van der Weerd, L., Barbier, E. L. & van Osch, M. J. P. (2018) Transit time mapping in the mouse brain using time-encoded pCASL. *NMR in Biomedicine*. 31 (2), e3855.
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R. & Hagmann, P. (2009) Predicting human resting-state functional connectivity from structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*. 106 (6), 2035–2040.
- Hoshaw, B. A., Evans, J. C., Mueller, B., Valentino, R. J. & Lucki, I. (2006) Social competition in rats: Cell proliferation and behavior. *Behavioural Brain Research*. 175 (2), 343–351.

- Houdebine, L.-M. (2007) ‘Transgenic Animal Models in Biomedical Research’, in *Target Discovery and Validation Reviews and Protocols*. New Jersey: Humana Press. pp. 163–202.
- Hsu, L.-M., Liang, X., Gu, H., Brynildsen, J. K., Stark, J. A., Ash, J. A., Lin, C.-P., Lu, H., Rapp, P. R., Stein, E. A. & Yang, Y. (2016) Constituents and functional implications of the rat default mode network. *Proceedings of the National Academy of Sciences*. 113 (31), E4541–E4547.
- Hsu, L.-M., Wang, S., Ranadive, P., Ban, W., Chao, T.-H. H., Song, S., Cerri, D. H., Walton, L. R., Broadwater, M. A., Lee, S.-H., Shen, D. & Shih, Y.-Y. I. (2020) Automatic Skull Stripping of Rat and Mouse Brain MRI Data Using U-Net. *Frontiers in Neuroscience*. 14568614.
- Huang, C. M., Lee, S. H., Hsiao, I. T., Kuan, W. C., Wai, Y. Y., Ko, H. J., Wan, Y. L., Hsu, Y. Y. & Liu, H. L. (2010) Study-specific EPI template improves group analysis in functional MRI of young and older adults. *Journal of Neuroscience Methods*. 189 (2), 257–266.
- Hughes, H. C. & Lang, C. M. (1978) Basic principles in selecting animal species for Research projects. *Clinical Toxicology*. 13 (5), 611–621.
- Hüttenrauch, M., Salinas, G. & Wirths, O. (2016) Effects of long-term environmental enrichment on anxiety, memory, hippocampal plasticity and overall brain gene expression in C57BL6 mice. *Frontiers in Molecular Neuroscience*. 9 (AUG), 62.
- Hyder, F., Behar, K. L., Martin, M. A., Blamire, A. M. & Shulman, R. G. (1994) Dynamic Magnetic Resonance Imaging of the Rat Brain during Forepaw Stimulation. *Journal of Cerebral Blood Flow & Metabolism*. 14 (4), 649–655.
- Jack, C. R., Lowe, V. J., Senjem, M. L., Weigand, S. D., Kemp, B. J., Shiung, M. M., Knopman, D. S., Boeve, B. F., Klunk, W. E., Mathis, C. A. & Petersen, R. C. (2008) ¹¹C PiB and structural MRI provide complementary information in imaging of Alzheimer’s disease and amnesic mild cognitive impairment. *Brain*. 131 (3), 665–680.

- Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Boeve, B. F., Tangalos, E. G. & Kokmen, E. (2000) Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology*. 55 (4), 484–489.
- Jeune, B. (2002) Living longer — but better? *Aging Clinical and Experimental Research*. 14 (2), 72–93.
- Johnson, S. A., Turner, S. M., Santacrose, L. A., Carty, K. N., Shafiq, L., Bizon, J. L., Maurer, A. P. & Burke, S. N. (2017) Rodent age-related impairments in discriminating perceptually similar objects parallel those observed in humans. *Hippocampus*. 27 (7), 759–776.
- Jonckers, E., Van Audekerke, J., De Visscher, G., Van der Linden, A. & Verhoye, M. (2011) Functional Connectivity fMRI of the Rodent Brain: Comparison of Functional Connectivity Networks in Rat and Mouse. *PLoS ONE*. 6 (4), e18876.
- Jones, D. T., Machulda, M. M., Vemuri, P., McDade, E. M., Zeng, G., Senjem, M. L., Gunter, J. L., Przybelski, S. A., Avula, R. T., Knopman, D. S., Boeve, B. F., Petersen, R. C. & Jack, C. R. (2011) Age-related changes in the default mode network are more advanced in Alzheimer disease. *Neurology*. 77 (16), 1524–1531.
- Jordan, M. I. & Mitchell, T. M. (2015) Machine learning: Trends, perspectives, and prospects. *Science*. 349 (6245), 255–260.
- Kang, H., Choi, D. H., Kim, S. K., Lee, J. & Kim, Y. J. (2016) Alteration of Energy Metabolism and Antioxidative Processing in the Hippocampus of Rats Reared in Long-Term Environmental Enrichment. *Developmental Neuroscience*. 38 (3), 186–194.
- Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., Renbing, X. & Peck, A. (1988) Clinical, pathological, and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. *Annals of Neurology*. 23 (2), 138–144.

- Kempermann, G. (2019) Environmental enrichment, new neurons and the neurobiology of individuality. *Nature Reviews Neuroscience*. 20 (4), 235–245.
- Kempermann, G., Gast, D. & Gage, F. H. (2002a) Neuroplasticity in old age: Sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. *Annals of Neurology*. 52 (2), 135–143.
- Kempermann, G., Gast, D. & Gage, F. H. (2002b) Neuroplasticity in old age: Sustained fivefold induction of hippocampal neurogenesis by long-term environmental enrichment. *Annals of Neurology*. 52 (2), 135–143.
- Kennedy, B. K., Berger, S. L., Brunet, A., Campisi, J., Cuervo, A. M., Epel, E. S., Franceschi, C., Lithgow, G. J., Morimoto, R. I., Pessin, J. E., Rando, T. A., Richardson, A., Schadt, E. E., Wyss-Coray, T. & Sierra, F. (2014) Geroscience: Linking Aging to Chronic Disease. *Cell*. 159 (4), 709–713.
- Kennedy, K. M. & Raz, N. (2015) *Normal Aging of the Brain*. Vol. 3. Elsevier Inc.
- Kiecolt-Glaser, J. K., Preacher, K. J., MacCallum, R. C., Atkinson, C., Malarkey, W. B. & Glaser, R. (2003) Chronic stress and age-related increases in the proinflammatory cytokine IL-6. *Proceedings of the National Academy of Sciences of the United States of America*. 100 (15), 9090–9095.
- Klose, U., Batra, M. & Nägele, T. (2013) Age-dependent changes in the histogram of apparent diffusion coefficients values in magnetic resonance imaging. *Frontiers in Aging Neuroscience*. 5 (NOV), 1–6.
- Kundu, P., Santin, M. D., Bandettini, P. A., Bullmore, E. T. & Petiet, A. (2014) Differentiating BOLD and non-BOLD signals in fMRI time series from anesthetized rats using multi-echo EPI at 11.7T. *NeuroImage*. 102 (P2), 861–874.
- Kundu, P., Voon, V., Balchandani, P., Lombardo, M. V., Poser, B. A. & Bandettini, P. A. (2017) Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. *NeuroImage*. 15459–80.
- Latora, V. & Marchiori, M. (2001) Efficient Behavior of Small-World Networks.

- Physical Review Letters*. 87 (19), 198701.
- Lauterbur, P. C. (1973) Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature*. 242 (5394), 190–191.
- Leal-Galicia, P., Castañeda-Bueno, M., Quiroz-Baez, R. & Arias, C. (2008) Long-term exposure to environmental enrichment since youth prevents recognition memory decline and increases synaptic plasticity markers in aging. *Neurobiology of Learning and Memory*. 90 (3), 511–518.
- Lee, H., Nakamura, K., Narayanan, S., Brown, R. A. & Arnold, D. L. (2019) Estimating and accounting for the effect of MRI scanner changes on longitudinal whole-brain volume change measurements. *NeuroImage*. 184555–565.
- Li, X. (2020) *dicm2nii*. [online]. Available from: <https://github.com/xiangruili/dicm2nii> (Accessed 20 March 2020).
- Lim, A., Krajina, K. & Marsland, A. L. (2013) ‘Peripheral Inflammation and Cognitive Aging’, in *Inflammation in Psychiatry*. S. Karger AG. pp. 175–187.
- Lin, A. L., Zhang, W., Gao, X. & Watts, L. (2015) Caloric restriction increases ketone bodies metabolism and preserves blood flow in aging brain. *Neurobiology of Aging*. 36 (7), 2296–2303.
- Liu, T. T. (2016) Noise contributions to the fMRI signal: An overview. *NeuroImage*. 143141–151.
- Liu, Y., Unsal, H. S., Tao, Y. & Zhang, N. (2020) Automatic Brain Extraction for Rodent MRI Images. *Neuroinformatics*. 18 (3), 395–406.
- Livingston, G., Sommerlad, A., Orgeta, V., et al. (2017) Dementia prevention, intervention, and care. *The Lancet*. 390 (10113), 2673–2734.
- Logan, J. M., Sanders, A. L., Snyder, A. Z., Morris, J. C. & Buckner, R. L. (2002) Under-Recruitment and Nonselective Recruitment. *Neuron*. 33 (5), 827–840.
- Lohmeier, J., Kaneko, T., Hamm, B., Makowski, M. R. & Okano, H. (2019)

