HHS Public Access

Author manuscript

Neuroimage. Author manuscript; available in PMC 2017 March 01.

Published in final edited form as:

Neuroimage. 2016 March ; 128: 11–20. doi:10.1016/j.neuroimage.2015.12.045.

Accumulation of Iron in the Putamen Predicts its Shrinkage in Healthy Older Adults: A multi-occasion longitudinal study

Ana M. Daugherty[1,*] and **Naftali Raz**[1,2]

[1]Institute of Gerontology, Wayne State University, Detroit, MI

[2]Psychology Department, Wayne State University, Detroit, MI

Abstract

Accumulation of non-heme iron is believed to play a major role in neurodegeneration of the basal ganglia. In healthy aging, however, the temporal relationship between change in brain iron content and age-related volume loss is unclear. Here, we present the first long-term longitudinal multi-occasion investigation of changes in iron content and volume in the neostriatum in a sample of healthy middle-aged and older adults (N = 32; ages 49–83 years at baseline). Iron content, estimated via R2* relaxometry, increased in the putamen, but not the caudate nucleus. In the former, the rate of accumulation was coupled with change in volume. Moreover, greater baseline iron content predicted faster shrinkage and smaller volumes seven years later. Older age partially accounted for individual differences in neostriatal iron content and volume, but vascular risk did not. Thus, brain iron content may be a promising biomarker of impending decline in normal aging.

Keywords

aging; oxidative stress; R2*; striatum; susceptibility weighted imaging

1.0 Introduction

[1]The brain changes with age but the mechanisms of change remain obscure (Raz and Kennedy 2009). An influential hypothesis of brain aging postulates that age-related losses of brain parenchyma and reduction in functional capacity are driven by cumulative damage produced by build-up of reactive oxygen species (ROS) and ensuing oxidative stress (Harman 1956; Dröge and Schipper 2007; Sohal and Orr 2012) and chronic neuroinflammation (Finch et al. 1969; Finch and Crimmins 2004; Grammas 2011). ROS originate in organelles, such as the mitochondria and peroxisomes (Murphy 2009; Brown and Borutaite 2012), and are part of normal metabolism (Görlach et al. 2015). However, excessive accumulation of ROS upsets the normal equilibrium and accelerates the rate of oxidative stress that degrades mitochondrial membranes, impedes energy production in the

[*]Corresponding author: ana.daugherty@wayne.edu; 87 E. Ferry St., 226 Knapp Bldg., Detroit, MI 48202; telephone +01 313-664-2649; fax +01 313-664-2666.

mitochondria, promotes DNA mutations, and hastens apoptosis (Sohal and Orr 2012). Paradoxically, one of the major sources of intracellular ROS is iron, an essential participant in normal metabolic function, including synthesis of high-energy phosphate in the mitochondria (Halliwell 1992; Mills et al. 2010; Ward et al. 2014). By producing highly reactive ROS via Fenton reaction, non-heme iron exerts detrimental effects on the cell (Zecca et al. 2004; Mills et al. 2010; Hare et al. 2013). Because of its major role in abetting ROS-related cellular damage, brain iron that can be estimated by noninvasive neuroimaging makes a plausible proxy of the processes that otherwise are very difficult to assess *in vivo*.

Since recent advances in magnetic resonance imaging (MRI) methods for iron estimation, studies of lifetime differences in brain iron content have proliferated (see Haacke et al. 2005; Daugherty and Raz 2013, 2015 for reviews). The cumulative record thus far supports the proposition that brain iron accumulation may be a meaningful biomarker of impending structural and cognitive declines in aging and disease (Schenck and Zimmerman 2004; Walsh et al. 2013; Ward et al. 2014; Daugherty and Raz 2015). Nonetheless, the temporal relationship between iron accumulation and structural changes in the brain is unclear.

