Charting human subcortical maturation across the adult lifespan with *in vivo* 7 T MRI

Steven Miletic^{a,1,*}, Pierre-Louis Bazin^{a,b,1}, Scott J. S. Isherwood^a, Max C. Keuken^a,
Anneke Alkemade^{a,2}, Birte U. Forstmann^{a,2,*}

^{1,2}Authors contributed equally

^aUniversity of Amsterdam, Department of Psychology, Integrative Model-based Cognitive Neuroscience research unit (IMCN), Nieuwe Achtergracht 129B, Amsterdam, The Netherlands

^bMax Planck Institute for Human Cognitive and Brain Sciences, Departments of Neurophysics and Neurology, Stephanstraße 1A, Leipzig, Germany

*Corresponding authors:

Steven Miletic, Nieuwe Achtergracht 129B, 1001 NK Amsterdam, The Netherlands, s.miletic@uva.nl

Birte U. Forstmann, Nieuwe Achtergracht 129B, 1001 NK Amsterdam, The Netherlands, buforstmann@gmail.com

Declarations of interest: None

17 **Abstract**

18 The human subcortex comprises hundreds of unique structures. Subcortical functioning
19 is crucial for behavior, and disrupted function is observed in common neurodegenerative
20 diseases. Despite their importance, human subcortical structures continue to be difficult
21 to study *in vivo*. Here we provide a detailed account of 17 prominent subcortical struc-
22 tures, describing their approximate iron and myelin contents, morphometry, and their
23 age-related changes across the normal adult lifespan. The results provide compelling
24 insights into the heterogeneity and intricate age-related alterations of these structures.
25 They also show that the locations of many structures shift across the lifespan, which is of
26 direct relevance for the use of standard magnetic resonance imaging atlases. The results
27 further our understanding of subcortical morphometry and neuroimaging properties,
28 and of normal aging processes which ultimately can improve our understanding of
29 neurodegeneration.

30 **1 Introduction**

31 The human subcortex comprises hundreds of unique structures (Alkemade, Keuken, &
32 Forstmann, 2013; Forstmann, de Hollander, Van Maanen, Alkemade, & Keuken, 2017)
33 which receive interest from a broad range of neuroscientific disciplines (e.g. Lozano
34 et al., 2019; Raznahan et al., 2014; G. M. Shepherd, 2013; Tian, Margulies, Breakspear, &
35 Zalesky, 2020). Subcortical functioning is crucial for normal behavior and physiology
36 including decision making (Ding & Gold, 2013), reward processing (O’Doherty et al.,
37 2004; Schultz, Dayan, & Montague, 1997), and motor behavior (Mink, 1996). Disruption
38 of subcortical structures is observed in common neurodegenerative diseases including
39 Parkinson’s (Hirsch, Graybiel, & Agid, 1988) and Alzheimer’s disease (Ehrenberg et al.,
40 2017; German, White, & Sparkman, 1987). Subcortical structures are also of interest
41 as (potential) deep brain stimulation (DBS) targets in Parkinson’s disease (Fasano &
42 Lozano, 2015; Limousin et al., 1995) and other disorders such as major depression and
43 epilepsy (Lozano et al., 2019).

44 Research into the subcortex depends on the imaging of individual subcortical struc-
45 tures. However, visualizing subcortical structures using *in vivo* methods such as mag-
46 netic resonance imaging (MRI) is challenging due to their close spatial proximity, bio-
47 physical properties, and morphometry (Keuken, Isaacs, Trampel, van der Zwaag, &
48 Forstmann, 2018). As a consequence, our understanding of the subcortex remains lim-
49 ited, and lags behind our understanding of the cortex. Quantitative ultra-high field 7

50 Tesla MRI provides a method to overcome the challenges associated with visualizing
51 subcortical structures (Bazin, Alkemade, Mulder, Henry, & Forstmann, 2020; Keuken
52 et al., 2018), which we use here to provide a cross-sectional account of the subcortex
53 across the adult lifespan.

54 The biophysical properties that determine the appearance of brain structures on MR
55 images include the iron and myelin contents, which influence the main sources of
56 contrast in MRI: the longitudinal and effective transverse relaxation rates, and the local
57 susceptibility to magnetic fields. Furthermore, iron and myelin are highly biologically
58 relevant: Myelin plays an important role in plasticity and development (e.g. Fields,
59 2015; Hill, Li, & Grutzendler, 2018; Turner, 2019), and iron is crucial for normal tissue
60 functioning (e.g. Zecca, Youdim, Riederer, Connor, & Crichton, 2004). Iron deposition
61 (Daugherty & Raz, 2013; Hallgren & Sourander, 1958; Raz & Rodrigue, 2006; Ward,
62 Zucca, Duyn, Crichton, & Zecca, 2014; Zecca et al., 2004) and decreased myelination (Raz
63 & Rodrigue, 2006; Shen et al., 2008) are part of normal aging processes, but excessive iron
64 accumulation and myelin degradation are prominent in diseases including Parkinson's
65 and Alzheimer's disease (e.g. Mancini et al., 2020; Zecca et al., 2004). A description of
66 normal age-related changes in iron and myelin content can therefore provide a frame
67 of reference to contrast pathological iron accumulation and myelin degradation, and to
68 refine methods for the early detection of pathological alterations using MRI measures as
69 biomarkers.

70 An additional factor determining the appearance of the human subcortex is the
71 small size of the individual structures. Prominent subcortical structures such as the
72 subthalamic nucleus are as small as a few millimeters thick, limiting the number of voxels
73 they encompass on MR images commonly used in research and in the clinic. Moreover,
74 voxels at the border of structures likely include tissue from adjacent structures (*partial*
75 *voluming*), which can lead to biases especially when voxel sizes are large relative to
76 the structure (Mulder, Keuken, Bazin, Alkemade, & Forstmann, 2019). Structure size
77 should therefore be taken into account when imaging the subcortex. An important
78 additional consideration here is the development of atrophy with increasing age, which
79 is reflected in reduced volume of grey matter structures (Cherubini, Pérán, Caltagirone,
80 Sabatini, & Spalletta, 2009; Courchesne et al., 2000; Herting et al., 2018; Lemaitre et al.,
81 2012; Raz, 2004; Raz & Rodrigue, 2006; Walhovd et al., 2005) and which results in more
82 cerebrospinal fluid (CSF) and larger ventricles (Good et al., 2001; Greenberg et al., 2008;
83 Stafford, Albert, Naeser, Sandor, & Garvey, 1988; Walhovd et al., 2005). In addition to
84 volume changes, atrophy can result in a shift in the location of structures (Keuken et al.,
85 2013; Keuken et al., 2017; Kitajima et al., 2008).

These factors combined hamper visualization of the subcortex when using conventional magnetic resonance imaging techniques. Furthermore, the age-related alterations in these factors alter the appearance of the subcortex with increasing age. In this study, we provide a detailed account of 17 subcortical structures using data from 105 healthy participants across the adult lifespan obtained with *in vivo* methods tailored for studying the human subcortex (Alkemade, Mulder, et al., 2020). We first estimate the iron and myelin distributions and size of individual structures. We next analyze age-related changes: iron deposition, myelin degradation, and the morphometric changes that are the result of atrophy (changes in size, shape, and location).

2 Results

One hundred and five healthy volunteers were scanned using an ultra-high field 7 Tesla MRI scanner as part of the openly available Amsterdam ultra-high field adult lifespan database project (AHEAD; Alkemade, Mulder, et al., 2020). A quantitative, multi-modal MP2RAGE-ME sequence (Caan et al., 2019) with 0.7 mm isotropic resolution was used to simultaneously estimate R1, R2* and quantified susceptibility mapping (QSM) values in a single scanning sequence. For each participant, 17 subcortical structures (see Table 1) were delineated using the Multi-contrast Anatomical Subcortical Structures Parcellation method (MASSP; Bazin et al., 2020).

We analyzed each structure by first estimating the *iron* and *myelin* concentrations, using simplified biophysical models that translate the measured R1, R2*, and QSM values into the most likely corresponding iron and myelin concentrations (see Methods). We obtained both the (within-structure) median of iron and myelin distributions, and the interquartile range (IQR) which reflects image noise and tissue (in)homogeneity. Second, we analyzed the structure morphometry by estimating *volume* and *thickness*. Thickness is defined as twice the distance between the boundary and the internal skeleton of the structure. As a local measure (contrary to volume), it is defined for every voxel in a structure, and it depends on the structure's shape. Also for thickness, we determined both the median and IQR, the latter reflecting the regularity of the structure's shape: Regularly shaped structures (e.g., the red nucleus) have a similar thickness at each voxel's location, resulting in lower between-voxel IQRs compared to complex shaped structures (e.g., the striatum). Third, we determined the *location* (center of mass in 3 Cartesian coordinates) of each structure. Center of mass was determined after applying an affine transformation to a group template, to account for inter-individual differences in intracranial volume and shape, while retaining inter-individual variability in distances

Table 1: Regions of interest. Midline structures were parcellated as a single structure, all other structures (indicated by bold-faced letters) were parcellated separately per hemisphere. Abbreviations in italics indicate white matter structures.

AMG: Amygdala	SN: Substantia nigra
CL: Claustrum	STN: Subthalamic nucleus
<i>fx:</i> Fornix	STR: Striatum
GPe: Globus Pallidus Externa	THA: Thalamus
GPI: Globus Pallidus Interna	VTA: Ventral Tegmental Area
<i>ic:</i> Internal Capsule	LV: Lateral ventricle
PAG: Periaqueductal grey	3V: Third ventricle
PPN: Pedunculopontine nucleus	4V: Fourth ventricle
RN: Red nucleus	

relative to the neurocranium.

The distributions of iron, myelin, and volumes revealed a large between-structure heterogeneity in the human subcortex (Figure 1). The globus pallidus externa and interna, red nucleus, substantia nigra, and subthalamic nucleus displayed the highest iron concentrations (both median and IQR), corroborating earlier reports (Haacke et al., 2005; Hallgren & Sourander, 1958; Ramos et al., 2014). In line with expectations, low iron concentrations in combination with high myelin concentrations were observed in the white matter structures under study: the internal capsule and the fornix. The estimated myelin concentrations of the subthalamic nucleus, red nucleus, and ventral tegmental area were relatively high, which causes the limited visibility of these structures on T1-weighted images (Keuken et al., 2018). For comparison, the estimated myelin concentrations of the striatum and amygdala were substantially lower, resulting in intensities comparable to cortical grey matter on T1-weighted images. The within-structure IQR of iron roughly scaled with the median estimates, indicating that structures with high iron concentrations tended to be less homogeneous. This also held for the myelin concentrations, with the exception of the fornix, which had a particularly high IQR. This could potentially have been caused by partial voluming with the lateral ventricles, decreasing the myelin estimates at voxels near the boundary of the fornix. Finally, the striatum, thalamus, lateral ventricles, and internal capsule were the largest structures, and the subthalamic nucleus the smallest.