- atlasBRET: Automated template-derived brain extraction in animal MRI. *Scientific Reports*. 9 (1), 1–9.
- Lu, H., Zou, Q., Gu, H., Raichle, M. E., Stein, E. A. & Yang, Y. (2012) Rat brains also have a default mode network. *Proceedings of the National Academy of Sciences*. 109 (10), 3979–3984.
- MacKay, A., Laule, C., Vavasour, I., Bjarnason, T., Kolind, S. & Mädler, B. (2006) Insights into brain microstructure from the T2 distribution. *Magnetic Resonance Imaging*. 24 (4), 515–525.
- Maggio, M., Guralnik, J. M., Longo, D. L. & Ferrucci, L. (2006) Interleukin-6 in Aging and Chronic Disease: A Magnificent Pathway. *The Journals of Gerontology: Series A*. 61 (6), 575–584.
- Maier, F. C., Wehrl, H. F., Schmid, A. M., Mannheim, J. G., Wiehr, S., Lerdkrai, C., Calaminus, C., Stahlschmidt, A., Ye, L., Burnet, M., Stiller, D., Sabri, O., Reischl, G., Staufenbiel, M., Garaschuk, O., Jucker, M. & Pichler, B. J. (2014) Longitudinal PET-MRI reveals β 2-amyloid deposition and rCBF dynamics and connects vascular amyloidosis to quantitative loss of perfusion. *Nature Medicine*. 20 (12), 1485–1492.
- Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. (2006) Statistical practice in high-throughput screening data analysis. *Nature Biotechnology*. 24 (2), 167–175.
- Manjón, J. V., Coupé, P., Martí-Bonmatí, L., Collins, D. L. & Robles, M. (2010) Adaptive non-local means denoising of MR images with spatially varying noise levels. *Journal of Magnetic Resonance Imaging*. 31 (1), 192–203.
- Martin, B., Mattson, M. P. & Maudsley, S. (2006) Caloric restriction and intermittent fasting: Two potential diets for successful brain aging. *Ageing Research Reviews* 5 (3) p.332–353.
- Maslov, S. & Sneppen, K. (2002) Specificity and stability in topology of protein

- networks. *Science*. 296 (5569), 910–913.
- Mata, R., Josef, A. K., Samanez-Larkin, G. R. & Hertwig, R. (2011) Age differences in risky choice: A meta-analysis. *Annals of the New York Academy of Sciences*. 1235 (1), 18–29.
- Matta, B. F., Heath, K. J., Tipping, K. & Summors, A. C. (1999) Direct Cerebral Vasodilatory Effects of Sevoflurane and Isoflurane. *Anesthesiology*. 91 (3), 677–677.
- McCarrey, A. C., An, Y., Kitner-Triolo, M. H., Ferrucci, L. & Resnick, S. M. (2016) Sex differences in cognitive trajectories in clinically normal older adults. *Psychology and Aging*. 31 (2), 166–175.
- McCutcheon, J. E. & Marinelli, M. (2009) Age matters. *European Journal of Neuroscience*. 29 (5), 997–1014.
- Melani, R., Chelini, G., Cenni, M. C. & Berardi, N. (2017) Enriched environment effects on remote object recognition memory. *Neuroscience*. 352296–305.
- Meunier, D., Achard, S., Morcom, A. & Bullmore, E. (2009) Age-related changes in modular organization of human brain functional networks. *NeuroImage*. 44 (3), 715–723.
- Michielse, S., Coupland, N., Camicioli, R., Carter, R., Seres, P., Sabino, J. & Malykhin, N. (2010) Selective effects of aging on brain white matter microstructure: A diffusion tensor imaging tractography study. *NeuroImage*. 52 (4), 1190–1201.
- Mikels, J. A., Shuster, M. M. & Thai, S. T. (2015) Aging and Decision Making. *Aging and Decision Making*. 169–188.
- Miller, M. G., Thangthaeng, N. & Shukitt-Hale, B. (2017) A Clinically Relevant Frailty Index for Aging Rats. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*. 72 (7), 892–896.
- Modlinska, K. & Pisula, W. (2020) The natural history of model organisms the norway rat, from an obnoxious pest to a laboratory pet. *eLife*. 91–13.

- Montembeault, M., Joubert, S., Doyon, J., Carrier, J., Gagnon, J. F., Monchi, O., Lungu, O., Belleville, S. & Brambati, S. M. (2012) The impact of aging on gray matter structural covariance networks. *NeuroImage*. 63 (2), 754–759.
- Moore, T. L., Killiany, R. J., Herndon, J. G., Rosene, D. L. & Moss, M. B. (2006) Executive system dysfunction occurs as early as middle-age in the rhesus monkey. *Neurobiology of Aging*. 27 (10), 1484–1493.
- Morrison, J. H. & Baxter, M. G. (2012) The ageing cortical synapse: hallmarks and implications for cognitive decline. *Nature reviews. Neuroscience*. 13 (4), 240–250.
- Mosconi, L. (2013) Glucose metabolism in normal aging and Alzheimer’s disease: methodological and physiological considerations for PET studies. *Clinical and Translational Imaging*. 1 (4), 217–233.
- Moseley, M. (2002) Diffusion tensor imaging and aging - A review. *NMR in Biomedicine*. 15 (7–8), 553–560.
- Muir, J. L., Everitt, B. J. & Robbins, T. W. (1996) The Cerebral Cortex of the Rat and Visual Attentional Function: Dissociable Effects of Medial Frontal, Cingulate, Anterior Dorsolateral, and Parietal Cortex Lesions on a Five-Choice Serial Reaction Time Task. *Cerebral Cortex*. 6 (3), 470–481.
- Muldoon, S. F., Bridgeford, E. W. & Bassett, D. S. (2016) Small-world propensity and weighted brain networks. *Scientific Reports*. 6.
- Mullarky, E. & Cantley, L. C. (2015) ‘Diverting Glycolysis to Combat Oxidative Stress’, in *Innovative Medicine*. Springer Japan. pp. 3–23.
- Murphy, K., Birn, R. M. & Bandettini, P. A. (2013) Resting-state fMRI confounds and cleanup. *NeuroImage*. 80 349–359.
- Naritomi, H., Meyer, J. S., Sakai, F., Yamaguchi, F. & Shaw, T. (1979) Effects of Advancing Age on Regional Cerebral Blood Flow: Studies in Normal Subjects and Subjects with Risk Factors for Atherothrombotic Stroke. *Archives of Neurology*. 36 (7), 410–416.

- Narvacan, K., Treit, S., Camicioli, R., Martin, W. & Beaulieu, C. (2017) Evolution of deep gray matter volume across the human lifespan. *Human Brain Mapping*. 38 (8), 3771–3790.
- Neal, Z. P. (2017) How small is it? Comparing indices of small worldliness. *Network Science*. 5 (1), 30–44.
- Newman, M. E. J. (2006) Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*. 103 (23), 8577–8582.
- Ng, K. K., Lo, J. C., Lim, J. K. W., Chee, M. W. L. & Zhou, J. (2016) Reduced functional segregation between the default mode network and the executive control network in healthy older adults: A longitudinal study. *NeuroImage*. 133321–330.
- Nilsson, M., Perfilieva, E., Johansson, U., Orwar, O. & Eriksson, P. S. (1999) Enriched environment increases neurogenesis in the adult rat dentate gyrus and improves spatial memory. *Journal of Neurobiology*. 39 (4), 569–578.
- Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U. & Bäckman, L. (2012) Memory aging and brain maintenance. *Trends in Cognitive Sciences*. 16 (5), 292–305.
- Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., Lind, J., Pudas, S., Persson, J. & Nilsson, L. G. (2010) Longitudinal evidence for diminished frontal cortex function in aging. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (52), 22682–22686.
- Nyuyki, K. D., Beiderbeck, D. I., Lukas, M., Neumann, I. D. & Reber, S. O. (2012) Chronic subordinate colony housing (CSC) as a model of chronic psychosocial stress in male rats. Yvette Tache (ed.). *PloS one*. 7 (12), e52371.
- O'Brien, J. L., O'Keefe, K. M., LaViolette, P. S., DeLuca, A. N., Blacker, D., Dickerson, B. C. & Sperling, R. A. (2010) Longitudinal fMRI in elderly reveals loss of hippocampal activation with clinical decline. *Neurology*. 74 (24), 1969–1976.