Postmortem studies show that subcortical regions vulnerable to age-related volume loss (e.g., the neostriatum) evidence greater iron content in older brains (Hallgren and Sourander 1958; Thomas et al. 1993; Aquino et al. 2009) and cross-sectional MRI investigations largely replicate these findings (Antonini et al. 1993; Bartzokis et al. 1994; Xu et al. 2008; Cherubini et al. 2009; Peran et al. 2009; Sullivan et al. 2009; Haacke et al. 2010; Pfefferbaum et al. 2010; Penke et al. 2012). The cumulative evidence of age-related differences in iron content has been quantified in a recent meta-analysis of MRI studies, which identified the greatest age-related differences in the caudate nucleus and putamen (Daugherty and Raz 2013). Cross-sectional studies, however, are not informative about the dynamics of continuous processes of change and individual variations in trajectories of aging (Raz and Lindenberger 2011). Thus longitudinal studies are necessary to determine the potential contribution of iron accumulation to typical brain aging, and yet, when it comes to age-related iron accumulation, such studies are particularly scarce.

A single longitudinal study of healthy adults found an increase in iron content in the caudate nucleus and putamen after two years (Daugherty et al. 2015). A study of neurodegenerative disease that followed a control group of younger and middle-aged adults over two years showed increase in iron content in the putamen and globus pallidus, but not the caudate nucleus (Walsh et al. 2014), whereas a small control group of middle-aged and older adults evidenced no change in striatal iron content (Ulla et al. 2013). The mixed evidence with regards to regional vulnerability notwithstanding, an increase in iron content in the basal ganglia appears to occur in normal aging and further longitudinal study is warranted.

Moreover, because all extant longitudinal *in vivo* studies of brain iron involved only two measurement occasions, the temporal order of regional iron accumulation and loss of volume in the brain could not be assessed (see Daugherty and Raz 2015 for a review). The variance partitioning approach in cross-sectional mediation analyses (e.g., Rodrigue et al. 2013) cannot reveal the temporal order of events and the relationship between them (Lindenberger et al. 2011). A sole longitudinal study of healthy adults showed that iron

accumulation in the neostriatum can explain its shrinkage (Daugherty et al. 2015), but as a two-occasion study it could not examine the lead-lag relations between the variables. Thus, multiple occasions of measurement are required to test the hypothesis of iron accumulation as a driver of shrinkage.

A plausible alternative hypothesis is that the age-related increase in iron content is not an independent phenomenon that precedes regional shrinkage, but instead is a relative shift in concentration due to shrinkage. Although we found no support for this hypothesis in our previous study (Daugherty et al. 2015), the two measurement occasions design limited testing of change-change associations.

Thus, the present study was designed to expand upon our previous longitudinal study (Daugherty et al. 2015) by testing these hypotheses in a new sample of middle-aged and older healthy adults who were assessed up to four times over seven years. Iron content was estimated via R2* relaxometry in the neostriatum, a technique that has demonstrated reliability and validity (Daugherty and Raz 2015; Daugherty et al. 2015). These measures of iron content were combined with regional volumes in models evaluated with a latent-variable longitudinal modeling technique—simple and parallel change latent growth curve analyses. This statistical approach produces error-free estimates of change in iron content and volume, individual differences therein, and allows testing the precedence of change in one factor predicting change in the other. We hypothesized that longitudinal increase in iron in the caudate nucleus and putamen would precede and predict shrinkage of both regions.