2.1 Maturation effects

We next studied the age-related alterations in iron, myelin, and morphometry across the adult lifespan. We fit a set of 24 regression model specifications (with, as predictor

Figure 1: Across-participant distributions of within-structure median and IQR of iron and myelin, and volumes per structure. The center line in each box marks the median, box limits are the across-participant IQR, and whiskers are at 1.5 times the IQR below and above the box limits. Error bars drawn inside boxes indicate 95% confidence intervals around the median, obtained by bootstrapping with 10,000 iterations. Colors indicate hemisphere (LH = left hemisphere, RH = right hemisphere, Single = structures that are continuous across the hemispheres), ppm = parts per million, mm = millimeter.

variables, linear and/or quadratic effects of age, plus sex and potential interactions between sex and age) for all structures and measures individually. As we had no *a priori* hypothesis on lateralization, we collapsed across hemispheres to reduce the total number of fitted models. The model specification that showed the most parsimonious trade-off between quality of fit and model complexity (as quantified using the Bayesian information criterion; Schwarz, 1978) was considered the winning model and used for further analyses. To help navigate the winning models of each structure and measure (including $R1$, $R2^*$, and QSM values), we developed an online interactive app, which is accessible at <https://subcortex.eu/app> (see also Figure S1).

We observed (median) iron accumulation in all structures except for the claustrum, globus pallidus interna, and periaqueductal gray, which instead showed stable iron concentrations (Figure 2). With the exception of the globus pallidus interna, the iron-rich basal ganglia appeared to accumulate most iron during aging in absolute terms. The IQRs increased with age for all structures, revealing a global decrease in structure homogeneity. Since this decrease in homogeneity was also present in the structures where no median iron increase was observed, it likely partially reflects an increase in image noise. However, the increases in IQR were substantially higher in the structures that accumulate most iron, suggesting that the (median) iron accumulation for these structures was not homogeneously distributed within the structure. This decrease in

162 homogeneity was particularly strong in striatum and the red nucleus.

163 In line with expectations, we observed a general myelin degradation (see Figure 3),
 164 except for in the amygdala, claustrum, and substantia nigra, where no alterations in
 165 myelin concentrations were detected. The largest (absolute) reduction of myelin was
 166 present in the fornix; the other white matter structure, internal capsule, showed a smaller
 167 decrease in myelin. The globus pallidus interna, periaqueductal grey, pedunclopontine
 168 nucleus, substantia nigra, and ventral tegmental area showed slightly higher median
 169 myelin concentrations in females than in males. Like in the case of iron, the increases in
 170 IQR of myelin reflected a trend of decreasing structure homogeneity across structures.
 171 Since these IQR increases were present for structures that did not show any change in me-
 172 dian myelin content, they likely partially reflect increases in image noise. However, the
 173 structures with large median myelin decreases also tended to show the largest increases
 174 in IQR, suggesting that these IQR increases also reflect increasing inhomogeneities in
 175 myelin concentrations.

176 Next, we analyzed the effects of atrophy (Figure 4). The lateral and third ventricle
 177 showed a substantial volume increase with age, which can at least partially be explained
 178 by the filling of the intracranial space created by atrophied brain tissue. Contrary
 179 to expectations, the volume of the fourth ventricle decreased rather than increased.
 180 Inspection of the mesh of the fourth ventricle in the elderly suggests this may be caused
 181 by shrinkage of the superior part. Volume decreases were also found in the striatum,
 182 thalamus, amygdala, ventral tegmental area, periaqueductal grey, pedunclopontine
 183 nucleus, and red nucleus, likely reflecting atrophy. The internal capsule, fornix and
 184 globus pallidus interna showed a small increase in volume with age, suggesting white
 185 matter swelling, which could be caused by neuroinflammatory processes.

186 Atrophy of specific subparts of a structure, as a result of increased vulnerability to
 187 atrophy in that part, could result in shape changes (Ho et al., 2020; Raznahan et al.,
 188 2014). Shape changes can be detected by analyzing changes the median and IQR of
 189 thickness, which depend on the structure's shape. Specifically, when changes in the
 190 median thickness and volume point in the same direction (as is the case in, e.g., the lateral
 191 ventricles, striatum), this suggests overall thickening or thinning of a structure. Instead,
 192 increases in median thickness combined with decreases in volume can indicate atrophy
 193 in a thinner part of the structure, as this would decrease the amount of voxels with
 194 relatively low thickness, increasing the median thickness. This specific effect appeared to
 195 be present in the ventral tegmental area, pedunclopontine nucleus and periaqueductal
 196 grey. Furthermore, increases in IQR indicate decreases in structure regularity, which
 197 was observed in the globus pallidus interna, substantia nigra, periaqueductal grey,

Figure 2: Age-related changes in iron content. The meshes are based on the young (18–30 years old; left) and elderly (70–80 years old; right) participants after a non-linear transformation to MNI2009b space. Mesh colors illustrate the model predictions for the median and IQR of iron distributions at 25 (left) and 75 (right) years old. Colors in the top-left meshes of all structures indicate model predictions at 25 years old. In case the winning model did not include sex as a predictor variable, the model predictions are shown in black lines; otherwise, green and orange lines are used for the predictions for women and men, respectively. The total amount of change in median (Med.) and IQR are shown in each scatterplot. The ventricles are assumed to have no iron and are excluded from this graph.

Figure 3: Age-related changes in myelin content. The meshes are based on the young (18–30 years old; left) and elderly (70–80 years old; right) participants after a non-linear transformation to MNI2009b space. Mesh colors illustrate the model predictions for the median and IQR of myelin distributions at 25 (left) and 75 (right) years old. Colors in the top-left meshes of all structures indicate model predictions at 25 years old. In case the winning model did not include sex as a predictor variable, the model predictions are shown in black lines; otherwise, green and orange lines are used for the predictions for women and men, respectively. The total amount of change in median (Med.) and IQR are shown in each scatterplot. The ventricles are assumed to have no myelin and are excluded from this graph.

198 pedunculo pontine nucleus, and red nucleus.

199 A third potential effect of atrophy is a change in the location of individual structures
200 relative to the neurocranium (Keuken et al., 2013; Keuken et al., 2017; Kitajima et al.,
201 2008): As the brain atrophies, the resulting physical space is filled with CSF, leading
202 to location shifts of other brain structures. For the majority of brain structures under
203 investigation, we observed location shifts in the lateral and inferior direction (Figure 5).
204 The center of mass of the lateral and third ventricles and the claustrum also shifted in
205 the posterior direction; the fornix and striatum shifted in anterior direction.

206 Combined, we observed age-related changes in all measures: iron, myelin, volume,
207 thickness, and location. These effects were in line with the expected effects of iron
208 accumulation, myelin degradation, and atrophy, but there appeared to be strong between-
209 region variability in the degree to which regions change with age, which we focus on in
210 the next section.

211 2.2 Between-structure variability in maturation

212 Because the winning models of age-related change trajectories included either linear or
213 quadratic influences of age, the parameter estimates of the different models cannot be
214 directly compared. To provide a quantity that summarizes the amount of age-related
215 change (irrespective of the underlying model specification), we calculated the sum of the
216 absolute yearly changes between 19 and 75 years old, relative to the model's predicted
217 value at 19 years old to take into account baseline differences. This *total age-related*
218 *change* metric (also shown in Figures 2—4 and S2) revealed that the strongest age-related
219 changes, across metrics, were observed in the striatum and red nucleus, driven by
220 tissue degeneration and iron accumulation (Figure 6). The pedunculo pontine nucleus,
221 claustrum, and internal capsule remained relatively stable across the age range.

222 Sex differences in aging effects were present in the measures that reflect atrophy, where
223 we observed larger ventricular volume and thickness increases in males compared
224 to females, and slightly larger volume and thickness decreases in striatum, ventral
225 tegmental area, and red nucleus in females compared to males. Sex differences in
226 atrophy rates have been observed before, although the size and direction of sex effects
227 appears to depend on the specific structure (Li et al., 2014; Wang, Xu, Luo, Hu, & Zuo,
228 2019; Xu et al., 2000). In the iron and myelin measures, sex differences were relatively
229 modest and confined to specific regions. The most prominent were in the thalamus
230 (males but not females showed a decrease in the IQR of myelin), the periaqueductal
231 grey (males but not females showed median myelin decrease and IQR increase), internal

Figure 4: Age-related changes in structure morphometry. The meshes are based on the young (18–30 years old; left) and elderly (70–80 years old; right) participants after a non-linear transformation to MNI2009b space. The lines in each scatterplot visualize the winning model predictions. In case the winning model did not include sex as a predictor variable, the model predictions are shown in black lines; otherwise, green and orange lines are used for the predictions for women and men, respectively. The total amount of change in median (Med.) and IQR of thickness and volume (Vol.) are shown in each scatterplot.

Figure 5: Age-related changes in structure location, posterior view. Meshes were based on the young (18–30 years old) participants after non-linearly transforming to MNI2009b space. Arrows depict the model predictions for the location shift, starting at the center of mass of each structure in MNI2009b space, and pointing to the predicted center of mass of the structure at 75 years old. The left graph shows all 17 subcortical structures under investigation, the right graph excludes the lateral ventricles, internal capsule, and thalamus, to improve the visibility of the smaller structures.

Figure 6: Total change per morphometric measure between 19 and 75 years, relative to the measure's predicted value at 19. Error bars indicate standard errors, obtained by bootstrapping with 10,000 iterations.

capsule (males but not females showed median myelin decrease), and pedunclopontine nucleus (males but not females showed alterations in iron and myelin). These sex differences, however, were modest, and the overall aging patterns were comparable across sexes.

Finally, we tested whether groups of structures with correlated aging patterns could be identified. We computed all pairwise correlations between structures using their total age-related changes (in median and IQR in iron, myelin, and thickness, as well as volume) as outcome variables (i.e., using the matrix visualized in Figure S2 as input, excluding the location changes). The correlation matrix was then organized using data-driven hierarchical clustering, the result of which is shown in Figure 7. The predominantly positive correlations reiterate the presence of a global, across-region aging process. Furthermore, there are clusters of regions that covary strongly. One such cluster consists of the ventricular system, which ages differently compared to other regions, due to the volume increases rather than decreases. Contrary to expectations, the other clusters that were identified did not group anatomically or functionally related regions such as the basal ganglia or the dopaminergic midbrain structures, although some functionally related nodes do show high pairwise correlations (e.g., subthalamic nucleus and substantia nigra). Instead, the largest cluster of regions with similar aging patterns included the striatum, red nucleus, claustrum, internal capsule, globus pallidus externa.