- O'Connor, E. E. & Zeffiro, T. A. (2019) Why is Clinical fMRI in a Resting State? *Frontiers in Neurology*. 10 (APR), 420.
- O'Sullivan, M., Jones, D. K., Summers, P. E., Morris, R. G., Williams, S. C. & Markus, H. S. (2001) Evidence for cortical 'disconnection' as a mechanism of age-related cognitive decline. *Neurology*. 57 (4), 632–638.
- Ogawa, S., Lee, T. M., Kay, A. R. & Tank, D. W. (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*. 87 (24), 9868–9872.
- Oguz, I., Zhang, H., Rumple, A. & Sonka, M. (2014) RATS: Rapid Automatic Tissue Segmentation in rodent brain MRI. *Journal of Neuroscience Methods*. 221175–182.
- Ohata, M., Sundaram, U., Fredericks, W. R., London, E. D. & Rapoport, S. I. (1981) Regional cerebral blood flow during development and ageing of the rat brain. *Brain*. 104 (2), 319–332.
- Onoda, K. & Yamaguchi, S. (2013) Small-worldness and modularity of the resting-state functional brain network decrease with aging. *Neuroscience Letters*. 556104–108.
- Orfanos, S., Toygar, T., Berthold-Losleben, M., Chechko, N., Durst, A., Laoutidis, Z. G., Vocke, S., Weidenfeld, C., Schneider, F., Karges, W., Beckmann, C. F., Habel, U. & Kohn, N. (2018) Investigating the impact of overnight fasting on intrinsic functional connectivity: a double-blind fMRI study. *Brain Imaging and Behavior*. 12 (4), 1150–1159.
- Paasonen, J., Stenroos, P., Salo, R. A., Kiviniemi, V. & Gröhn, O. (2018) Functional connectivity under six anesthesia protocols and the awake condition in rat brain. *NeuroImage*. 1729–20.
- Pagani, E., Agosta, F., Rocca, M. A., Caputo, D. & Filippi, M. (2008) Voxel-based analysis derived from fractional anisotropy images of white matter volume changes with aging. *NeuroImage*. 41 (3), 657–667.

- Pan, W.-J., Billings, J. C. W., Grooms, J. K., Shakil, S. & Keilholz, S. D. (2015) Considerations for resting state functional MRI and functional connectivity studies in rodents. *Frontiers in Neuroscience*. 9 (JUL), 269.
- Papp, E. A., Leergaard, T. B., Calabrese, E., Johnson, G. A. & Bjaalie, J. G. (2014) Waxholm Space atlas of the Sprague Dawley rat brain. *NeuroImage*. 97374–386.
- Parikh, I., Guo, J., Chuang, K. H., Zhong, Y., Rempe, R. G., Hoffman, J. D., Armstrong, R., Bauer, B., Hartz, A. M. S. & Lin, A. L. (2016) Caloric restriction preserves memory and reduces anxiety of aging mice with early enhancement of neurovascular functions. *Aging*. 8 (11), 2814–2826.
- Park, D. C. & Reuter-Lorenz, P. (2009) The adaptive brain: aging and neurocognitive scaffolding. *Annual review of psychology*. 60173–196.
- Pei, M., Nguyen, T. D., Thimmappa, N. D., Salustri, C., Dong, F., Cooper, M. A., Li, J., Prince, M. R. & Wang, Y. (2015) Algorithm for fast monoexponential fitting based on Auto-Regression on Linear Operations (ARLO) of data. *Magnetic Resonance in Medicine*. 73 (2), 843–850.
- Persson, J., Pudas, S., Nilsson, L. G. & Nyberg, L. (2014) Longitudinal assessment of default-mode brain function in aging. *Neurobiology of Aging*.
- Peters, A., Sethares, C. & Luebke, J. I. (2008) Synapses are lost during aging in the primate prefrontal cortex. *Neuroscience*. 152 (4), 970–981.
- Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G. & Kokmen, E. (1999) Mild Cognitive Impairment. *Archives of Neurology*. 56 (3), 303.
- Pfefferbaum, A., Rohlfing, T., Rosenbloom, M. J., Chu, W., Colrain, I. M. & Sullivan, E. V. (2013) Variation in longitudinal trajectories of regional brain volumes of healthy men and women (ages 10 to 85years) measured with atlas-based parcellation of MRI. *NeuroImage*. 65176–193.
- Pfefferbaum, A. & Sullivan, E. V. (2015) Cross-sectional versus longitudinal estimates of age-related changes in the adult brain: Overlaps and discrepancies.

- Neurobiology of Aging*. 36 (9), 2563–2567.
- Pike, C. J. (2017) Sex and the development of Alzheimer's disease. *Journal of Neuroscience Research*. 95 (1–2), 671–680.
- Poldrack, R. A. (2017) Precision Neuroscience: Dense Sampling of Individual Brains. *Neuron*. 95 (4), 727–729.
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J. B., Vul, E. & Yarkoni, T. (2017) Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*. 18 (2), 115–126.
- Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M. L., Elghahwagi, B., Richards, T., Dager, S. R. & Kiselev, V. G. (1999) Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magnetic Resonance in Medicine*. 42 (1), 87–97.
- Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. & Petersen, S. E. (2013) Evidence for hubs in human functional brain networks. *Neuron*. 79 (4), 798–813.
- Prins, N. D. & Scheltens, P. (2015) White matter hyperintensities, cognitive impairment and dementia: an update. *Nature Reviews Neurology*. 11 (3), 157–165.
- Puzianowska-Kuźnicka, M., Owczarz, M., Wieczorowska-Tobis, K., Nadrowski, P., Chudek, J., Slusarczyk, P., Skalska, A., Jonas, M., Franek, E. & Mossakowska, M. (2016) Interleukin-6 and C-reactive protein, successful aging, and mortality: The PolSenior study. *Immunity and Ageing*. 13 (1), .
- Qiu, L. R., Fernandes, D. J., Szulc-Lerch, K. U., Dazai, J., Nieman, B. J., Turnbull, D. H., Foster, J. A., Palmert, M. R. & Lerch, J. P. (2018) Mouse MRI shows brain areas relatively larger in males emerge before those larger in females. *Nature Communications*. 9 (1), .
- Quinn, R. (2005) Comparing rat's to human's age: How old is my rat in people years? *Nutrition*. 21 (6), 775–777.

- Raichle, M. E. (1998) Behind the scenes of functional brain imaging: A historical and physiological perspective. *Proceedings of the National Academy of Sciences*. 95 (3), 765–772.
- Raichle, M. E. (2015) The Brain's Default Mode Network. *Annual Review of Neuroscience*. 38 433–447.
- Raichle, M. E. & Gusnard, D. A. (2002) Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences of the United States of America* [online]. Available from: www.pnas.org/cgi/doi/10.1073/pnas.172399499 (Accessed 17 December 2020). 99 (16) p.10237–10239.
- Ramos, B. P., Birnbaum, S. G., Lindenmayer, I., Newton, S. S., Duman, R. S. & Arnsten, A. F. T. (2003) Dysregulation of protein kinase A signaling in the aged prefrontal cortex: New strategy for treating age-related cognitive decline. *Neuron*. 40 (4), 835–845.
- Raz, N. (1997) Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*. 7 (3), 268–282.
- Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A. & Acker, J. D. (2004a) Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: Replicability of regional differences in volume. *Neurobiology of Aging*. 25 (3), 377–396.
- Raz, N. & Lindenberger, U. (2011) Only time will tell: Cross-sectional studies offer no solution to the age–brain–cognition triangle: Comment on Salthouse (2011). *Psychological Bulletin*. 137 (5), 790–795.
- Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D. & Acker, J. D. (2005) Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*. 15 (11), 1676–1689.

- Raz, N. & Rodrigue, K. M. (2006) Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews*. 30 (6), 730–748.
- Raz, N., Rodrigue, K. M. & Haacke, E. M. (2007) Brain aging and its modifiers: Insights from in vivo neuromorphometry and susceptibility weighted imaging. *Annals of the New York Academy of Sciences*. 109784–93.
- Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M. & Acker, J. D. (2004b) Differential aging of the medial temporal lobe: A study of a five-year change. *Neurology*. 62 (3), 433–438.
- Reijmer, Y. D., Schultz, A. P., Leemans, A., O’Sullivan, M. J., Gurol, M. E., Sperling, R., Greenberg, S. M., Viswanathan, A. & Hedden, T. (2015) Decoupling of structural and functional brain connectivity in older adults with white matter hyperintensities. *NeuroImage*. 117222–229.
- Restom, K., Bangen, K. J., Bondi, M. W., Perthen, J. E. & Liu, T. T. (2007) Cerebral blood flow and BOLD responses to a memory encoding task: A comparison between healthy young and elderly adults. *NeuroImage*. 37 (2), 430–439.
- Reuter-Lorenz, P. A. & Park, D. C. (2014) How Does it STAC Up? Revisiting the Scaffolding Theory of Aging and Cognition. *Neuropsychology Review*. 24 (3), 355–370.
- Reuter-Lorenz, P. & Mikels, J. A. (2006) ‘The aging mind and brain: Implications of enduring plasticity for behavioral and cultural change’, in *Lifespan development and the brain: The perspective of biocultural coconstructivism*. pp. 255–276.
- Rissman, J., Gazzaley, A. & D’Esposito, M. (2004) Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*. 23 (2), 752–763.
- Robinson, S. D., Dymerska, B., Bogner, W., Barth, M., Zaric, O., Goluch, S., Unther Grabner, G., Deligianni, X., Bieri, O. & Trattnig, S. (2017) Combining Phase Images from Array Coils Using a Short Echo Time Reference Scan (COMPOSER). *Magn Reson Med*. 77318–327.