2.0 Materials and Methods

2.1 Participants

Middle-aged and older adults (N = 32; 58% female) were recruited from the Metro Detroit area as part of a long-term longitudinal study. Participants (age 49–83 years at baseline) were assessed two to four times over 7 years (average delay between baseline and the first follow-up = 15.69 months, *SD* = 1.28; between the first and second follow-up = 15.40 months, *SD* = 2.79; and between the second and third follow-ups = 58.05 months, *SD* = 5.28), see Table 1 for a demographic profile of the sample and Figure 1 for a graphic display of the assessment schedule. The participants were screened for neurological and cardiovascular pathology, thyroid disorder, endocrine disease, psychiatric disease, drug and alcohol abuse, and head injury. Participants reported right-hand dominance (Edinburgh Handedness Questionnaire; Oldfield, 1971) and were screened for vision and hearing problems at each assessment. For inclusion, participants scored less than 16 on the Center for Epidemiologic Study depression scale (CES-D; Radloff 1977) and at least 26 on the mini-mental state examination (MMSE; Folstein et al. 1975) at enrollment and each follow-up.

The sample used in this study consisted of cases with complete data at baseline and the first follow-up assessment (i.e., at least two assessments). In addition to the selected sample of N = 32, 23 persons were enrolled in the study at baseline and follow-up but were excluded from analyses. Seven cases were dropped because upon retrospective evaluation they were found to violate the health criteria set at enrollment. The remaining 16 cases had incomplete

MRI data at the first two assessments due to either incorrect acquisition or excessive artifacts.

Of the retained sample (N = 32), 13 persons (46% female) had missing longitudinal data (n = 2 at the second follow-up, n = 11 at the third follow-up). These participants did not differ from the 19 with complete four measurement occasions with respect to age (t = 0.20, p = 0.85), MMSE (t = 0.27, p = 0.79), CES-D (t = −0.42, p = 0.68), or years of education (t = −0.20, p = 0.84) at baseline. MMSE scores in the retained sample were stable over the course of study (mean latent change = −0.02, p = 0.66; individual differences in change = 0.02, p = 0.15) and remained high (see Table 1). Thus, the missing longitudinal data were treated as missing at random and were handled via full information maximum likelihood (FIML), a method that utilizes all available data to optimize estimation without imputation (Muthén et al. 1987; Larsen 2011).

In the course of the longitudinal study, 9 participants developed conditions that would have been exclusionary at baseline: 2 persons developed bladder cancer, 1 person a thyroid colloid cyst, 5 persons were diagnosed with a heart condition (e.g., murmur, arrhythmia), and 1 person reported possible binge drinking behaviors. In order to maintain the sample size, these 9 cases were kept in the primary analyses and later removed to confirm that development of health conditions alone did not account for the effects.

The inclusion criteria allowed for having a diagnosis of hypertension and taking antihypertensive and cholesterol lowering medications. By the last assessment, 12 persons were taking at least one such medication: 3 participants were on an ace inhibitor, 2 on a beta-blocker, 2 on an angiotensin II receptor blocker, 2 on a diuretic, 2 on a calcium antagonist, and 6 on statins. At each assessment occasion, systolic and diastolic blood pressure (in mmHg) was calculated as the average of two measurements taken via a mercury sphygmomanometer (BMS 12-S25) with a standard blood pressure cuff (Omron Professional) on the left arm with the participant seated in a quiet room. Although the sample was, on average, normotensive, in the course of study, as many as 11 participants had clinically diagnosed or observed hypertension (blood pressure measurement at least 140 mm Hg systolic or 90 mmHg diastolic; see Table 1). Hypertension was determined categorically at each assessment and composited into a frequency score—ranging 0 (normotensive at all assessments) to 1 (hypertensive at all assessments)—that was tested as a cardiovascular risk factor to potentially explain age-related variability in iron accumulation and shrinkage (see Daugherty et al. 2015; Daugherty and Raz 2015).

2.2 R2* Relaxometry

A four-echo susceptibility weighted imaging (SWI) sequence was acquired on a 1.5T Siemens Sonata scanner with the following parameters: echo time (TE) = 10–40 ms with an inter-echo interval of 10 ms; 1 × 1 × 2 mm^3 voxel; repetition time (TR) = 100 ms; flip angle (FA) = 30°; bandwidth = 170 Hz/pixel; field of view (FOV) = 256 × 256. Images were acquired, viewed and processed in the axial plane. Noise was excluded from all measurements by applying a threshold that selectively isolated intensity values corresponding to image inhomogeneity.