3 Discussion

Interest in the human subcortex is rapidly growing in cognitive and clinical neuroscience due to the relevance of subcortical regions as (potential) targets for DBS and their role in cognition. Here, we studied 17 subcortical structures in terms of their iron and myelin contents, their sizes, as well as the intricate age-related alterations. Our results highlight the heterogeneity in the subcortex, presenting the strong variability in iron, myelin, and morphometry that exists between structures. Furthermore, our results indicate global effects of iron accumulation, myelin degradation, and atrophy in the subcortex across the normal adult lifespan, and strong variability in the extent to which the different structures are affected by each type of age-related change. The red nucleus and striatum show most profound changes, whereas some other structures (such as the claustrum and pedunclopontine nucleus) appear less affected by aging processes.

Furthermore, we observe clusters of regions with similar aging patterns, but these clusters (except for the ventricular system) do not show similarity between the aging

Figure 7: Correlation matrix indicating the similarities and differences between structures' aging patterns. Regions are ordered using hierarchical clustering, of which the dendrograms are shown on the left and top of the correlation matrix. Colors indicate correlation coefficients.

266 patterns of groups of regions that are functionally or anatomically related. For example,
 267 it could be hypothesized that there are similarities between the aging processes of the
 268 globus pallidus externa and subthalamic nucleus, based on their structural connectivity,
 269 or between the ventral tegmental area and substantia nigra, as these are the dopaminergic
 270 midbrain areas. However, the correlations in aging patterns between such related regions
 271 were not particularly strong, and substantially weaker than those between regions that
 272 are anatomically and functionally less related such as the thalamus and the ventral
 273 tegmental area. This suggests that functional or anatomical relations between regions do
 274 not imply similar aging patterns, and that inferences about aging of one region cannot
 275 be drawn on the basis of aging patterns of similar regions.

276 To better navigate the rich landscape of subcortical aging, we also share our results in
 277 an online app (Figure S1, <https://subcortex.eu/app>) that can be used to create interactive
 278 and intuitive 3D visualizations of the human subcortex across the lifespan and across
 279 modalities. It allows for inspection and reuse of the underlying models and data of each
 280 individual structure. The app was designed in a flexible way, so that it can be augmented
 281 with more structures and contrasts to expand it to a comprehensive chart of the human
 282 subcortex. The underlying data can readily be downloaded for further analyses.

283 Understanding the aging processes in the subcortex is paramount in research and in
 284 clinical practice. While iron accumulation and myelin degradation are part of normal

aging processes, increased accumulation and myelin degradation are part of multiple neurodegenerative disorders including Parkinson's and Huntington's disease (Andersen, Johnsen, & Moos, 2014; Collingwood & Davidson, 2014; Ward et al., 2014; Zecca et al., 2004). An accurate description of the distributions of iron and myelin across the lifespan in health provides a frame of reference against which pathological iron accumulation and myelin degradation can be contrasted, and can prove useful in the development of biomarkers for disease (Guan, Xu, & Zhang, 2017; Mancini et al., 2020; Martin, Wieler, & Gee, 2008; Schenck & Zimmerman, 2004; Zecca et al., 2004).

Iron and myelin are also the two main determinants of MRI contrast. Many subcortical structures, such as the subthalamic nucleus, cannot readily be distinguished on conventional T1-weighted MRI images due to a lack of contrast with nearby regions. Because of the limited visibility of subcortical structures on conventional MR images, a common practice is to use atlases to locate individual structures (Devlin & Poldrack, 2007; Evans, Janke, Collins, & Baillet, 2012). Stereotactic atlases based on *post mortem* tissue are often used for planning DBS surgery, and probabilistic MRI atlases are conventionally used in cognitive neuroscientific research. Subcortical MRI atlases are growing in numbers (Keuken et al., 2014; Lau et al., 2020; Pauli, Nili, & Tyszka, 2018; Trutti et al., 2021; Ye et al., 2021) due to improvements in MRI resolution and contrasts. However, MRI atlases are typically developed using anatomical images obtained from young participants, which can cause biases when such atlases are subsequently used to infer anatomical information in older participants or patient populations (Evans et al., 2012; Keuken et al., 2013; Samanez-Larkin & D'Esposito, 2008). In cognitive neuroscience research, it is common to apply spatial normalization procedures to a group space to account for individual differences in anatomy, but consistent deviations from the group template are likely to introduce normalization errors (Samanez-Larkin & D'Esposito, 2008). These biases can result from iron accumulation and myelin degradation (which change the contrast of images) and from atrophy (which change the size and the location of structures). Our results can help understand the biases that could occur when conventional MRI atlases, based on young participants, are used to analyze data from older participants, and call for the development of age-specific MRI atlases of the subcortex to remedy these biases.

The between-region variation in iron contents has important consequences for blood oxygenation level dependent (BOLD) functional MRI. Because iron decreases $T2^*$ relaxation times, on which contrast-to-noise ratios (CNR) of BOLD-fMRI sequences depend, BOLD CNR varies substantially between regions. For instance, within young participants, the CNR in the red nucleus is expected to be 72% lower than in the amygdala,

when using an echo time of 42 ms at 7 T (corresponding to the $T2^*$ of the amygdala in young participants), solely due to the differences in $T2^*$ (see supplement for details). Age-related alterations in iron contents can have similar effects. For instance, the CNR in the red nucleus at 50 years old is 28% lower than at 19 years old when using an echo time of 18 ms at 7 T (corresponding to the $T2^*$ of the red nucleus at 19 years old). Thus, iron deposition can confound fMRI studies into age-related changes of BOLD responses.

However, substantial gains in CNR can be achieved by optimizing the echo time to meet the specific requirements of studying a structure of interest (see also de Hollander, Keuken, van der Zwaag, Forstmann, & Trampel, 2017; Miletić, Bazin, et al., 2020). For instance, when targeting the red nucleus, decreasing the echo time to 18 ms (corresponding to the $T2^*$ of the red nucleus at 19 years old) is expected to lead to a 58% higher CNR compared to an echo time of 42 ms (which would be optimal to target the amygdala). Similarly, the echo time can be adjusted to partially mitigate the effects of age-related changes in $T2^*$: By decreasing the echo time from 18 ms to 13.7 ms (corresponding to the $T2^*$ of the red nucleus at 50 years old), a modest 5% increase in CNR can be expected. Using our online app as a resource for participant-specific predictions of $R1$, $R2^*$, and QSM values, we envision the use of MRI protocols tailored to the structure of interest and the participant's age and sex.

The present study has several limitations. First, iron and myelin contents were approximated using simplified biophysical models (c.f. Hametner et al., 2018; Mangeat, Govindarajan, Mainero, & Cohen-Adad, 2015; Marques, Khabipova, & Gruetter, 2017; Metere & Möller, 2018; Rooney et al., 2007; Stüber et al., 2014), which rely on population-average iron and myelin estimates found in the literature (Hallgren & Sourander, 1958; Metere & Möller, 2018; Randall, 1938). Since the amount of structures for which myelin concentrations are known is limited, we also relied on *post mortem* materials to infer myelin concentrations. While the biophysical models provided a good fit to our data, the resulting iron and myelin estimates could be biased by the relatively limited information on myelin concentrations in the literature. Such potential biases should be present for all participants and therefore not affect the age-related alterations we report. The approach of estimating myelin and iron concentrations based on quantitative MRI using simplified biophysical models would benefit from more quantitative myelin and iron estimates obtained through analyzing *post mortem* tissue.

Second, we cannot exclude age-related changes in parcellation accuracy. We relied the MASSP algorithm (Bazin et al., 2020) to parcellate the 17 subcortical regions in each participant individually. The performance of MASSP (compared to manual delineations) tends to decrease with age. Fortunately, the impact of age-related biases in parcellation

was shown to be limited for the quantitative MRI measures (Bazin et al., 2020) on which the iron and myelin estimates are based, suggesting that the age-related changes in myelin and iron are unlikely to be caused by age-related differences in parcellation performance. On the other hand, size estimates (volume and to a lesser extent thickness) and more susceptible to the age-related changes in parcellation quality. Here, we used an improved version of MASSP to mitigate these effects, although we cannot exclude any remaining age-related decreases in parcellation accuracy.

Our results extend previous studies into aging patterns of the subcortex, which focus on a smaller number of typically large subcortical areas, often based on MRI with lower field strengths (Aquino et al., 2009; Cherubini et al., 2009; Daugherty & Raz, 2013, 2016; Fjell et al., 2013; Greenberg et al., 2008; Herting et al., 2018; Keuken et al., 2017; Li et al., 2014; Raz, 2004; Raz et al., 2005; Raz & Rodrigue, 2006; Raznahan et al., 2014; Walhovd et al., 2005; Wang et al., 2019). Experiments using very large numbers of participants detected complex nonlinear age-related changes in some subcortical structures (Fjell et al., 2013; Raznahan et al., 2014), suggesting the presence of critical ages that demark changes in the rates of atrophy in some subcortical areas (Fjell et al., 2013). Our study had a more modest sample size, which did not allow to evaluate complex non-linear trends. On the other hand, by leveraging an open database of ultra-high field 7 T quantitative MRI, we could provide a first view on many structures and variables at once, which may be refined as more 7 T quantitative MRI data sets become available. As such, our study provides a richer and more extensible description of subcortical composition, morphometry and aging.

4 Methods

4.1 Participants

We used the Amsterdam ultra-high field adult lifespan database (AHEAD; Alkemade, Mulder, et al., 2020), which consists of multimodal MRI data from 105 healthy participants. Inclusion criteria were age 18–80 years and self-reported health at the time of inclusion. Exclusion criteria were any factors that could potentially interfere with MRI scanning, including MRI incompatibility (e.g., pacemakers), pregnancy, and self-reported claustrophobia. At least six males and females were included in each age decade to ensure full coverage of the adult lifespan. All participants gave written informed consent prior to the onset of data collection. The local ethics board approved the study.

4.2 MRI scanning

Images were acquired at the Spinoza Centre for Neuroimaging in Amsterdam, the Netherlands, using a Philips Achieva 7 T MRI scanner with a 32-channel phased-array coil. Routine quality checks were performed on a weekly basis at the scanner site. T1-weighted, T2* contrasts were obtained using a MP2RAGEME (multi-echo magnetization-prepared rapid gradient echo) sequence (Caan et al., 2019). The MP2RAGEME is an extension of the MP2RAGE sequence (Marques et al., 2010) and consists of two rapid gradient echo (GRE_{1,2}) images that are acquired in the sagittal plane after a 180 degrees inversion pulse and excitation pulses with inversion times $TI_{1,2} = [670 \text{ ms}, 3675.4 \text{ ms}]$. A multi-echo readout was added to the second inversion at four echo times ($TE_1 = 3 \text{ ms}$, $TE_{2,1-4} = 3, 11.5, 19, 28.5 \text{ ms}$). Other scan parameters include flip angles $FA_{1,2} = [4, 4]$ degrees; $TR_{GRE1,2} = [6.2 \text{ ms}, 31 \text{ ms}]$; bandwidth = 404.9 MHz; $TR_{MP2RAGE} = 6778 \text{ ms}$; acceleration factor SENSE PA = 2; FOV = $205 \times 205 \times 164 \text{ mm}$; acquired voxel size = $0.7 \times 0.7 \times 0.7 \text{ mm}$; acquisition matrix was 292×290 ; reconstructed voxel size = $0.64 \times 0.64 \times 0.7 \text{ mm}$; turbo factor (TFE) = 150 resulting in 176 shots; Total acquisition time = 19.53 min.