- Rodgers, Z. B., Detre, J. A. & Wehrli, F. W. (2016) MRI-based methods for quantification of the cerebral metabolic rate of oxygen. *Journal of Cerebral Blood Flow & Metabolism*. 36 (7), 1–21.
- Rodrigue, K. M., Kennedy, K. M. & Park, D. C. (2009) Beta-amyloid deposition and the aging brain. *Neuropsychology Review* 19 (4) p.436–450.
- Rogalski, E. J., Gefen, T., Shi, J., Samimi, M., Bigio, E., Weintraub, S., Geula, C. & Mesulam, M.-M. (2013) Youthful Memory Capacity in Old Brains: Anatomic and Genetic Clues from the Northwestern SuperAging Project. *Journal of Cognitive Neuroscience*. 25 (1), 29–36.
- Rosenzweig, M. R. & Bennett, E. L. (1996) Psychobiology of plasticity: Effects of training and experience on brain and behavior. *Behavioural Brain Research*. 78 (1), 57–65.
- Rowe, J. W. & Kahn, R. L. (1987) Human aging: Usual and successful. *Science*. 237 (4811), 143–149.
- Roy, C. & Sherrington, C. (1890) On the Regulation of the Blood-Supply of the Brain. *The Journal of physiology*. 11 (1–2), 85–158.
- Royston, P., Altman, D. G. & Sauerbrei, W. (2006) Dichotomizing continuous predictors in multiple regression: A bad idea. *Statistics in Medicine*. 25 (1), 127–141.
- Rubinov, M. & Sporns, O. (2010) Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*. 52 (3), 1059–1069.
- Rumple, A., McMurray, M., Johns, J., Lauder, J. & Makam, P. (2013) 3-Dimensional Diffusion Tensor Imaging (DTI) Atlas of the Rat Brain. *PLoS ONE*. 8 (7), 67334.
- Sala-Llonch, R., Bartrés-Faz, D. & Junqué, C. (2015) Reorganization of brain networks in aging: a review of functional connectivity studies. *Frontiers in Psychology*. 6663.

- Sala-Llonch, R., Junqué, C., Arenaza-Urquijo, E. M., Vidal-Piñeiro, D., Valls-Pedret, C., Palacios, E. M., Domènech, S., Salvà, A., Bargalló, N. & Bartrés-Faz, D. (2014) Changes in whole-brain functional networks and memory performance in aging. *Neurobiology of Aging*. 35 (10), 2193–2202.
- Salat, D. H. (2004) Thinning of the Cerebral Cortex in Aging. *Cerebral Cortex*. 14 (7), 721–730.
- Salat, D. H., Kaye, J. A. & Janowsky, J. S. (1999) Prefrontal gray and white matter volumes in healthy aging and Alzheimer disease. *Archives of Neurology*. 56 (3), 338–344.
- Salthouse, T. A. (1986) ‘Functional age: Examination of a concept.’, in *Age, health, and employment*. pp. 78–92.
- Salthouse, T. A. (2011) Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin*. 137 (5), 753–784.
- Salthouse, T. A. (2010) Selective review of cognitive aging. *Journal of the International Neuropsychological Society*. 16 (5), 754–760.
- Salthouse, T. A. (2019) Trajectories of normal cognitive aging. *Psychology and Aging*. 34 (1), 17–24.
- Salthouse, T. A. (2009) When does age-related cognitive decline begin? *Neurobiology of Aging* 30 (4) p.507–514.
- Salthouse, T. A. (2018) Why is cognitive change more negative with increased age? *Neuropsychology*. 32 (1), 110–120.
- Samanez-Larkin, G. R. & Knutson, B. (2015) Decision making in the ageing brain: Changes in affective and motivational circuits. *Nature Reviews Neuroscience*. 16 (5), 278–289.
- Sams-Dodd, F. (2006) Strategies to optimize the validity of disease models in the drug discovery process. *Drug Discovery Today* 11 (7–8) p.355–363.
- Schaie, K. W. & Hofer, S. M. (2001) ‘Longitudinal studies in aging research’, in

Handbook of the psychology of aging. pp. 53–77.

- Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K. & Fischl, B. (2004) A hybrid approach to the skull stripping problem in MRI. *NeuroImage*. 22 (3), 1060–1075.
- Seibenhener, M. L. & Wooten, M. C. (2015) Use of the Open Field Maze to Measure Locomotor and Anxiety-like Behavior in Mice. *Journal of Visualized Experiments*. (96), 52434.
- Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K. J., Vértes, P. E., Wagstyl, K., Kirkpatrick Reardon, P., Clasen, L., Liu, S., Messinger, A., Leopold, D. A., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., Raznahan, A. & Bullmore, E. T. (2018) Morphometric Similarity Networks Detect Microscale Cortical Organization and Predict Inter-Individual Cognitive Variation. *Neuron*. 97 (1), 231–247.e7.
- Sengupta, P. (2013) The Laboratory Rat: Relating Its Age With Human's. *International journal of preventive medicine*. 4 (6), 624–630.
- Simon, N. W., Lasarge, C. L., Montgomery, K. S., Williams, M. T., Ian, A., Setlow, B. & Bizon, J. (2010) Good things come to those who wait: attenuated discounting of delayed rewards in aged Fischer 344 rats. *Neurobiology of Aging*. 31 (5), 853–862.
- Simpson, J. & Kelly, J. P. (2011) The impact of environmental enrichment in laboratory rats—Behavioural and neurochemical aspects. *Behavioural Brain Research*. 222 (1), 246–264.
- Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., Ebmeier, K. P., Ferrie, J. E. & Dugravot, A. (2012) Timing of onset of cognitive decline: results from Whitehall II prospective cohort study. *BMJ*. 344 (jan04 4), d7622–d7622.
- Slob, A. K. & van der Werff Ten Bosch, J. J. (1975) Sex differences in body growth

- in the rat. *Physiology and Behavior*. 14 (3), 353–361.
- Smail, M. A., Smith, B. L., Nawreen, N. & Herman, J. P. (2020) Differential impact of stress and environmental enrichment on corticolimbic circuits. *Pharmacology Biochemistry and Behavior* 197.
- Smith, S. M. (2002) Fast robust automated brain extraction. *Human Brain Mapping*. 17 (3), 143–155.
- Smith, S. & Nichols, T. (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. 44 (1), 83–98.
- Snowdon, C. T. (1990) Language capacities of nonhuman animals. *American Journal of Physical Anthropology*. 33 (S11), 215–243.
- Snowdon, D. A. (2003) Healthy aging and dementia: findings from the Nun Study. *Annals of internal medicine*. 139 (5 Pt 2), 450–454.
- Speisman, R. B., Kumar, A., Rani, A., Pastoriza, J. M., Severance, J. E., Foster, T. C. & Ormerod, B. K. (2013) Environmental enrichment restores neurogenesis and rapid acquisition in aged rats. *Neurobiology of Aging*. 34 (1), 263–274.
- Spreng, R. N., Stevens, W. D., Viviano, J. D. & Schacter, D. L. (2016) Attenuated anticorrelation between the default and dorsal attention networks with aging: evidence from task and rest. *Neurobiology of Aging*. 45149–160.
- Stebbins, G. T., Carrillo, M. C., Dorfman, J., Dirksen, C., Desmond, J. E., Turner, D. A., Bennett, D. A., Wilson, R. S., Glover, G. & Gabrieli, J. D. E. (2002) Aging effects on memory encoding in the frontal lobes. *Psychology and Aging*. 17 (1), 44–55.
- Steffener, J., Brickman, A. M., Habeck, C. G., Salthouse, T. A. & Stern, Y. (2013) Cerebral blood flow and gray matter volume covariance patterns of cognition in aging. *Human Brain Mapping*. 34 (12), 3267–3279.
- Stern, Y. (2009) Cognitive reserve. *Neuropsychologia*. 47 (10), 2015–2028.