All images were pre-processed in Signal Processing in NMR (SPIN), an in-house written software package available at http://mrinnovations.com/index.php?site=spin (last accessed 10/07/2015). T2* maps were interpolated to common space and estimated by a maximum-likelihood fit function for exponential decay dependent upon TE with respect to individual scan image inhomogeneity. The scanner underwent standard maintenance over the duration of the study and T2* image inhomogeneity was consistent across the four occasions in the 19 persons who had complete longitudinal data: error coefficient of variation, $F(3,54) = 1.20$, $p = 0.32$. Boundaries of the caudate nucleus and putamen were manually traced with a stylus on a 21-inch digitizing tablet (Wacom Cintiq) by a single rater (A.M.D); see Figure 2 for examples of tracings, and Daugherty et al. (2015) for a description of tracing procedures. The T2* images from all four assessments were processed simultaneously in pairs randomly assigned across time points. Rater reliability was confirmed for both regions with an intraclass correlation coefficient (ICC(3); Shrout and Fleiss 1979) exceeding 0.90 in a sub-sample of 10 cases measured twice with a delay of 2 weeks. Average T2* (ms) was estimated across all slices traced of a given region and converted to R2* [s^{-1}] = (1/T2*) × 1000, for which higher values correspond to greater iron content. This procedure has been demonstrated to provide reliable measures of R2* relaxometry that correlate highly with reported post-mortem measures of non-heme iron content (see Daugherty et al. 2015; Daugherty and Raz 2015).

2.3 Volumetry

In addition to the multi-echo SWI sequence for T2*, a T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence was acquired for volumetry with the following parameters: TE = 3.93 ms, interpolated voxel 0.5 × 0.5 × 0.5 mm^3, TR = 800 ms, inversion time = 420 ms, GRAPPA = 2, FOV = 192 × 192 mm^2, bandwidth = 130 Hz/pixel, acquisition matrix = 256 × 256, FA = 20°. With Analyze software (v10 and later; Biomedical Imaging Resource, Mayo Clinic College of Medicine), the operators viewed the caudate nucleus and putamen in the coronal plane on images manually realigned perpendicular to the anterior-posterior commissural axis (see Figure 2). In addition to regional measures, intracranial volume (ICV) was obtained through manual segmentation and used in adjustment of the regional volumes for head size and possible effects of changes in scanner-related variables over time via analysis of covariance (Jack et al. 1989; Raz et al. 2004). Seven expert raters manually demarcated regional boundaries for volumetry using a stylus on a 21-inch digitizing tablet (Wacom Cintiq). All attained among them high inter-rater reliability measured via an ICC(2) formula that assumes random raters (Shrout and Fleiss 1979). Across rater pairs and regions ICC(2) of at least 0.90 was attained. See our previous publications (Raz et al. 2005; Daugherty et al. 2015) for a description of tracing procedures. To guard against rater drift and to ensure raters were blind to time in the study, all four occasions of MRI data were traced in pairs, randomly assigned across measurement occasions.

2.4 Latent Growth Curve Modeling

Hypotheses were tested via latent growth curve models (LGCM; Preacher et al. 2008) in a longitudinal structural equation modeling framework implemented in Mplus software (v7, Muthén and Muthén). A LGCM estimates the rate of latent longitudinal change, individual

differences therein, and the independent, explicit estimation of error that makes latent estimates, by definition, error-free. Prior to constructing latent models, left and right hemisphere measures were averaged and univariate outliers were winsorized. To ensure measurement invariance over time, the observed measures were normed to baseline values and centered at 0, and measurement variances within a latent construct were estimated and constrained to be equal across time. Within and across latent constructs, measures were constrained to not correlate, except for models of putamen iron content, which allowed the association between baseline and the last follow-up. All models were estimated with full information maximum likelihood (FIML) treating data missing at random—a method that exploits all available data without imputation to produce estimates of change and individual differences (Muthén et al. 1987; Larsen 2011). The LGCMs were constructed as fixed linear slopes with coefficient loadings reflecting the average delay between assessments. Simple LGCMs estimated the latent longitudinal change and individual differences for a single regional measure (i.e., R2* or volume) with the slope centered at baseline (fixed factor loading of the measurement at 0) and interpreted as the rate of change from baseline.