4.3 Quantitative MRI modeling and parcellation

The MP2RAGEME consists of two interleaved MPRAGEs with different inversions and four echoes in the second inversion. Based on these images, we estimated quantitative MR parameters of R1, R2* and QSM as follows. First, we took advantage of the redundancy in the MP2RAGEME sequence to perform a PCA-based denoising with LCPCA (Bazin et al., 2019). R1 maps were then computed using a look-up table (Marques et al., 2010). R2*-maps were computed by least-squares fitting of the exponential signal decay over the four echoes of the second inversion. QSM images were obtained from the phase maps of the second, third, and fourth echoes of the second inversion with TGV-QSM (Langkammer et al., 2015). Skull stripping, required for QSM, was performed on the second inversion, first echo magnitude image (Bazin et al., 2014).

The anatomical regions of interest were defined with the MASSP automated algorithm (Bazin et al., 2020) on the basis of the R1, R2* and QSM image maps. The algorithm combines location, shape, and quantitative MRI priors to define 17 subcortical anatomical regions, listed in Table 1. Separate masks for left and right hemisphere were obtained except for 3V, 4V, and fx.

For this study, the MASSP algorithm was trained on renormalized versions of the quantitative contrasts using a fuzzy C-means clustering of intensities, and linearly

interpolating between cluster centroids. The renormalized contrasts were thus less sensitive to the intensity variations induced by aging. Additionally, the registration to the MASSP atlas was performed in two successive steps, producing more accurate alignment of the anatomical priors with each subject.

4.4 Iron and myelin approximation

Iron and myelin are main determinants of MR image contrast (Stüber et al., 2014). Several lines of research indicate that the concentrations of iron and myelin are approximately linearly related to qMRI metrics $R1$, $R2^*$ and QSM (Hametner et al., 2018; Mangeat et al., 2015; Marques et al., 2017; Metere & Möller, 2018; Rooney et al., 2007; Stüber et al., 2014). Assuming a linear relationship between iron and myelin on the one hand, and qMRI on the other, linear models can be fit and used to predict iron and myelin contents based on qMRI values (Metere & Möller, 2018):

$$\begin{aligned} Iron &= Intercept + w_{i,qsm} * QSM + w_{i,R2^*} * R2^* + w_{i,R1} * R1 \\ Myelin &= Intercept + w_{m,qsm} * QSM + w_{m,R2^*} * R2^* + w_{m,R1} * R1 \end{aligned} \quad (1)$$

Estimating the parameters w of these models requires population-average estimates of iron and myelin content for a variety of regions of interest that cover the range of $R1$, $R2^*$, and QSM values observed across the brain. Following the approach by Metere and Möller (2018), we obtained these values from the literature (Hallgren & Sourander, 1958; Metere & Möller, 2018; Randall, 1938), and supplemented those values using observations in *post mortem* tissue (detailed below). For iron estimates, Hallgren and Sourander (1958) provided quantifications in regions of interest that vary in both iron contents and qMRI values, which allows for stable estimators of the weights in Equation 1. An iron concentration of approximately 0 in the ventricles was assumed.

As a reference for myelin concentrations, we used work by Randall (1938), which provides lipid concentrations for the corona radiata, frontal and parietal white matter, brain stem, thalamus, caudate, and frontal and parietal grey matter. Following Metere and Möller (2018), we assumed that these lipid concentrations reflect myelin concentrations. Unfortunately, the reported regions do not include iron-rich nuclei, which limits the range of (especially) $R2^*$ and QSM values with known corresponding lipid concentrations. Using a limited number of regions of interest to estimate the myelin model could limit the generalizability of the estimated parameters to structures with lower $R2^*$ and/or QSM values, which would bias myelin estimates in iron-rich structures like some basal ganglia nodes (e.g., based on using only Randall's (1938) lipid

454 concentrations, Metere and Möller (2018) obtained *negative* myelin concentrations in
455 various basal ganglia structures).

456 To supplement the literature-based myelin concentrations, we approximated the
457 myelin contents of other regions of interest using a *post mortem* specimen. Here, we
458 made the following assumptions:

- 459 1. The optic density of tissue in our silver stains is approximately linearly related to
460 the concentration of myelin in that tissue in our regions of interest (see Figure 8).
461 Here, we confirmed that silver stains were not saturated even in the white matter
462 regions;
- 463 2. The myelin concentrations in the *post mortem* specimen are representative of the
464 population and comparable to the concentrations reported by Randall (1938);
- 465 3. The myelin concentrations in white matter reported by Randall (1938) are com-
466 parable to the myelin concentration in the internal capsule. Similarly, the myelin
467 concentrations in parietal cortex are comparable to insular cortex.

468 Seven 200 μ m coronal sections of a single specimen were stained according to the
469 method described by Bielschowsky (for details, see Alkemade, Pine, et al., 2020). Sections
470 included the caudate nucleus, thalamus, internal capsule, and insular cortex, in which we
471 estimated the median intensity of the lightness of the stain (the optic density). Randall
472 (1938) reports quantified lipid concentrations of the caudate nucleus and thalamus,
473 which can be directly compared to the stain intensities, as well as of parietal grey and
474 white matter. The caudate nucleus, thalamus, and parietal grey and white matter (as
475 reported by Randall (1938)) were not visible in the same histological section, and we
476 therefore used insular cortex as a reference region for grey matter, and the internal
477 capsule as a reference region for white matter. For each section separately, we then
478 created a linear calibration curve, which allowed us to determine lipid concentrations
479 based on the stain intensity (Figure 8) for putamen, globus pallidus, subthalamic nucleus,
480 red nucleus, and substantia nigra.

481 For the region of which population-averaged iron and myelin contents were known,
482 we estimated the qMRI values using the MRI data. Median qMRI values were calculated
483 using the MASSP masks for subcortical regions, and a MGDM and CRUISE parcellation
484 was used to obtain individual masks for brain stem, cerebellum, and cortex (Bazin et al.,
485 2014). We included only participants of 30 years and older to match the ages of the
486 specimens on which the iron and myelin estimates are based. For brain regions for
487 which we had estimated the myelin content using our *post mortem* specimen, we only

Figure 8: Procedure of estimating myelin contents using a *post mortem* specimen. In each section (seven in total, first four shown), the stain intensities corresponding to CAU, THA, insular cortex (CTX) and the internal capsule (ic/wm) were estimated. For each section individually, a calibration curve was estimated to map stain intensity to myelin values (solid blue lines and equations). Within the range of interest, the relation between stain intensity and myelin content could be approximated with a linear trend. Then, the intensity values for putamen (PUT), GP, STN, RN, and SN were estimated per section (colored dashed lines; note that not all sections contained all structures), and the corresponding myelin values were calculated. Boxplots in the right panel show across-section variability in estimated myelin contents and suggest agreement across sections. The center line in each box marks the median, box limits are the across-section interquartile range, and whiskers are at 1.5 times the interquartile range below and above the box limits. ROI = Region of interest.

488 included AHEAD subjects of 70 years and older (17 participants total) to approximately
489 match ages of the MRI data and the specimen. Tables S1 and S2 list the iron and myelin
490 concentrations, respectively, and their corresponding qMRI values, that were used to
491 estimate the parameters in Equations 1.

492 To test whether all qMRI metrics were required as predictors to accurately predict
493 iron and myelin content, we fitted linear models with all eight possible combinations
494 of R1, R2*, and QSM. Models were fitted using ordinary least squares (OLS). For each
495 model, we estimated the Akaike information criterion (AIC; Akaike, 1973) to identify
496 the model that is expected to have the highest predictive performance, and used the
497 model with lowest AIC values (AIC and BIC values agreed on the winning model). We
498 used the AIC here instead of the BIC as the AIC is expected to select models with the
499 highest cross-validated predictive performance, whereas the BIC is expected to select
500 the data-generating model (Wagenmakers & Farrell, 2004). The parametrized winning
501 models were as follows:

$$\begin{aligned} Iron &= -3.828341 + 98.279472 * QSM + 0.243101 * R2^* \\ Myelin &= -7.989650 + 31.874833 * R1 - 0.111816 * R2^* \end{aligned} \quad (2)$$

502 Figure 9 visualizes quality of fit of the winning models and shows that a high propor-

Figure 9: Quality of fit of the myelin (left) and iron (right) model. The planes are given by Equations 2. Red dots illustrate data points, grey dots are the model predictions for these data points.

tion of variance in myelin and iron can be explained using the qMRI values. Note that the model weights cannot directly be compared to the weights from Stüber et al. (2014), which were obtained using formalin fixated *post mortem* tissue. Formalin fixation can change qMRI values (Birkel et al., 2016; Langkammer et al., 2012; Schmierer et al., 2008; T. M. Shepherd et al., 2009; Tovi & Ericsson, 1992). A second complicating factor is that qMRI values can vary between MRI sites (Mancini et al., 2020), suggesting the need to re-estimate model weights when using qMRI obtained at a different site.

Using these simplified biophysical models, we calculated whole-brain iron and myelin maps, and obtained participant-specific myelin and iron values for all structures using the MASSP masks. Iron and myelin maps of a representative participant are shown in Figure 10.

It is important to emphasize that the iron and myelin estimates we report are based on simplifying assumptions with regard to the linearity of the relation between qMRI and iron/myelin, and on the iron/myelin concentrations on which the biophysical models are fitted (detailed above). As such, the iron and myelin estimates should be not be interpreted as absolute measurements, but rather as approximations that serve to guide the interpretations of qMRI values in terms of the most likely underlying biological contributors to those values.

4.5 Thickness estimation

We calculated local structure thickness based on a medial skeleton representation: for each structure, we estimated the skeleton as the ridge equidistant to the structure boundaries. Thickness was defined as twice the distance between the skeleton and the closest boundary, using the method described in Bazin et al. (2020). In other words, local

Figure 10: Example of myelin (left) and iron map (right) of a representative participant. The top row shows the R1, R2*, and QSM maps, which were linearly combined into myelin and iron maps (middle and bottom row) using Equations 2. Note the hyperintense appearance of iron rich structures such as the rounded shape of the red nucleus.

thickness measures at every location inside the structure the distance between the two closest boundaries of that structure, extending the concept of cortical thickness to more complex shapes. Contrary to volume, thickness can be determined at the position of each voxel within a structure, thus providing local information. A similar thickness measure was also used in Ho et al. (2020) to detect subtle shape differences.