- Stern, Y. (2002) What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society : JINS*. 8 (3), 448–460.
- Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P. & Vuoksimaa, E. (2020) Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*. 16 (9), 1305–1311.
- Stone, A. J. & Blockley, N. P. (2017) A streamlined acquisition for mapping baseline brain oxygenation using quantitative BOLD. *NeuroImage*. 14779–88.
- Strøm, C., Rasmussen, L. S. & Sieber, F. E. (2014) Should general anaesthesia be avoided in the elderly? *Anaesthesia*. 69 Suppl 1 (SUPPL. 1), 35–44.
- Stüber, C., Morawski, M., Schäfer, A., Labadie, C., Wähnert, M., Leuze, C., Streicher, M., Barapatre, N., Reimann, K., Geyer, S., Spemann, D. & Turner, R. (2014) Myelin and iron concentration in the human brain: A quantitative study of MRI contrast. *NeuroImage*. 93 (P1), 95–106.
- Sullivan, E. V., Adalsteinsson, E., Hedehus, M., Ju, C., Moseley, M., Lim, K. O. & Pfefferbaum, A. (2001) Equivalent disruption of regional white matter microstructure in ageing healthy men and women. *NeuroReport*. 12 (1), 99–104.
- Sullivan, E. V. & Pfefferbaum, A. (2006) Diffusion tensor imaging and aging. *Neuroscience and Biobehavioral Reviews*. 30 (6), 749–761.
- Sullivan, G. M. & Feinn, R. (2012) Using Effect Size—or Why the P Value Is Not Enough . *Journal of Graduate Medical Education*. 4 (3), 279–282.
- Sun, F. W., Stepanovic, M. R., Andreano, J., Barrett, L. F., Touroutoglou, A. & Dickerson, B. C. (2016) Youthful brains in older adults: Preserved neuroanatomy in the default mode and salience networks contributes to youthful

- memory in superaging. *Journal of Neuroscience*. 36 (37), 9659–9668.
- Sun, J., Tong, S. & Yang, G. Y. (2012) Reorganization of brain networks in aging and age-related diseases. *Aging and Disease* [online]. Available from: /pmc/articles/PMC3377830/?report=abstract (Accessed 13 December 2020). 3 (2) p.181–193.
- Syddall, H., Cooper, C., Martin, F., Briggs, R. & Sayer, A. A. (2003) Is grip strength a useful single marker of frailty? *Age and Ageing*. 32 (6), 650–656.
- Takeshita, H., Yamamoto, K., Nozato, S., Inagaki, T., Tsuchimochi, H., Shirai, M., Yamamoto, R., Imaizumi, Y., Hongyo, K., Yokoyama, S., Takeda, M., Oguro, R., Takami, Y., Itoh, N., Takeya, Y., Sugimoto, K., Fukada, S. I. & Rakugi, H. (2017) Modified forelimb grip strength test detects aging-associated physiological decline in skeletal muscle function in male mice. *Scientific Reports*. 7 (1), 42323.
- Tarumi, T. & Zhang, R. (2018) Cerebral blood flow in normal aging adults: cardiovascular determinants, clinical implications, and aerobic fitness. *Journal of Neurochemistry*. 144 (5), 595–608.
- Tauffmanberger, A., Fiumelli, H., Almustafa, S. & Magistretti, P. J. (2019) Lactate and pyruvate promote oxidative stress resistance through hormetic ROS signaling. *Cell Death and Disease*. 10 (9), .
- Timmers, I., Roebroek, A., Bastiani, M., Jansma, B., Rubio-Gozalbo, E. & Zhang, H. (2016) Assessing Microstructural Substrates of White Matter Abnormalities: A Comparative Study Using DTI and NODDI Pew-Thian Yap (ed.). *PLOS ONE*. 11 (12), e0167884.
- Todorovic, S. T., Smiljanic, K. R., Ruzdijic, S. D., Djordjevic, A. N. M. & Kanazir, S. D. (2018) Effects of Different Dietary Protocols on General Activity and Frailty of Male Wistar Rats During Aging. *The Journals of Gerontology: Series A*. 73 (8), 1036–1044.

- Toescu, E. C. (2005) Normal brain ageing: models and mechanisms. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 360 (1464), 2347–2354.
- Tomasi, D. & Volkow, N. D. (2012) Aging and functional brain networks. *Mol Psychiatry*. 17 (5), 471–558.
- Traag, V. A., Waltman, L. & van Eck, N. J. (2019) From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*. 9 (1), 1–12.
- Tucker, A. M. & Stern, Y. (2011) Cognitive reserve in aging. *Current Alzheimer research*. 8 (4), 354–360.
- Turner, G. R. & Spreng, R. N. (2012) Executive functions and neurocognitive aging: Dissociable patterns of brain activity. *Neurobiology of Aging*. 33 (4), 826.e1-826.e13.
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A. & Gee, J. C. (2010) N4ITK: Improved N3 bias correction. *IEEE Transactions on Medical Imaging*. 29 (6), 1310–1320.
- Vaishnavi, S. N., Vlassenko, A. G., Rundle, M. M., Snyder, A. Z., Mintun, M. A. & Raichle, M. E. (2010) Regional aerobic glycolysis in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (41), 17757–17762.
- Valdés-Hernández, P. A., Sumiyoshi, A., Nonaka, H., Haga, R., Aubert-Vásquez, E., Ogawa, T., Iturria-Medina, Y., Riera, J. J. & Kawashima, R. (2011) An in vivo MRI template set for morphometry, tissue segmentation, and fMRI localization in rats. *Frontiers in Neuroinformatics*. 5 (November), 1–19.
- Vaupel, J. W. (2010) Biodemography of human ageing. *Nature* 464 (7288) p.536–542.
- Vincent, R. D., Neelin, P., Khalili-Mahani, N., Janke, A. L., Fonov, V. S., Robbins, S. M., Baghdadi, L., Lerch, J., Sled, J. G., Adalat, R., Macdonald, D., Zijdenbos, A. P., Collins, D. L. & Evans, A. C. (2016) MINC 2.0: A flexible format for

- multi-modal images. *Frontiers in Neuroinformatics*. 10 (AUG), .
- de Vries, N. M., Staal, J. B., van Ravensberg, C. D., Hobbelen, J. S. M., Olde Rikkert, M. G. M. & Nijhuis-van der Sanden, M. W. G. (2011) Outcome instruments to measure frailty: A systematic review. *Ageing Research Reviews*. 10 (1), 104–114.
- Vutskits, L. & Xie, Z. (2016) Lasting impact of general anaesthesia on the brain: mechanisms and relevance. *Nature Reviews Neuroscience*. 17 (11), 705–717.
- Wakana, S., Jiang, H., Nagae-Poetscher, L. M., van Zijl, P. C. M. & Mori, S. (2004) Fiber Tract-based Atlas of Human White Matter Anatomy. *Radiology*. 230 (1), 77–87.
- Walhovd, K. B., Fjell, A. M. & Espeseth, T. (2014) Cognitive decline and brain pathology in aging - need for a dimensional, lifespan and systems vulnerability view. *Scandinavian Journal of Psychology*. 55 (3), 244–254.
- Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., Quinn, B. T., Salat, D., Makris, N. & Fischl, B. (2005) Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging*. 26 (9), 1261–1270.
- Wang, J., Weygand, J., Hwang, K. P., Mohamed, A. S. R., Ding, Y., Fuller, C. D., Lai, S. Y., Frank, S. J. & Zhou, J. (2016) Magnetic Resonance Imaging of Glucose Uptake and Metabolism in Patients with Head and Neck Cancer. *Scientific Reports*. 6 (1), 1–7.
- Wansapura, J. P., Holland, S. K., Dunn, R. S. & Ball, W. S. (1999) NMR relaxation times in the human brain at 3.0 tesla. *Journal of Magnetic Resonance Imaging*. 9 (4), 531–538.
- Watson, C. G. (2018) *brainGraph: Graph Theory Analysis of Brain MRI Data*. [online]. Available from: <https://github.com/cwatson/brainGraph>.
- Watts, D. J. & Strogatz, S. H. (1998) Collective dynamics of ‘small-world’ networks. *Nature*. 393 (6684), 440–442.

- Wearn, A. R., Nurdal, V., Saunders-Jennings, E., Knight, M. J., Isotalus, H. K., Dillon, S., Tsivos, D., Kauppinen, R. A. & Coulthard, E. J. (2020) T2 heterogeneity: A novel marker of microstructural integrity associated with cognitive decline in people with mild cognitive impairment. *Alzheimer's Research and Therapy*. 12 (1), 105.
- Wei, Z., Chen, L., Hou, X., van Zijl, P. C. M., Xu, J. & Lu, H. (2020) Age-Related Alterations in Brain Perfusion, Venous Oxygenation, and Oxygen Metabolic Rate of Mice: A 17-Month Longitudinal MRI Study. *Frontiers in Neurology*. 11559.
- Weindruch, R. & Sohal, R. S. (1997) Seminars in medicine of the Beth Israel Deaconess Medical Center. Caloric intake and aging. Jeffrey S. Flier & Lisa H. Underhill (eds.). *The New England journal of medicine*. 337 (14), 986–994.
- Weissman, D. H., Roberts, K. C., Visscher, K. M. & Woldorff, M. G. (2006) The neural bases of momentary lapses in attention. *Nature Neuroscience*. 9 (7), 971–978.
- Wen, W., Sachdev, P. S., Li, J. J., Chen, X. & Anstey, K. J. (2009) White matter hyperintensities in the forties: Their prevalence and topography in an epidemiological sample aged 44-48. *Human Brain Mapping*. 30 (4), 1155–1167.
- Westphal, R., Simmons, C., Mesquita, M. B., Wood, T. C., Williams, S. C. R., Vernon, A. C. & Cash, D. (2017) Characterization of the resting-state brain network topology in the 6-hydroxydopamine rat model of Parkinson's disease Satoru Hayasaka (ed.). *PLOS ONE*. 12 (3), e0172394.
- Whitney, W. O. & Mehlhaff, C. J. (1987) High-rise syndrome in cats. *Journal of the American Veterinary Medical Association*.
- van Wijk, B. C. M., Stam, C. J. & Daffertshofer, A. (2010) Comparing Brain Networks of Different Size and Connectivity Density Using Graph Theory Olaf Sporns (ed.). *PLoS ONE*. 5 (10), e13701.
- Wollny, G., Kellman, P., Ledesma-Carbayo, M.-J., Skinner, M. M., Hublin, J.-J. &