2.4.1 Parallel Change LGCM—Parallel change LGCMs tested the specific lead-lag hypothesis that change in iron content precedes and predicts change in volume. In these models, the lead-lag contrast was imposed upon the slope functions by centering the slope of volume at the last follow-up. In this model, the intercept of iron content slope is equal to the baseline value, the corresponding slope is interpreted as the rate of change in iron content *from* baseline, the slope of volume as the rate of change *towards* the last follow-up and the intercept of this slope corresponds to the volume at the last assessment taken approximately seven years from baseline. This model construction better accommodates testing temporal precedence in the lead-lag relationship with contemporaneous measurements while still estimating the same rates of change as in the simple LGCMs.

2.4.2 Reverse models: Alternative hypothesis testing—To test the alternative hypothesis that shrinkage accounts for iron accumulation, reverse parallel change LGCMs were constructed to have change in volume (slope centered at baseline) precede and predict change in iron content (slope centered at the last follow-up). Poor model fit as compared to other models tested or non-significant effects provides more evidence for the direction of effects.

2.4.3 Model evaluation—Model fit was determined by a set of indices (Raykov and Marcoulides 2006): normal theory weighted chi-square statistic (a non-significant value indicates good fit); root mean square error of approximation (RMSEA ≤ 0.10 indicates good fit); comparative fit index (CFI ≥ 0.90 indicates excellent fit); standardized root mean residual (SRMR ≤ 0.10 supports good fit), or for models with categorical variables, weighted root mean residual (WRMR ≤ 0.80 supports good fit). To optimize model fit, final models constrained non-significant effects that were not necessary control covariates (e.g., non-significant correlations with average delay between assessments were maintained in the models). Unless otherwise stated, all reported effects are standardized. To avoid spurious results due to a small sample size, models were bootstrapped with bias correction (5000 iterations of the whole sample; Hayes and Scharkow 2013) to estimate 95% confidence

intervals (BS 95% CI) of unstandardized effects. Finally, the limited sample size required separate models by region and a Bonferroni correction (α') was used to adjust for multiple comparisons.

3.0 Results

3.1 Longitudinal Measurement Invariance

Because we used single indicators of iron content and volume at each assessment occasion, we assessed the consistency of measures over time outside of the LGCM by testing the homogeneity of variance and Pearson correlations between the values at baseline and subsequent occasions. Measures of average putamen volume had similar variance at each time point (all $F \leq 1.03$, all $p \geq 0.59$) and high correlations between occasions ($r = 0.85$–0.96), as did average caudate nucleus volume (variance: all $F \leq 1.29$, all $p \geq 0.63$; $r = 0.89$–0.97). Measures of average R2* were less consistent over time. In the putamen, variance in average R2* was similar between baseline and the second occasion ($F = 0.55$, $p = 0.10$), but was larger at the third ($F = 0.28$, $p = 0.001$) and fourth occasions ($F = 0.24$, $p < 0.001$), and longitudinal measures were moderately correlated ($r = 0.63$–0.75). In the caudate nucleus, variance in average R2* was larger at the third occasion as compared to baseline ($F = 0.21$, $p < 0.001$) but was similar at the other occasions (both $F \leq 1.58$, $p \geq 0.22$), and the measures correlated moderately over time ($r = 0.66$–0.81). Thus, to ensure longitudinal measurement invariance in the LGCMs, indices were normed to baseline, centered at 0 and measurement variances within each construct were estimated and constrained to be equal.