4.6 Center of mass

For all structures, we calculated the center of masses in Cartesian x, y, and z coordinates per participant after an affine transformation to group space, using ANTs (Avants, Epstein, Grossman, & Gee, 2008) with mutual information. The affine transformation was applied to correct for inter-individual differences in intracranial volume and neurocranium shape, while retaining inter-individual variability in anatomy relative to the neurocranium.

4.7 Age-related change modelling

We describe the age-related changes in iron concentration, myelin content, volume, and thickness, as well as in the center of mass in x, y and z coordinates. For iron, myelin and thickness, we report both a median reflecting the central tendency and interquartile range reflecting structure homogeneity. For thickness, the interquartile range reflects the within-structure variability of thickness, quantifying the regularity of the shape. We also analyzed the R1, R2*, and QSM values, which can be found in the online app (<https://subcortex.eu/app>).

Exploratory modelling of the between-hemisphere differences per structure suggested no between-hemisphere difference in aging patterns for most structures. Therefore, we subsequently assumed that the age-related changes in each structure were the same in both hemispheres, to reduce the total number of models fitted. We collapsed across hemispheres by taking the mean value across both hemispheres per structure and participant.

Prior to fitting the aging models, we excluded outliers based on their Mahalanobis distance (cut-off 10.827, corresponding to $p < 0.001$, 0.69% of all data points). Per ROI and dependent variable, we then fit the following set of 24 potential models, with all possible combinations of the following predictors: A linear influence of age, a quadratic influence of age, sex, an interaction between age and sex, and an interaction between a quadratic influence of age and sex. We excluded models with both interaction terms, as this would imply implausibly large between-sex differences in aging patterns.

Models were fit with OLS as implemented in statsmodels (Seabold & Perktold, 2010) for the Python programming language. Models were compared with the Bayesian Information Criterion (BIC; Schwarz, 1978), which quantifies the quality of fit penalized for model complexity. Lower BIC values indicate more parsimonious trade-offs between quality of fit and model complexity and are preferred. Based on the winning model, we removed influential data points using Cook's distance (cut-off 0.2, 0.18% of all data points; we used a more conservative cut-off than $4/n$, which is sometimes recommended [Rawlings, Pantula, and Dickey, 1998]). We then refitted all models on the data excluding the influential data points, and performed a new model comparison.

Using the winning age-related change models, we quantified the total age-related changes by summing the yearly absolute changes between 19 and 75 years old (Figures 2–4, S2), divided by the model's prediction at 19 years old. This quantity summarizes the amount of change relative to the baseline value at 19, and allows for comparing models with different specifications (e.g., linear vs quadratic models). For winning models that included sex as a predictor (or interactions between sex and age), the cumulative yearly absolute changes were calculated for both sexes and then averaged. For models where the overall direction of the age-related change was negative (identified by comparing the predicted values at 19 and 75 years old), we took the negative of the total age-related changes. As our data contains only a single data point above 75 years old, we chose to limit the range of ages that we interpret to 19–75 years old, where we expect our models' predictions to be most robust.

To identify similarities in the aging patterns of groups of regions, we calculated all pairwise correlations between regions (based on Figure S2 as input, excluding the location changes)). The resulting correlation matrix was ordered using hierarchical clustering with the UPGMA algorithm with the Euclidean distance as a metric (as default settings in the seaborn visualization package, Waskom and development team, 2020), shown in Figure 7.

4.8 Confidence intervals and standard errors

Confidence intervals in Figure 1 were obtained using a bootstrapping procedure with 10,000 iterations. We iteratively sampled 105 random observations with replacement from the data, based on which we estimated the median, and took the 2.5th and 97.5th percentile of the 10,000 medians as the 95% confidence interval. The standard errors in Figure 6 were obtained using a similar bootstrapping procedure, in which winning model specifications were iteratively fit on 10,000 random samples (drawn with replacement)

from the data. Per iteration, cumulative change metrics were estimated. The standard deviation of the cumulative change metrics across iterations was used as an estimator of the standard error. For winning models that do not include age as a predictor variable, the standard error is 0 since the cumulative age-related change metric is 0 in each iteration.

Open Science

A prior version of the individual qMRI maps has been released as AHEAD (Alkemade, Mulder, et al., 2020) can be found at <https://doi.org/10.21942/uva.10007840.v1>. All derived participant-wise and region-wise measures can be downloaded from our app at <https://subcortex.eu/app>. All code used to estimate the models and produce the figures can be found at <https://osf.io/mvdbbe/>.

Acknowledgements

This research was financially supported by STW/NWO (#14017; BUF and AA), ERC PoC (BUF), NWO Vici (016.Vici.185.052; BUF).

CRediT author statement

Steven Miletić: Conceptualisation, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – Original draft, Writing – Review & Editing, Visualisation, **Pierre-Louis Bazin:** Conceptualisation, Methodology, Investigation, Software, Supervision, Writing – Original draft, Writing – Review & Editing, Formal analysis, Data curation, **Scott Isherwood:** Writing – Review & Editing, **Max C. Keuken:** Writing – Review & Editing, **Anneke Alkemade:** Conceptualisation, Methodology, Validation, Investigation, Resources, Data curation, Writing – Original draft, Writing – Review & Editing, Supervision, Project administration, Funding acquisition, **Birte U. Forstmann:** Conceptualisation, Methodology, Validation, Investigation, Resources, Data curation, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. Petrov & F. Caski (Eds.), *Proceedings of the second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

- 622 Alkemade, A., Keuken, M. C., & Forstmann, B. U. (2013). A perspective on terra incognita:
623 uncovering the neuroanatomy of the human subcortex. *Frontiers in Neuroanatomy*,
624 7(2), 267–277. doi:10.3389/fnana.2013.00040
- 625 Alkemade, A., Mulder, M. J., Groot, J. M., Isaacs, B. R., van Berendonk, N., Lute, N., ...
626 Forstmann, B. U. (2020). The Amsterdam Ultra-high field adult lifespan database
627 (AHEAD): A freely available multimodal 7 Tesla submillimeter magnetic resonance
628 imaging database. *NeuroImage*, 221, 117200. doi:10.1016/j.neuroimage.2020.117200
- 629 Alkemade, A., Pine, K., Kirilina, E., Keuken, M. C., Mulder, M. J., Balesar, R., ... Forstmann,
630 B. U. (2020). 7 Tesla MRI Followed by Histological 3D Reconstructions in Whole-
631 Brain Specimens. *Frontiers in Neuroanatomy*, 14(October), 1–11. doi:10.3389/fnana.
632 2020.536838
- 633 Andersen, H. H., Johnsen, K. B., & Moos, T. (2014). Iron deposits in the chronically
634 inflamed central nervous system and contributes to neurodegeneration. *Cellular
635 and Molecular Life Sciences*, 71(9), 1607–1622. doi:10.1007/s00018-013-1509-8
- 636 Aquino, D., Bizzi, A., Grisoli, M., Garavaglia, B., Bruzzone, M. G., Nardocci, N., ... Chi-
637 apparini, L. (2009). Age-related Iron Deposition in the Basal Ganglia: Quantitative
638 analysis in healthy subjects. *Radiology*, 252(1), 165–172.
- 639 Avants, B., Epstein, C., Grossman, M., & Gee, J. (2008). Symmetric diffeomorphic image
640 registration with cross-correlation: Evaluating automated labeling of elderly and
641 neurodegenerative brain. *Medical Image Analysis*, 12(1), 26–41. doi:10.1016/j.media.
642 2007.06.004
- 643 Bazin, P.-L., Alkemade, A., Mulder, M. J., Henry, A. G., & Forstmann, B. U. (2020). Multi-
644 contrast anatomical subcortical structures parcellation. *eLife*, 9, 1–23. doi:10.7554/
645 eLife.59430
- 646 Bazin, P.-L., Alkemade, A., van der Zwaag, W., Caan, M., Mulder, M. J., & Forstmann,
647 B. U. (2019). Denoising High-Field Multi-Dimensional MRI With Local Complex
648 PCA. *Frontiers in Neuroscience*, 13(October), 1–10. doi:10.3389/fnins.2019.01066
- 649 Bazin, P.-L., Weiss, M., Dinse, J., Schäfer, A., Trampel, R., & Turner, R. (2014). A com-
650 putational framework for ultra-high resolution cortical segmentation at 7 Tesla.
651 *NeuroImage*, 93, 201–209. doi:10.1016/j.neuroimage.2013.03.077
- 652 Birkel, C., Langkammer, C., Golob-Schwarzl, N., Leoni, M., Haybaeck, J., Goessler, W., ...
653 Ropele, S. (2016). Effects of formalin fixation and temperature on MR relaxation
654 times in the human brain. *NMR in Biomedicine*, 29(4), 458–465. doi:10.1002/nbm.
655 3477

- 656 Caan, M. W. A., Bazin, P.-L., Marques, J. P., Hollander, G., Dumoulin, S. O., & Zwaag,
657 W. (2019). MP2RAGEME: T1, T2*, and QSM mapping in one sequence at 7 tesla.
658 *Human Brain Mapping*, 40(6), 1786–1798. doi:10.1002/hbm.24490
- 659 Cherubini, A., Péran, P., Caltagirone, C., Sabatini, U., & Spalletta, G. (2009). Aging
660 of subcortical nuclei: Microstructural, mineralization and atrophy modifications
661 measured in vivo using MRI. *NeuroImage*, 48(1), 29–36. doi:10.1016/j.neuroimage.
662 2009.06.035
- 663 Collingwood, J. F. & Davidson, M. R. (2014). The role of iron in neurodegenerative disor-
664 ders: Insights and opportunities with synchrotron light. *Frontiers in Pharmacology*,
665 5 AUG(August), 1–19. doi:10.3389/fphar.2014.00191
- 666 Courchesne, E., Chisum, H. J., Townsend, J., Cowles, A., Covington, J., Egaas, B., ...
667 Press, G. A. (2000). Normal brain development and aging: Quantitative analysis at
668 in vivo MR imaging in healthy volunteers. *Radiology*, 216(3), 672–682. doi:10.1148/
669 radiology.216.3.r00au37672
- 670 Daugherty, A. M. & Raz, N. (2013). Age-related differences in iron content of subcortical
671 nuclei observed in vivo: A meta-analysis. *NeuroImage*, 70, 113–121. doi:10.1016/j.
672 neuroimage.2012.12.040
- 673 Daugherty, A. M. & Raz, N. (2016). Accumulation of iron in the putamen predicts its
674 shrinkage in healthy older adults: A multi-occasion longitudinal study. *NeuroImage*,
675 128, 11–20. doi:10.1016/j.neuroimage.2015.12.045
- 676 de Hollander, G., Keuken, M. C., van der Zwaag, W., Forstmann, B. U., & Trampel,
677 R. (2017). Comparing functional MRI protocols for small, iron-rich basal ganglia
678 nuclei such as the subthalamic nucleus at 7 T and 3 T. *Human Brain Mapping*, 38(6),
679 3226–3248. doi:10.1002/hbm.23586
- 680 Devlin, J. T. & Poldrack, R. A. (2007). In praise of tedious anatomy. *NeuroImage*, 37(4),
681 1033–1041. doi:10.1016/j.neuroimage.2006.09.055
- 682 Ding, L. & Gold, J. I. (2013). The basal ganglia’s contributions to perceptual decision
683 making. *Neuron*, 79(4), 640–649. doi:10.1016/j.neuron.2013.07.042
- 684 Ehrenberg, A. J., Nguy, A. K., Theofilas, P., Dunlop, S., Suemoto, C. K., Di Lorenzo Alho,
685 A. T., ... Grinberg, L. T. (2017). Quantifying the accretion of hyperphosphorylated
686 tau in the locus coeruleus and dorsal raphe nucleus: the pathological building
687 blocks of early Alzheimer’s disease. *Neuropathology and Applied Neurobiology*, 43(5),
688 393–408. doi:10.1111/nan.12387
- 689 Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases.
690 *NeuroImage*, 62(2), 911–922. doi:10.1016/j.neuroimage.2012.01.024