- Hierl, T. (2013) MIA - A free and open source software for gray scale medical image analysis. *Source Code for Biology and Medicine*. 8 (1), 20.
- Wood, T. C. (2018) QUIT: QUantitative Imaging Tools. *Journal of Open Source Software*. 3 (26), 656.
- Wood, T. C., Lythgoe, D. J. & Williams, S. C. R. (2013) 'rBET: Making BET work for Rodent Brains', in *Proc Intl Soc Mag Reson Med 21 (2013)*. 2013 p. 2707.
- Wu, L., Zhao, H., Weng, H. & Ma, D. (2019) Lasting effects of general anesthetics on the brain in the young and elderly: “mixed picture” of neurotoxicity, neuroprotection and cognitive impairment. *Journal of Anesthesia*. 33 (2), 321–335.
- Wurtman, J. J. & Miller, S. A. (1976) Effect of litter size on weight gain in rats. *Journal of Nutrition*. 106 (5), 697–701.
- Yassa, M. A., Lacy, J. W., Stark, S. M., Albert, M. S., Gallagher, M. & Stark, C. E. L. (2011) Pattern separation deficits associated with increased hippocampal CA3 and dentate gyrus activity in nondemented older adults. *Hippocampus*. 21 (9), 968–979.
- Yeatman, J. D., Wandell, B. A. & Mezer, A. A. (2014) Lifespan maturation and degeneration of human brain white matter. *Nature Communications*. 5 (1), 4932.
- Ylikoski, R., Ylikoski, A., Keskivaara, P., Tilvis, R., Sulkava, R. & Erkinjuntti, T. (1999) Heterogeneity of cognitive profiles in aging: successful aging, normal aging, and individuals at risks for cognitive decline. *European Journal of Neurology*. 6 (6), 645–652.
- Yorke, A., Kane, A. E., Hancock Friesen, C. L., Howlett, S. E. & O'Blenes, S. (2017) Development of a Rat Clinical Frailty Index. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*. 72 (7), 897–903.
- Yu, Y., Herman, P., Rothman, D. L., Agarwal, D. & Hyder, F. (2018) Evaluating the gray and white matter energy budgets of human brain function. *Journal of Cerebral Blood Flow and Metabolism*. 38 (8), 1339–1353.

- Zahodne, L. B., Manly, J. J., MacKay-Brandt, A. & Stern, Y. (2013) Cognitive declines precede and predict functional declines in aging and Alzheimer's disease. *PloS One*. 8 (9), e73645.
- Zalesky, A., Fornito, A. & Bullmore, E. T. (2010) Network-based statistic: Identifying differences in brain networks. *NeuroImage*. 53 (4), 1197–1207.
- Zhang, H., Schneider, T., Wheeler-Kingshott, C. A. & Alexander, D. C. (2012) NODDI: Practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage*. 61 (4), 1000–1016.
- Zhang, Y., Kuang, Y., Xu, K., Harris, D., Lee, Z., Lamanna, J. & Puchowicz, M. A. (2013) Ketosis proportionately spares glucose utilization in brain. *Journal of Cerebral Blood Flow and Metabolism*. 33 (8), 1307–1311.
- Zhao, T., Cao, M., Niu, H., Zuo, X. N., Evans, A., He, Y., Dong, Q. & Shu, N. (2015) Age-related changes in the topological organization of the white matter structural connectome across the human lifespan. *Human Brain Mapping*. 36 (10), 3777–3792.
- Zimmermann, J., Ritter, P., Shen, K., Rothmeier, S., Schirner, M. & McIntosh, A. R. (2016) Structural architecture supports functional organization in the human aging brain at a regionwise and network level. *Human Brain Mapping*. 37 (7), 2645–2661.

APPENDICES

Appendix 2-1: Environmental enrichment sets

Set 1
Set 2
Set 3
Set 4

Appendix 2-1: Environmental enrichment sets that were rotated weekly throughout the course of the study after the 3-month time point.

Appendix 2-2: Frailty index score sheet

Rat ID:	Body Weight (g):			Rating: 0 = absent 0.5 = mild 1 = severe
<u>Integument</u>				Comments:
1. Alopecia	0	0.5	1	
2. Dermatitis	0	0.5	1	
3. Coat condition	0	0.5	1	
<u>Physical/Musculoskeletal</u>				
4. Tumours	0	0.5	1	
5. Distended abdomen	0	0.5	1	
6. Hunched posture	0	0.5	1	
7. Body condition score	0	0.5	1	
8. Gait disorder	0	0.5	1	
9. Tremor	0	0.5	1	
<u>Vestibulocochlear/Auditory</u>				
10. Hearing loss	0	0.5	1	
<u>Ocular/Nasal</u>				
11. Cataracts	0	0.5	1	
12. Chromodacryorrhea/porphyrin	0	0.5	1	
13. Exophthalmos	0	0.5	1	
14. Microphthalmos	0	0.5	1	
15. Corneal opacity	0	0.5	1	
<u>Neurological</u>				
16. Head tilt	0	0.5	1	
<u>Digestive/Urogenital</u>				
17. Malocclusion	0	0.5	1	
18. Diarrhea	0	0.5	1	
19. Jaundice	0	0.5	1	
20. Penile prolapse	0	0.5	1	
21. Rectal prolapse	0	0.5	1	
<u>Pain/Discomfort</u>				
23. Piloerection	0	0.5	1	
24. Unusual sounds	0	0.5	1	

TOTAL SCORE =

TOTAL SCORE/MAX SCORE =

<u>Symptom</u>	<u>Description</u>
Alopecia	Acquired hair loss due to inflammation, endocrine disorder, or idiopathic disease.
Skin lesions/dermatitis	Excessive scratching, self-mutilation, or skin conditions leading to open sores on the body.
Coat condition	Ungroomed appearance: fur appears ruffled and matted.
Tumours	Presence of neoplastic growths.
Distended abdomen	Asymmetric/enlarged abdomen. May be due to neoplastic growths, organ enlargement or peritoneal
fluid accumulation.	
Hunched posture	Presence of hunched posture (head down, feet together); reduced mobility.
Body condition	Visual signs of emaciation or obesity. Based on the amount of flesh covering the vertebral column
and dorsal pelvis.	
Gait disorders	Abnormal locomotion: slow movement, lack of coordination, stumbling, falling, or limping.
Tremor	Involuntary shaking at rest.
Hearing loss	Impaired acoustic startle reflex; associated with loss of hearing sensitivity.
Cataracts	Opaque spot in the center of the eye; clouding of the lens.
Chromodacryorrhea	Porphyrim staining around the eyes/nose.
Exophthalmos	Abnormal protrusion of the eye.
Microphthalmos	Abnormally small eye. Sunken in appearance.
Corneal opacity	Cornea appears white or clouded.
Head tilt	Abnormal/asymmetric head position associated with a central nervous system disturbance.
Malocclusion	Abnormal occlusion due to uneven or overgrown incisors.
Diarrhea	Increased frequency and decreased consistency of bowel movements. Fecal smearing in cage.
Jaundice	Yellowing of the feet, nose, ears and tail associated with accumulation of bilirubin.
Penile/vaginal prolapse	Tissue protruding from the penis or vagina.
Rectal prolapse	Tissue protruding from the rectum.
Piloerection	Fur standing on end
Unusual sounds	Acute vocalization in response to touch.

Appendix 3-1: Example morphometric similarity adjacency matrix

Appendix 3-1: An example morphometric similarity matrix before thresholding for a single subject at the 3-month time point.

Appendix 3-2: Supplemental figures for univariate voxelwise comparisons

Volumetric differences at baseline: A) Grey matter volume; B) grey matter concentration; C) absolute log Jacobian determinant, D) relative log-Jacobian determinant. There are small differences where the control group had a slightly higher mean GMV and absolute Jacobian determinant than the EEDR group but no observed differences between groups for the relative Jacobian determinant.

Relaxometry differences at baseline: E) T1 shows no differences between groups at the baseline time point; F) T2 shows no differences between groups; G) T₂ is longer in the frontal regions in EEDR rats and longer in ventrocaudal regions in controls but these differences do not pass the significance threshold*

Diffusion differences at baseline: H) V_{ic} is higher in the EEDR. I) V_{iso} differs by region in both directions J) ODI is higher in some regions in the EEDR group. There are no group differences in K) FA or L) MD at baseline.

Appendix 3-3: Supplemental figures for thresholding structural network graphs

Sparsity (%)	Time point							
	T1		T2		T3		T4	
	CE	SWP	CE	SWP	CE	SWP	CE	SWP
5	0.019	0.976	0.012	0.982	0.008	0.983	0.008	0.987
10	0.126	0.918	0.111	0.922	0.114	0.910	0.119	0.914
15	0.177	0.927	0.170	0.921	0.174	0.925	0.173	0.923
20	0.194	0.925	0.189	0.925	0.192	0.922	0.193	0.927
25	0.193	0.918	0.189	0.920	0.193	0.919	0.194	0.923
30	0.184	0.911	0.181	0.914	0.184	0.910	0.185	0.913
35	0.168	0.904	0.166	0.906	0.168	0.902	0.169	0.902
40	0.148	0.897	0.146	0.899	0.148	0.895	0.148	0.9893
45	0.123	0.889	0.123	0.891	0.123	0.886	0.124	0.885
50	0.097	0.881	0.096	0.884	0.097	0.878	0.097	0.877

Appendix 3-3: Global cost-efficiency (CE) and small-world propensity (SWP) calculated across various sparsity thresholds for each time point. Bold emphasis denotes largest value for time point.