3.2 Seven-Year Longitudinal Change in Regional Iron Content and Volume

The simple LGCMs of change in R2* in each region had excellent fit: $\chi^2 \leq 15.46$, $p \geq 0.05$, RMSEA ≤ 0.16, CFI ≥ 0.91, SRMR ≤ 0.16. Only the putamen evidenced significant increase in iron content over seven years (see Figure 3) and participants varied in the rate of accumulation (see Table 2). In the caudate nucleus, there was no mean change, but variance in change was significant (see Figure 4 for individual change plots). Throughout the approximate seven years of study, the rank order of regional iron content estimates remained the same—the putamen had greater iron content than the caudate nucleus at baseline ($t(29) = -5.17$, $p < 0.001$) and the last follow-up ($t(19) = -6.03$, $p < 0.001$).

The simple LGCMs of change in volume in each region also evidenced excellent fit: $\chi^2 \leq 14.64$, $p \geq 0.07$, RMSEA ≤ 0.16, CFI ≥ 0.96, SRMR ≤ 0.11. The putamen and caudate nucleus shrank at a similar rate (see Figure 3). As with measures of longitudinal iron accumulation, the rate of shrinkage evidenced significant individual variability (see Figure 4 and Table 2). To explain the individual differences in the rate of change, we added baseline age, sex, and frequency of hypertension across occasions to the models.

3.3 Individual Differences at Baseline and in Change

Baseline age, converted to a z-score (centered at the sample mean), was entered as a correlate of baseline measures and rate of change in iron content and volume of each region. These models had good fit: $\chi^2 \leq 18.56$, $p \geq 0.04$, RMSEA ≤ 0.18, CFI ≥ 0.91, SRMR ≤ 0.13. In the putamen, older age at baseline was associated with smaller volumes (-0.58, $p < 0.001$,

$\alpha' = 0.03$; BS 95% CI: −0.81/−0.32) and greater iron content (0.39, p = 0.02, $\alpha' = 0.03$; BS 95% CI: 0.06/0.65). The same effects on caudate volume (−0.33, p = 0.04, $\alpha' = 0.03$; BS 95% CI: −0.57/−0.09) and iron content (0.25, p = 0.10; BS 95% CI: −0.06/0.53) showed trends in the same direction but failed to reach Bonferroni-corrected level of significance. Older adults experienced a slower rate of shrinkage of the putamen (−0.38, p = 0.02, $\alpha' = 0.03$; BS 95% CI: −0.08/−0.01), but not of the caudate nucleus (0.07, p = 0.82; BS 95% CI: −0.02/0.03). Advanced age was related to faster accumulation of caudate iron (0.97, p < 0.001, $\alpha' = 0.03$; BS 95% CI: 0.05/0.12), but unrelated to change in putamen iron content (0.76, p = 0.11; BS 95% CI: 0.01/0.10). Thus, baseline age was included as a covariate in all models.

Sex differences were evaluated by adding a categorical covariate to the models of each regional measure. The models of measures in the caudate nucleus and putamen iron had excellent fit: $\chi^2 \leq 14.71$, p ≥ 0.33, RMSEA ≤ 0.06, CFI ≥ 0.93, WRMR ≤ 0.56; but the model of effects on putamen volume had poorer fit: $\chi^2 = 16.64$, p = 0.22, RMSEA = 0.09, CFI = 0.80, WRMR = 0.38. Sex differences were identified only in baseline iron content of the caudate—men had less iron than women did (−0.45, p = 0.01, $\alpha' = 0.03$; BS 95% CI: −0.71/−0.17). There were no sex differences in change in caudate iron content (0.05, p = 0.24) or in volume (−0.33 and 0.03, p ≥ 0.17, baseline and change, respectively), nor in any measure within the putamen (all −0.83–0.02, p ≥ 0.45). Due to the limited effect, and its detrimental influence on model fit for the putamen, sex was removed from additional model testing.