- 691 Fasano, A. & Lozano, A. M. (2015). Deep brain stimulation for movement disorders.
692 *Current Opinion in Neurology*, 28(4), 423–436. doi:10.1097/WCO.0000000000000226
- 693 Fields, R. D. (2015). A new mechanism of nervous system plasticity: Activity-dependent
694 myelination. *Nature Reviews Neuroscience*, 16(12), 756–767. doi:10.1038/nrn4023
- 695 Fjell, A. M., Westlye, L. T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., ...
696 Walhovd, K. B. (2013). Critical ages in the life course of the adult brain: Non-
697 linear subcortical aging. *Neurobiology of Aging*, 34(10), 2239–2247. doi:10.1016/j.
698 neurobiolaging.2013.04.006
- 699 Forstmann, B. U., de Hollander, G., Van Maanen, L., Alkemade, A., & Keuken, M. C.
700 (2017). Towards a mechanistic understanding of the human subcortex. *Nature*
701 *Reviews Neuroscience*, 18, 57–65.
- 702 German, D. C., White, C. L., & Sparkman, D. R. (1987). Alzheimer's disease: Neurofibril-
703 lary tangles in nuclei that project to the cerebral cortex. *Neuroscience*, 21(2), 305–312.
704 doi:10.1016/0306-4522(87)90123-0
- 705 Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak,
706 R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult
707 human brains. *NeuroImage*, 14(1 I), 21–36. doi:10.1006/nimg.2001.0786
- 708 Gowland, P. A. & Bowtell, R. (2007). Theoretical optimization of multi-echo fMRI data
709 acquisition. *Physics in Medicine and Biology*, 52(7), 1801–1813. doi:10.1088/0031-
710 9155/52/7/003
- 711 Greenberg, D. L., Messer, D. F., Payne, M. E., MacFall, J. R., Provenzale, J. M., Steffens,
712 D. C., & Krishnan, R. R. (2008). Aging, gender, and the elderly adult brain: An
713 examination of analytical strategies. *Neurobiology of Aging*, 29(2), 290–302. doi:10.
714 1016/j.neurobiolaging.2006.09.016
- 715 Guan, X., Xu, X., & Zhang, M. (2017). Region-Specific Iron Measured by MRI as a
716 Biomarker for Parkinson's Disease. *Neuroscience Bulletin*, 33(5), 561–567. doi:10.
717 1007/s12264-017-0138-x
- 718 Haacke, E. M., Cheng, N. Y., House, M. J., Liu, Q., Neelavalli, J., Ogg, R. J., ... Obenaus,
719 A. (2005). Imaging iron stores in the brain using magnetic resonance imaging.
720 *Magnetic Resonance Imaging*, 23(1), 1–25. doi:10.1016/j.mri.2004.10.001
- 721 Hallgren, B. & Sourander, P. (1958). the Effect of Age on the Non-Haemin Iron in the
722 Human Brain. *Journal of Neurochemistry*, 3(1), 41–51. doi:10.1111/j.1471-4159.1958.
723 tb12607.x
- 724 Hametner, S., Endmayr, V., Deistung, A., Palmrich, P., Prihoda, M., Haimburger, E., ...
725 Grabner, G. (2018). The influence of brain iron and myelin on magnetic susceptibil-

ity and effective transverse relaxation - A biochemical and histological validation study. *NeuroImage*, 179(May), 117–133. doi:10.1016/j.neuroimage.2018.06.007

Herting, M. M., Johnson, C., Mills, K. L., Vijayakumar, N., Dennison, M., Liu, C., ... Tamnes, C. K. (2018). Development of subcortical volumes across adolescence in males and females: A multisample study of longitudinal changes. *NeuroImage*, 172(January), 194–205. doi:10.1016/j.neuroimage.2018.01.020

Hill, R. A., Li, A. M., & Grutzendler, J. (2018). Lifelong cortical myelin plasticity and age-related degeneration in the live mammalian brain. *Nature Neuroscience*, 21(5), 683–695. doi:10.1038/s41593-018-0120-6

Hirsch, E. C., Graybiel, A. M., & Agid, Y. (1988). Melanized dopaminergic neurons are differentially affected in Parkinson's disease. *Nature*, 334, 345–348.

Ho, T. C., Gutman, B., Pozzi, E., Grabe, H. J., Hosten, N., Wittfeld, K., ... Schmaal, L. (2020). Subcortical shape alterations in major depressive disorder: Findings from the ENIGMA major depressive disorder working group. *Human Brain Mapping*, (December 2019), 1–11. doi:10.1002/hbm.24988

Keuken, M. C., Bazin, P. L., Schäfer, A., Neumann, J., Turner, R., & Forstmann, B. U. (2013). Ultra-high 7T MRI of structural age-related changes of the subthalamic nucleus. *Journal of Neuroscience*, 33(11), 4896–4900. doi:10.1523/JNEUROSCI.3241-12.2013

Keuken, M. C., Bazin, P.-L., Backhouse, K., Beekhuizen, S., Himmer, L., Kandola, A., ... Forstmann, B. U. (2017). Effects of aging on T1, T2*, and QSM MRI values in the subcortex. *Brain Structure and Function*, 222(6), 2487–2505. doi:10.1007/s00429-016-1352-4

Keuken, M. C., Isaacs, B. R., Trampel, R., van der Zwaag, W., & Forstmann, B. U. (2018). Visualizing the Human Subcortex Using Ultra-high Field Magnetic Resonance Imaging. *Brain Topography*, 31(4), 513–545. doi:10.1007/s10548-018-0638-7

Keuken, M. C., Müller-Axt, C., Langner, R., Eickhoff, S. B., Forstmann, B. U., & Neumann, J. (2014). Brain networks of perceptual decision-making: an fMRI ALE meta-analysis. *Frontiers in Human Neuroscience*, 8. doi:10.3389/fnhum.2014.00445

Kitajima, M., Korogi, Y., Kakeda, S., Moriya, J., Ohnari, N., Sato, T., ... Yamashita, Y. (2008). Human subthalamic nucleus: Evaluation with high-resolution MR imaging at 3.0 T. *Neuroradiology*, 50(8), 675–681. doi:10.1007/s00234-008-0388-4

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*, 154(March), 59–80. doi:10.1016/j.neuroimage.2017.03.033

- 761 Langkammer, C., Bredies, K., Poser, B. A., Barth, M., Reishofer, G., Fan, A. P., ... Ropele, S.
762 (2015). Fast quantitative susceptibility mapping using 3D EPI and total generalized
763 variation. *NeuroImage*, 111, 622–630. doi:10.1016/j.neuroimage.2015.02.041
- 764 Langkammer, C., Schweser, F., Krebs, N., Deistung, A., Goessler, W., Scheurer, E., ...
765 Reichenbach, J. R. (2012). Quantitative susceptibility mapping (QSM) as a means to
766 measure brain iron? A post mortem validation study. *NeuroImage*, 62(3), 1593–1599.
767 doi:10.1016/j.neuroimage.2012.05.049
- 768 Lau, J. C., Xiao, Y., Haast, R. A., Gilmore, G., Uludağ, K., MacDougall, K. W., ... Khan,
769 A. R. (2020). Direct visualization and characterization of the human zona incerta
770 and surrounding structures. *Human Brain Mapping*, 41(16), 4500–4517. doi:10.1002/
771 hbm.25137
- 772 Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A.,
773 Weinberger, D. R., & Mattay, V. S. (2012). Normal age-related brain morphometric
774 changes: Nonuniformity across cortical thickness, surface area and gray matter
775 volume? *Neurobiology of Aging*, 33(3), 617.e1–617.e9. doi:10.1016/j.neurobiolaging.
776 2010.07.013
- 777 Li, W., van Tol, M. J., Li, M., Miao, W., Jiao, Y., Heinze, H. J., ... Walter, M. (2014). Regional
778 specificity of sex effects on subcortical volumes across the lifespan in healthy aging.
779 *Human Brain Mapping*, 35(1), 238–247. doi:10.1002/hbm.22168
- 780 Limousin, P., Pollak, P., Benazzouz, A., Hoffmann, D., Le Bas, J.-F., Perret, J., ... Brous-
781 solle, E. (1995). Effect on parkinsonian signs and symptoms of bilateral subthalamic
782 nucleus stimulation. *The Lancet*, 345(8942), 91–95. doi:10.1016/S0140-6736(95)90062-
783 4
- 784 Lozano, A. M., Lipsman, N., Bergman, H., Brown, P., Chabardes, S., Chang, J. W., ...
785 Krauss, J. K. (2019). Deep brain stimulation: current challenges and future direc-
786 tions. *Nature Reviews Neurology*, 15(3), 148–160. doi:10.1038/s41582-018-0128-2
- 787 Mancini, M., Karakuzu, A., Cohen-Adad, J., Cercignani, M., Nichols, T. E., & Stikov, N.
788 (2020). An interactive meta-analysis of MRI biomarkers of Myelin. *eLife*, 9, 1–23.
789 doi:10.7554/eLife.61523
- 790 Mangeat, G., Govindarajan, S. T., Mainiero, C., & Cohen-Adad, J. (2015). Multivariate
791 combination of magnetization transfer, T2* and B0 orientation to study the myelo-
792 architecture of the in vivo human cortex. *NeuroImage*, 119, 89–102. doi:10.1016/j.
793 neuroimage.2015.06.033
- 794 Marques, J. P., Khabipova, D., & Gruetter, R. (2017). Studying cyto and myeloarchitecture
795 of the human cortex at ultra-high field with quantitative imaging: R1, R2* and

796 magnetic susceptibility. *NeuroImage*, 147(December 2016), 152–163. doi:10.1016/j.
797 neuroimage.2016.12.009

798 Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., &
799 Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved
800 segmentation and T1-mapping at high field. *NeuroImage*, 49(2), 1271–1281. doi:10.
801 1016/j.neuroimage.2009.10.002

802 Martin, W. R., Wieler, M., & Gee, M. (2008). Midbrain iron content in early Parkinson
803 disease: A potential biomarker of disease status. *Neurology*, 70(16 PART 2), 1411–
804 1417. doi:10.1212/01.wnl.0000286384.31050.b5

805 Metere, R. & Möller, H. E. (2018). Quantifying the myelin and iron contents of the brain
806 in vivo using a linear model of relaxation. *Proceedings of the International Society for*
807 *Magnetic Resonance in Medicine*, 26.