Appendix 4-1: Mean AUC of functional connectivity metrics

	EEDR				Control			
Age (months)	3	5	11	17	3	5	11	17
E_{glob}	14.0604	12.6647	10.1856	9.0317	15.5378	12.2441	10.9575	11.3970
Overall connectivity	46.3616	38.9416	29.6006	25.0063	53.1886	37.8778	31.5856	30.3673
Q	31.8830	32.5103	33.5912	33.0045	30.8729	34.2699	33.0506	33.6688
SWP	32.1113	32.5725	33.0399	33.5940	31.5486	32.2740	33.9701	33.5113

Appendix 4-1: Area under the curve derived from the mean values of global efficiency (E_{glob}), overall functional connectivity, maximum modularity (Q), and small-world propensity (SWP), across subjects for a range of graph densities.

Appendix 4-2: Baseline group differences in overall functional connectivity and global efficiency

Appendix 4-2 Differences in the area under the curve for functional connectivity and global efficiency at baseline. A) curve for mean overall functional connectivity at each sparsity with error bars representing one standard deviation B) null distribution of overall functional connectivity with red line at the observed difference in AUC. C) curve for mean global efficiency at each sparsity with error bars representing one standard deviation B) null distribution of global efficiency with red line at the observed difference in AUC.

Appendix 4-3: Extended estimates for model predicting cerebral blood flow from age, region, and group

Model factor	β	Std. Error	df	P-value
Intercept: difference from 0				
control visual cortex at 3 months	94.021	4.928	54.98	0.0000
Main effects: difference from control visual cortex at 3 months				
5 months	-9.310	7.153	54.45	0.1985
11 months	20.129	7.383	50.59	0.0088
17 months	41.175	8.446	48.15	0.0000
parietal	3.640	2.083	25453.84	0.0806
retrosplenial	2.983	2.083	25453.84	0.1521
cingulate	7.889	2.083	25453.84	0.0002
motor	-2.826	2.083	25453.84	0.1748
sensory	-3.164	2.083	25453.84	0.1287
rhinal	3.053	2.083	25453.84	0.1427
frontal	-11.311	2.083	25453.84	0.0000
auditory	-7.720	2.083	25453.84	0.0002
piriform	-14.188	2.083	25453.84	0.0000
prefrontal	13.167	2.083	25453.84	0.0000
olfactory	18.704	2.083	25453.84	0.0000
insula	-9.587	2.083	25453.84	0.0000
basal forebrain	1.149	2.083	25453.84	0.5811
ventral hippocampus	7.495	2.083	25453.84	0.0003
dorsal hippocampus	-1.545	2.083	25453.84	0.4581
hypothalamus	12.208	2.083	25453.84	0.0000
thalamus	3.332	2.083	25453.84	0.1097
globus pallidus	-5.589	2.083	25453.84	0.0073
striatum	2.643	2.083	25453.84	0.2044
substantia nigra	15.760	2.083	25453.84	0.0000
colliculi	32.677	2.083	25453.84	0.0000
EEDR	4.146	6.936	53.96	0.5525
Interaction: ROI and time (difference from control visual cortex at 3 months)				
5 months: parietal	4.084	2.946	25453.84	0.1656
11 months: parietal	-3.248	3.194	25453.84	0.3093
17 months: parietal	-7.684	3.773	25453.84	0.0417
5 months: retrosplenial	8.032	2.946	25453.84	0.0064
11 months: retrosplenial	-3.045	3.194	25453.84	0.3406
17 months: retrosplenial	-5.124	3.773	25453.84	0.1745
5 months: cingulate	11.643	2.946	25453.84	0.0001
11 months: cingulate	2.640	3.194	25453.84	0.4085
17 months: cingulate	-5.231	3.773	25453.84	0.1656
5 months: motor	-0.052	2.946	25453.84	0.9860
11 months: motor	-7.458	3.194	25453.84	0.0196
17 months: motor	-10.824	3.773	25453.84	0.0041
5 months: sensory	3.751	2.946	25453.84	0.2029
11 months: sensory	-5.137	3.194	25453.84	0.1078
17 months: sensory	-12.186	3.773	25453.84	0.0012
5 months: rhinal	8.249	2.946	25453.84	0.0051
11 months: rhinal	9.296	3.194	25453.84	0.0036
17 months: rhinal	15.315	3.773	25453.84	0.0000
5 months: frontal	3.769	2.946	25453.84	0.2007
11 months: frontal	-7.877	3.194	25453.84	0.0137
17 months: frontal	-0.832	3.773	25453.84	0.8254
5 months: auditory	8.148	2.946	25453.84	0.0057
11 months: auditory	5.765	3.194	25453.84	0.0711
17 months: auditory	17.231	3.773	25453.84	0.0000
5 months: piriform	12.487	2.946	25453.84	0.0000
11 months: piriform	9.298	3.194	25453.84	0.0036
17 months: piriform	-10.855	3.773	25453.84	0.0040
5 months: prefrontal	8.937	2.946	25453.84	0.0024
11 months: prefrontal	-6.310	3.194	25453.84	0.0482
17 months: prefrontal	-14.632	3.773	25453.84	0.0001

5 months: olfactory	-2.921	2.946	25453.84	0.3215
11 months: olfactory	-24.451	3.194	25453.84	0.0000
17 months: olfactory	-35.652	3.773	25453.84	0.0000
5 months: insula	14.296	2.946	25453.84	0.0000
11 months: insula	5.539	3.194	25453.84	0.0829
17 months: insula	-1.906	3.773	25453.84	0.6135
5 months: basal forebrain	0.938	2.946	25453.84	0.7503
11 months: basal forebrain	-17.647	3.194	25453.84	0.0000
17 months: basal forebrain	-30.585	3.773	25453.84	0.0000
5 months: ventral hippocampus	4.565	2.946	25453.84	0.1212
11 months: ventral hippocampus	-5.033	3.194	25453.84	0.1152
17 months: ventral hippocampus	-2.188	3.773	25453.84	0.5620
5 months: dorsal hippocampus	1.497	2.946	25453.84	0.6113
11 months: dorsal hippocampus	-22.183	3.194	25453.84	0.0000
17 months: dorsal hippocampus	-23.362	3.773	25453.84	0.0000
5 months: hypothalamus	3.234	2.946	25453.84	0.2723
11 months: hypothalamus	-22.679	3.194	25453.84	0.0000
17 months: hypothalamus	-30.624	3.773	25453.84	0.0000
5 months: thalamus	-3.662	2.946	25453.84	0.2139
11 months: thalamus	-25.568	3.194	25453.84	0.0000
17 months: thalamus	-28.353	3.773	25453.84	0.0000
5 months: globus pallidus	-9.008	2.946	25453.84	0.0022
11 months: globus pallidus	-36.307	3.194	25453.84	0.0000
17 months: globus pallidus	-49.342	3.773	25453.84	0.0000
5 months: striatum	1.104	2.946	25453.84	0.7078
11 months: striatum	-18.689	3.194	25453.84	0.0000
17 months: striatum	-31.095	3.773	25453.84	0.0000
5 months: substantia nigra	-15.983	2.946	25453.84	0.0000
11 months: substantia nigra	-34.692	3.194	25453.84	0.0000
17 months: substantia nigra	-35.703	3.773	25453.84	0.0000
5 months: colliculi	-15.581	2.946	25453.84	0.0000
11 months: colliculi	-29.471	3.194	25453.84	0.0000
17 months: colliculi	-16.939	3.773	25453.84	0.0000
Interaction: group and time (difference from control visual cortex at 3 months)				
5 months: EEDR	3.551	10.071	53.49	0.7258
11 months: EEDR	-10.225	9.804	46.79	0.3023
17 months: EEDR	-19.401	10.758	43.56	0.0783
Interaction: ROI and group (difference from control visual cortex at 3 months)				
parietal: EEDR	-2.831	2.790	25453.84	0.3103
retrosplenial: EEDR	1.701	2.790	25453.84	0.5420
cingulate: EEDR	4.298	2.790	25453.84	0.1234
motor: EEDR	0.412	2.790	25453.84	0.8825
sensory: EEDR	1.435	2.790	25453.84	0.6071
rhinal: EEDR	0.543	2.790	25453.84	0.8458
frontal: EEDR	-7.436	2.790	25453.84	0.0077
auditory: EEDR	2.913	2.790	25453.84	0.2964
piriform: EEDR	-3.448	2.790	25453.84	0.2165
prefrontal: EEDR	0.572	2.790	25453.84	0.8375
olfactory: EEDR	-7.675	2.790	25453.84	0.0059
insula: EEDR	-1.298	2.790	25453.84	0.6416
basal forebrain: EEDR	-3.208	2.790	25453.84	0.2501
ventral hippocampus: EEDR	-1.269	2.790	25453.84	0.6493
dorsal hippocampus: EEDR	-1.273	2.790	25453.84	0.6482
hypothalamus: EEDR	6.379	2.790	25453.84	0.0222
thalamus: EEDR	3.698	2.790	25453.84	0.1849
globus pallidus: EEDR	-2.269	2.790	25453.84	0.4160
striatum: EEDR	-1.292	2.790	25453.84	0.6432
substantia nigra: EEDR	-5.258	2.790	25453.84	0.0594
colliculi: EEDR	6.714	2.790	25453.84	0.0161
Interaction: ROI, group, and time (difference from difference between control visual cortex at 3 months)				
5 months: parietal: EEDR	-2.294	3.945	25453.84	0.5609
11 months: parietal: EEDR	4.055	4.143	25453.84	0.3278
17 months: parietal: EEDR	-1.137	4.652	25453.84	0.8069
5 months: retrosplenial: EEDR	-5.226	3.945	25453.84	0.1853
11 months: retrosplenial: EEDR	-2.503	4.143	25453.84	0.5457