808 Miletić, S., Bazin, P.-L., Weiskopf, N., van der Zwaag, W., Forstmann, B. U., & Trampel, R.
809 (2020). fMRI protocol optimization for simultaneously studying small subcortical
810 and cortical areas at 7 T. *NeuroImage*, 219(May). doi:10.1016/j.neuroimage.2020.
811 116992

812 Miletić, S., Boag, R. J., & Forstmann, B. U. (2020). Mutual benefits: Combining reinforce-
813 ment learning with sequential sampling models. *Neuropsychologia*, 136, 107261.
814 doi:10.1016/j.neuropsychologia.2019.107261

815 Mink, J. W. (1996). The Basal Ganglia: Focused selection and inhibition of competing
816 motor programs. *Progress in Neurobiology*, 50(4), 381–425. doi:10.1016 / S0301-
817 0082(96)00042-1

818 Mulder, M. J., Keuken, M. C., Bazin, P.-L., Alkemade, A., & Forstmann, B. U. (2019). Size
819 and shape matter: The impact of voxel geometry on the identification of small
820 nuclei. *PLOS ONE*, 14(4), e0215382. doi:10.1371/journal.pone.0215382

821 O’Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J., & Dolan, R. J. (2004).
822 Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning.
823 *Science*, 304(5669), 452–454. doi:10.1126/science.1094285

824 Pauli, W. M., Nili, A. N., & Tyszka, J. M. (2018). A high-resolution probabilistic in vivo
825 atlas of human subcortical brain nuclei. *Scientific Data*, 5(1), 180063. doi:10.1038/
826 sdata.2018.63

827 Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensi-
828 tivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired
829 inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, 55(6), 1227–1235.
830 doi:10.1002/mrm.20900

831 Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M. L., ...
832 Kiselev, V. G. (1999). Enhancement of BOLD-Contrast Sensitivity by Single-Shot
833 Multi-Echo Functional MR Imaging. *Magnetic resonance in medicine*, 42(1), 87–97.
834 Puckett, A. M., Bollmann, S., Poser, B. A., Palmer, J., Barth, M., & Cunningham, R. (2018).
835 Using multi-echo simultaneous multi-slice (SMS) EPI to improve functional MRI
836 of the subcortical nuclei of the basal ganglia at ultra-high field (7T). *NeuroImage*,
837 172(May 2017), 886–895. doi:10.1016/j.neuroimage.2017.12.005
838 Ramos, P., Santos, A., Pinto, N. R., Mendes, R., Magalhães, T., & Almeida, A. (2014). Iron
839 levels in the human brain: A post-mortem study of anatomical region differences
840 and age-related changes. *Journal of Trace Elements in Medicine and Biology*, 28(1),
841 13–17. doi:10.1016/j.jtemb.2013.08.001
842 Randall, L. O. (1938). Chemical Topography of the Brain. *Journal of Biological Chemistry*,
843 124(2), 481–488.
844 Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). *Applied Regression Analysis* (J. O.
845 Rawlings, S. G. Pantula, & D. A. Dickey, Eds.). Springer Texts in Statistics. New
846 York: Springer-Verlag. doi:10.1007/b98890
847 Raz, N. (2004). The Aging Brain Observed in Vivo. In *Cognitive neuroscience of ag-*
848 *ing* (Vol. 15, 1, pp. 19–57). Oxford University Press. doi:10.1093/acprof:oso/
849 9780195156744.003.0002
850 Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A.,
851 ... Acker, J. D. (2005). Regional brain changes in aging healthy adults: General
852 trends, individual differences and modifiers. *Cerebral Cortex*, 15(11), 1676–1689.
853 doi:10.1093/cercor/bhi044
854 Raz, N. & Rodrigue, K. M. (2006). Differential aging of the brain: Patterns, cognitive
855 correlates and modifiers. *Neuroscience and Biobehavioral Reviews*, 30(6), 730–748.
856 doi:10.1016/j.neubiorev.2006.07.001
857 Raznahan, A., Shaw, P. W., Lerch, J. P., Clasen, L. S., Greenstein, D., Berman, R., ... Giedd,
858 J. N. (2014). Longitudinal four-dimensional mapping of subcortical anatomy in
859 human development. *Proceedings of the National Academy of Sciences*, 111(4), 1592–
860 1597. doi:10.1073/pnas.1316911111
861 Rooney, W. D., Johnson, G., Li, X., Cohen, E. R., Kim, S.-G., Ugurbil, K., & Springer, C. S.
862 (2007). Magnetic field and tissue dependencies of human brain longitudinal 1H2O
863 relaxation in vivo. *Magnetic Resonance in Medicine*, 57(2), 308–318. doi:10.1002/mrm.
864 21122

- 865 Samanez-Larkin, G. R. & D'Esposito, M. (2008). Group comparisons: Imaging the aging
866 brain. *Social Cognitive and Affective Neuroscience*, 3(3), 290–297. doi:10.1093/scan/
867 nsn029
- 868 Schenck, J. F. & Zimmerman, E. A. (2004). High-field magnetic resonance imaging of
869 brain iron: Birth of a biomarker? *NMR in Biomedicine*, 17(7), 433–445. doi:10.1002/
870 nbm.922
- 871 Schmierer, K., Wheeler-Kingshott, C. A., Tozer, D. J., Boulby, P. A., Parkes, H. G., Yousry,
872 T. A., ... Miller, D. H. (2008). Quantitative magnetic resonance of postmortem
873 multiple sclerosis brain before and after fixation. *Magnetic Resonance in Medicine*,
874 59(2), 268–277. doi:10.1002/mrm.21487
- 875 Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and
876 reward. *Science*, 275(5306), 1593–1599. doi:10.1126/science.275.5306.1593
- 877 Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2),
878 461–464. doi:10.1214/aos/1176344136
- 879 Seabold, S. & Perktold, J. (2010). Statsmodels: econometric and statistical modeling with
880 Python. *9th Python in Science Conference*, (Scipy), 57–61.
- 881 Shen, S., Sandoval, J., Swiss, V. A., Li, J., Dupree, J., Franklin, R. J., & Casaccia-Bonnel, P.
882 (2008). Age-dependent epigenetic control of differentiation inhibitors is critical for
883 remyelination efficiency. *Nature Neuroscience*, 11(9), 1024–1034. doi:10.1038/nn.2172
- 884 Shepherd, G. M. (2013). Corticostriatal connectivity and its role in disease. *Nature Reviews*
885 *Neuroscience*, 14(4), 278–291. doi:10.1038/nrn3469
- 886 Shepherd, T. M., Flint, J. J., Thelwall, P. E., Stanis, G. J., Mareci, T. H., Yachnis, A. T.,
887 & Blackband, S. J. (2009). Postmortem interval alters the water relaxation and
888 diffusion properties of rat nervous tissue - Implications for MRI studies of human
889 autopsy samples. *NeuroImage*, 44(3), 820–826. doi:10.1016/j.neuroimage.2008.09.054
- 890 Stafford, J. L., Albert, M. S., Naeser, M. A., Sandor, T., & Garvey, A. J. (1988). Age-Related
891 Differences in Computed Tomographic Scan Measurements. *Archives of Neurology*,
892 45(4), 409–415. doi:10.1001/archneur.1988.00520280055016
- 893 Stüber, C., Morawski, M., Schäfer, A., Labadie, C., Wähnert, M., Leuze, C., ... Turner, R.
894 (2014). Myelin and iron concentration in the human brain: A quantitative study of
895 MRI contrast. *NeuroImage*, 93(P1), 95–106. doi:10.1016/j.neuroimage.2014.02.026
- 896 Tian, Y., Margulies, D. S., Breakspear, M., & Zalesky, A. (2020). Topographic organization
897 of the human subcortex unveiled with functional connectivity gradients. *Nature*
898 *Neuroscience*, 23(11), 1421–1432. doi:10.1038/s41593-020-00711-6

- 899 Tovi, M. & Ericsson, A. (1992). Measurements of T1 and T2 over time in formalin-
900 fixed human whole-brain specimens. *Acta Radiologica*, 33(5), 400–404. doi:10.1080/
901 02841859209172021
- 902 Trutti, A. C., Fontanesi, L., Mulder, M. J., Bazin, P.-L., Hommel, B., & Forstmann, B. U.
903 (2021). A probabilistic atlas of the human ventral tegmental area (VTA) based on 7
904 Tesla MRI data. *Brain Structure and Function*. doi:10.1007/s00429-021-02231-w
- 905 Turner, B. M. (2019). Toward a Common Representational Framework for Adaptation.
906 *Psychological Review*. doi:10.1037/rev0000148
- 907 Wagenmakers, E.-J. & Farrell, S. (2004). AIC model selection using Akaike weights.
908 *Psychonomic bulletin & review*, 11(1), 192–196. doi:10.3758/BF03206482
- 909 Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., ...
910 Fischl, B. (2005). Effects of age on volumes of cortex, white matter and subcortical
911 structures. *Neurobiology of Aging*, 26(9), 1261–1270. doi:10.1016/j.neurobiolaging.
912 2005.05.020
- 913 Wang, Y., Xu, Q., Luo, J., Hu, M., & Zuo, C. (2019). Effects of age and sex on subcortical
914 volumes. *Frontiers in Aging Neuroscience*, 11(SEP), 1–12. doi:10.3389/fnagi.2019.
915 00259
- 916 Ward, R. J., Zucca, F. A., Duyn, J. H., Crichton, R. R., & Zecca, L. (2014). The role of
917 iron in brain ageing and neurodegenerative disorders. *The Lancet Neurology*, 13(10),
918 1045–1060. doi:10.1016/S1474-4422(14)70117-6
- 919 Waskom, M. & development team, S. (2020). Seaborn. Zenodo. doi:10.5281/zenodo.
920 592845
- 921 Xu, J., Kobayashi, S., Yamaguchi, S., Iijima, K. I., Okada, K., & Yamashita, K. (2000).
922 Gender effects on age-related changes in brain structure. *American Journal of Neuro-*
923 *radiology*, 21(1), 112–118.
- 924 Ye, R., Rua, C., O'Callaghan, C., Jones, P. S., Hezemans, F. H., Kaalund, S. S., ... Rowe,
925 J. B. (2021). An in vivo probabilistic atlas of the human locus coeruleus at ultra-
926 high field. *NeuroImage*, 225(October 2020), 117487. doi:10.1016/j.neuroimage.2020.
927 117487
- 928 Zecca, L., Youdim, M. B., Riederer, P., Connor, J. R., & Crichton, R. R. (2004). Iron, brain
929 ageing and neurodegenerative disorders. *Nature Reviews Neuroscience*, 5(11), 863–
930 873. doi:10.1038/nrn1537