17 months: retrosplenial: EEDR	-14.431	4.652	25453.84	0.0019
5 months: cingulate: EEDR	-5.469	3.945	25453.84	0.1656
11 months: cingulate: EEDR	-2.069	4.143	25453.84	0.6176
17 months: cingulate: EEDR	-12.972	4.652	25453.84	0.0053
5 months: motor: EEDR	-1.757	3.945	25453.84	0.6560
11 months: motor: EEDR	4.359	4.143	25453.84	0.2928
17 months: motor: EEDR	-8.830	4.652	25453.84	0.0577
5 months: sensory: EEDR	-1.485	3.945	25453.84	0.7067
11 months: sensory: EEDR	1.723	4.143	25453.84	0.6775
17 months: sensory: EEDR	-2.532	4.652	25453.84	0.5862
5 months: rhinal: EEDR	-4.354	3.945	25453.84	0.2697
11 months: rhinal: EEDR	-8.070	4.143	25453.84	0.0515
17 months: rhinal: EEDR	-8.007	4.652	25453.84	0.0852
5 months: frontal: EEDR	4.676	3.945	25453.84	0.2359
11 months: frontal: EEDR	11.472	4.143	25453.84	0.0056
17 months: frontal: EEDR	-9.645	4.652	25453.84	0.0381
5 months: auditory: EEDR	-0.899	3.945	25453.84	0.8197
11 months: auditory: EEDR	5.361	4.143	25453.84	0.1957
17 months: auditory: EEDR	-9.349	4.652	25453.84	0.0445
5 months: piriform: EEDR	-1.956	3.945	25453.84	0.6201
11 months: piriform: EEDR	-7.641	4.143	25453.84	0.0652
17 months: piriform: EEDR	4.116	4.652	25453.84	0.3763
5 months: prefrontal: EEDR	-1.875	3.945	25453.84	0.6345
11 months: prefrontal: EEDR	3.872	4.143	25453.84	0.3500
17 months: prefrontal: EEDR	-7.064	4.652	25453.84	0.1289
5 months: olfactory: EEDR	5.768	3.945	25453.84	0.1438
11 months: olfactory: EEDR	13.416	4.143	25453.84	0.0012
17 months: olfactory: EEDR	7.587	4.652	25453.84	0.1029
5 months: insula: EEDR	-2.305	3.945	25453.84	0.5591
11 months: insula: EEDR	2.332	4.143	25453.84	0.5736
17 months: insula: EEDR	-2.917	4.652	25453.84	0.5307
5 months: basal forebrain: EEDR	2.353	3.945	25453.84	0.5508
11 months: basal forebrain: EEDR	6.124	4.143	25453.84	0.1394
17 months: basal forebrain: EEDR	6.952	4.652	25453.84	0.1351
5 months: ventral hippocampus: EEDR	4.420	3.945	25453.84	0.2626
11 months: ventral hippocampus: EEDR	4.682	4.143	25453.84	0.2585
17 months: ventral hippocampus: EEDR	-7.394	4.652	25453.84	0.1119
5 months: dorsal hippocampus: EEDR	-0.574	3.945	25453.84	0.8843
11 months: dorsal hippocampus: EEDR	2.413	4.143	25453.84	0.5603
17 months: dorsal hippocampus: EEDR	-8.636	4.652	25453.84	0.0634
5 months: hypothalamus: EEDR	-3.507	3.945	25453.84	0.3740
11 months: hypothalamus: EEDR	-2.047	4.143	25453.84	0.6212
17 months: hypothalamus: EEDR	-4.852	4.652	25453.84	0.2969
5 months: thalamus: EEDR	-1.331	3.945	25453.84	0.7358
11 months: thalamus: EEDR	5.083	4.143	25453.84	0.2199
17 months: thalamus: EEDR	-6.561	4.652	25453.84	0.1584
5 months: globus pallidus: EEDR	1.313	3.945	25453.84	0.7394
11 months: globus pallidus: EEDR	6.725	4.143	25453.84	0.1046
17 months: globus pallidus: EEDR	4.699	4.652	25453.84	0.3124
5 months: striatum: EEDR	-0.569	3.945	25453.84	0.8853
11 months: striatum: EEDR	7.147	4.143	25453.84	0.0845
17 months: striatum: EEDR	3.944	4.652	25453.84	0.3966
5 months: substantia nigra: EEDR	10.788	3.945	25453.84	0.0063
11 months: substantia nigra: EEDR	-0.200	4.143	25453.84	0.9616
17 months: substantia nigra: EEDR	0.593	4.652	25453.84	0.8986
5 months: colliculi: EEDR	2.377	3.945	25453.84	0.5468
11 months: colliculi: EEDR	-2.111	4.143	25453.84	0.6104
17 months: colliculi: EEDR	-6.427	4.652	25453.84	0.1671

Appendix 4-3: Full results of the linear mixed-effects conditional growth model predicting median CBF from age, group, and ROI

Appendix 4-3: Estimated marginal means of the interaction between group and time across regions for cerebral blood flow

ROIs	Contrast (control – EEDR)	β	Std. Error	P-value
colliculi	5 months - 3 months	-5.9282	10.07	0.5561
	11 months - 5 months	18.2644	9.63	0.1602
	17 months - 11 months	13.4915	8.37	0.1602
striatum	5 months - 3 months	-2.9819	10.07	0.7672
	11 months - 5 months	6.06	9.63	0.7672
	17 months - 11 months	12.379	8.37	0.4168
globus pallidum	5 months - 3 months	-4.8636	10.07	0.6291
	11 months - 5 months	8.364	9.63	0.5775
	17 months - 11 months	11.2013	8.37	0.5417
hypothalamus	5 months - 3 months	-0.0439	10.07	0.9965
	11 months - 5 months	12.3167	9.63	0.3012
	17 months - 11 months	11.9802	8.37	0.3012
basal forebrain	5 months - 3 months	-5.9045	10.07	0.5577
	11 months - 5 months	10.0056	9.63	0.4775
	17 months - 11 months	8.3475	8.37	0.4775
insula	5 months - 3 months	-1.2464	10.07	0.9015
	11 months - 5 months	9.1402	9.63	0.5137
	17 months - 11 months	14.4235	8.37	0.254
olfactory	5 months - 3 months	-9.3186	10.07	0.5245
	11 months - 5 months	6.1278	9.63	0.5245
	17 months - 11 months	15.0044	8.37	0.2186
piriform	5 months - 3 months	-1.5954	10.07	0.8741
	11 months - 5 months	19.4616	9.63	0.1298
	17 months - 11 months	-2.5815	8.37	0.8741
rhinal	5 months - 3 months	0.8032	10.07	0.9364
	11 months - 5 months	17.4921	9.63	0.2078
	17 months - 11 months	9.1125	8.37	0.414
sensory	5 months - 3 months	-2.0663	10.07	0.8374
	11 months - 5 months	10.5688	9.63	0.4085
	17 months - 11 months	13.4306	8.37	0.3251
parietal	5 months - 3 months	-1.2572	10.07	0.9007
	11 months - 5 months	7.4281	9.63	0.6606
	17 months - 11 months	14.367	8.37	0.2577
visual	5 months - 3 months	-3.551	10.07	0.7244
	11 months - 5 months	13.7765	9.63	0.4091
	17 months - 11 months	9.1752	8.37	0.4091

Appendix 4.3: contrasting estimated marginal means of median cerebral blood flow. This table includes ROIs that did not have significant differences when comparing group (pairwise, control – EEDR) and time (consecutively)

Appendix 5-1: Physical and behavioural metrics between aged impaired and unimpaired groups

Appendix 5-1: comparing physical and behavioural metrics for rats classified as impaired or unimpaired on attentional set-shifting task

Appendix 5-2: Default mode connectivity by outcome

Appendix 5-2: Mean Fisher-transformed correlation coefficient for DMN regions for the end of study (upper triangle) and poor health groups (lower triangle) for 3, 5, and 11 months.