931 Supplementary Materials

932 CNR comparisons

933 Under the assumption of mono-exponential signal decay, the expected contrast-to-noise
934 ratio (CNR) of an $T2^*$ -weighted single echo, echo planar imaging protocol, per unit
935 change in $T2^*$, is given by (e.g. Poser, Versluis, Hoogduin, & Norris, 2006; Posse et al.,
936 1999):

$$CNR(TE, T2^*) = \frac{S_0}{\sigma_0} TE \exp \frac{-TE}{T2^*} \quad (3)$$

937 where S_0 and σ_0 are the signal and (temporal) variance, respectively, at echo time
938 $TE = 0$. In the following comparisons, we assume that S_0 and σ_0 are the same across
939 structures and protocols. This assumption is unlikely to be true for between-region CNR
940 comparisons, as S_0/σ_0 is typically lower in subcortical regions compared to cortical
941 regions in part due to the larger distance to the MRI receiver coils. We ignore any
942 potential between-region differences in the size of the $T2^*$ changes that result from
943 changes in the oxygenation levels. Under these assumptions, CNR ratios can be used
944 to compare the expected CNRs when using echo time A to study region n with the
945 expected CNR when using echo time B to study region m :

$$\frac{CNR(TE_A, T2_m^*)}{CNR(TE_B, T2_n^*)} = \frac{TE_A \exp \frac{-TE_A}{T2_m^*}}{TE_B \exp \frac{-TE_B}{T2_n^*}} \quad (4)$$

946 First, we can compare the expected CNR in the red nucleus with the CNR in the
947 amygdala, both in 19 years old participants, using $T2^*$ values obtained through our app
948 (<https://subcortex.eu/app>). Filling in $TE_A = TE_B = T2_{amg}^* = 0.041666... \text{ s}$ (i.e., in
949 both protocols, we use an echo time optimised for the amygdala at 19 years old), and
950 $T2_{rn}^* = 0.01826 \text{ s}$, we find $\frac{CNR(TE_A, T2_{rn}^*)}{CNR(TE_B, T2_{amg}^*)} = 0.277$, implying the CNR in the red nucleus
951 is approximately 72% lower than in the amygdala with this echo time. In practice, since
952 S_0/σ_0 is also likely to be lower in the red nucleus than in the amygdala, the CNR ratio
953 will be even lower.

954 Second, we can estimate the effect of age-related decreases in $T2^*$. The red nucleus
955 has an approximate $T2^* = 0.018 \text{ s}$ at 19 years old, and $T2^* = 0.0137 \text{ s}$ at 50 years old.
956 If an echo time of $TE_A = TE_B = 0.018 \text{ s}$ is used for a participant of 50 years old, then
957 $\frac{CNR(0.018, 0.0137)}{CNR(0.018, 0.018)} = 0.7168$, showing an approximately 28% loss in CNR compared to the
958 CNR that would be obtained with this echo time in a participant of 19 years old.

959 Third, we can compare the expected CNR in the red nucleus at 19 years old ($T2^* =$

0.018 s) when using the optimal echo time $TE_A = 0.018$ s with the CNR that would be obtained if an echo time is used that is optimal for the amygdala: $TE_B = 0.0042$ s. The ratio $\frac{CNR(0.018, 0.018)}{CNR(0.042, 0.018)} = 1.579$, showing that a substantial CNR gain can be expected from optimising the echo time to meet the specific requirements of that region.

Finally, we can compare the effect of adapting echo times to adjust for age-related changes in $T2^*$. Taking again the red nucleus as an example, the $T2^*$ decreases from 0.018 s to 0.0138 s between 19 and 50 years old. Suppose we analyze the red nucleus in a 50 year old participant using the corresponding optimal echo time (hence, $TE_A = T2^* = 0.0137$ s), then, compared with using an echo time optimal for young participants $TE_B = 0.018$, we find that $\frac{CNR(0.0137, 0.0137)}{CNR(0.018, 0.0137)} = 1.048$. Hence, adjusting for the age-related decrease in $T2^*$ leads to modest CNR gains.

In practice, changing the echo time may not always be possible due to hardware limitations (e.g., the slew rate of the MRI gradients limits the minimum echo time that can be achieved), and potentially requires undersampling of k-space (e.g., using GRAPPA, SENSE, or partial Fourier) or bandwidth changes. These additional changes will affect the protocols' S_0/σ_0 , complicating direct comparison between the expected performance of two candidate MRI protocols. Another option is to use multi echo protocols, in which data is acquired at multiple echo times, which can be optimised for multiple regions at the same time (Gowland & Bowtell, 2007; Kundu et al., 2017; Miletic, Boag, & Forstmann, 2020; Puckett et al., 2018). These additional factors should be taken into account when developing an MRI protocol in order to find the optimal trade-off between echo time settings and S_0/σ_0 .

Figure S1: Illustration of the interactive app which allows for visualizing all data and age-related change models. Top panel features a 3D mesh plot that includes all 17 subcortical structures under investigation. The age slider can be used to visualize the age-related changes in center of mass location (currently visible), or the median or interquartile range of iron, myelin, R1, R2*, and QSM values, color-coded on the structures. Bottom panel features scatterplots of the relation between age and twelve measures (median and interquartile range of myelin, iron, R1, R2*, QSM, and thickness) for all structures. Optional captions are included, and the data underlying the bottom panel can readily be downloaded as a csv file. The app can be accessed via <https://subcortex.eu/app>

Table S1: Iron estimates and corresponding qMRI values used for estimating linear model Equation 1. For all regions that were divided in subregions in one but not the other source (i.e. GP vs GPi/GPe; STR vs PUT/CAU; brain stem vs medulla oblongata; and cortical areas), we entered all subregions in the OLS model, using the more specific values where possible, and the global values otherwise. E.g., both GPe and GPi were in the model and shared iron values (Hallgren and Sourander report only GP), but different qMRI values. ^a From Hallgren and Sourander (1958); ^b cerebrospinal fluid estimates from Metere and Möller (2018); ^c We used the qMRI metrics of the internal capsule here; ^d Cortex, brain stem, cerebellum were parcellated using MGDM and CRUISE Bazin et al. (2014).

	Iron (ppm)	QSM (ppm)	R1 (Hz)	R2* (Hz)
Brainstem ^d	1.4 ^a	-0.000486	0.808	37.959
CAU	9.28 ^a	0.012768	0.66	48.64
Cerebellum ^d	3.35 ^a	$-7.80 * 10^{-5}$	0.505	34.069
Frontal Cortex ^d	2.92 ^a	$-4.00 * 10^{-6}$	0.537	32.722
Frontal White ^c	4.24 ^a	-0.020322	0.895	36.294
GPe	21.3 ^a	0.042671	0.848	85.107
GPi	21.3 ^a	0.047246	0.869	77.739
Motor Cortex ^d	5.03 ^a	$-4.00 * 10^{-6}$	0.537	32.722
Occipital Cortex ^d	4.55 ^a	$-4.00 * 10^{-6}$	0.537	32.722
PUT	13.32 ^a	0.012768	0.655	48.644
Parietal Cortex ^d	3.81 ^a	$-4.00 * 10^{-6}$	0.537	32.722
RN	19.48 ^a	0.063051	0.922	74.041
SN	18.46 ^a	0.058915	0.862	72.284
Sensory Cortex ^d	4.32 ^a	$-4.00 * 10^{-6}$	0.537	32.722
THA	4.76 ^a	0.002828	0.758	38.228
Temporal Cortex ^d	3.13 ^a	$-4.00 * 10^{-6}$	0.537	32.722
VENT	0.062 ^b	0.005553	0.341	6.259

Table S2: For all regions that were divided in subregions in one but not the other source (i.e. GP vs GPi/GPe; STR vs PUT/CAU; brain stem vs medulla oblongata; and cortical areas), we entered all subregions in the OLS model, using the more specific values where possible, and the global values otherwise. E.g., both CAU and PUT were in the model with separate myelin values, but identical qMRI values since the MASSP parcellation only reports STR. ^a From Hallgren and Sourander (1958); ^b cerebrospinal fluid estimates from Metere and Möller (2018); ^c We used the qMRI metrics of the internal capsule here; ^d Cortex, brain stem, cerebellum were parcellated using MGDM and CRUISE Bazin et al. (2014).

	Myelin (ppm)	QSM (ppm)	R1 (Hz)	R2* (Hz)
Brainstem ^d	15.36 ^a	-0.000486	0.808	37.959
CAU	6.21 ^a	0.012768	0.655	48.644
Frontal Cortex ^d	5.08 ^a	$-4.00 * 10^{-6}$	0.537	32.722
Frontal White ^c	16.26 ^a	-0.020322	0.895	36.294
GPe	10.404 ^a	0.044293	0.813	84.875
GPi	10.404 ^a	0.046384	0.838	75.216
Parietal Cortex ^d	5.42 ^a	$-4.00 * 10^{-6}$	0.537	32.722
PUT	5.611 ^a	0.012768	0.655	48.644
RN	13.442 ^a	0.059001	0.894	72.979
SN	7.404 ^a	0.05624	0.846	71.532
STN	14.423 ^a	0.058185	0.935	78.924
THA	11.4 ^a	0.002828	0.758	38.228
VENT	0.002 ^b	0.005553	0.232	6.259

Figure S2: Matrix of cumulative change in all dependent variables (columns) for individual structures (rows). Values indicate the summed absolute change between 19-75 years old. Negative values indicate decreases. Standard errors were obtained by bootstrapping with 10,000 iterations and are shown in parentheses. Location change in fornix was not included since the quality of parcellation of the fornix posterior horns increases with age, resulting in artificial shifts in the estimated center of mass with age. The ventricular system is assumed to have no iron or myelin concentration and is excluded from analysis.