Hypertension frequency across seven years (i.e., incidence of diagnosed or observed hypertension across the four assessments) was tested as a possible age-related cardiovascular risk factor. These models had poor fit: $\chi^2 \leq 27.93$, p ≥ 0.01, RMSEA ≤ 0.20, CFI ≥ 0.91, SRMR ≤ 0.13. Hypertension did not account for differences in putamen iron content (baseline = −0.12 and slope = −0.47, both p ≥ 0.10) or volume (baseline = 0.08 and slope = −0.22, both p ≥ 0.40), or in caudate iron content (baseline = −0.36 and slope = −0.10, both p ≥ 0.19) or volume (baseline = 0.02 and slope = −0.06, both p ≥ 0.86), and was removed from further model testing.

3.4 Iron Content at Baseline Predicts Seven-year Shrinkage and Volume of the Putamen

Parallel process LGCMs for the caudate nucleus and putamen tested the hypothesis that in each region, the accumulation of iron precedes and predicts regional shrinkage—the slope for iron was centered at baseline, whereas the slope of volume was centered at the last follow-up. For the putamen (see Figure 6A), the model had acceptable fit: $\chi^2 = 62.88$, p = 0.01, RMSEA = 0.13, CFI = 0.92, SRMR = 0.14. Faster iron accumulation was directly associated with faster shrinkage (−0.52, p < 0.01, $\alpha' = 0.03$; BS 95% CI: −2.55/−0.39). As indicated by the significant indirect effect (−0.36, p = 0.02, $\alpha' = 0.03$; BS 95% CI: −0.14/−0.02), individuals with greater iron content at baseline evidenced faster shrinkage (Figure 5A). Moreover, greater iron content at baseline predicted smaller volumes seven years later (−0.62, p < 0.001, $\alpha' = 0.03$; BS 95% CI: −1.52/−0.49; Figure 5B).

The parallel process model testing the same hypothesis for the caudate nucleus had poor fit: $\chi^2 = 68.77$, p < 0.001, RMSEA = 0.19, CFI = 0.85, SRMR = 0.19. The effects in the caudate

nucleus were in a similar direction as that in the putamen, but failed to reach significance (all $p \geq 0.45$). Based on poor model fit and lack of significance, the estimation of effects in this region was restructured. The better fitting models treated measures of iron content and volume separately (indices reported respectively): $\chi^2 = 12.74$ and 14.50, both $p \geq 0.31$, RMSEA = 0.07 and 0.06, CFI = 0.98 and 0.99, SRMR = 0.10 and 0.06. The slopes of these effects were each centered at the last follow-up to examine individual differences at that assessment occasion (see Figure 6 B1 and B2). Older age at baseline was associated with greater iron content (0.73, $p < 0.00$; $\alpha' = 0.03$; BS 95% CI: 0.57/1.02) and a nominally significant trend with smaller volumes (−0.33, $p = 0.047$; $\alpha' = 0.03$; BS 95% CI: −0.54/−0.08) at the last follow-up.

3.5 Reverse Model: Baseline Volume Does Not Predict Change in Iron Content

The alternative hypothesis that relative increase in iron concentration is a consequence of shrinkage was tested in a reverse model in which change in volume (slope centered at baseline) predicted change in iron (slope centered at the last follow-up). The reverse model for the putamen had incrementally worse fit than the model with change in iron content predicting change in volume: $\chi^2 = 67.02$, $p < 0.01$, RMSEA = 0.15, CFI = 0.90, SRMR = 0.20. The direct association between change in volume and change in iron content was not significant (−0.05, $p = 0.67$; BS 95% CI:−0.31/0.48) and there was no evidence of baseline volume predicting the rate of iron accumulation (indirect effect = 0.00, $p = 0.97$; BS 95% CI: −0.01/0.02). Moreover, larger volume at baseline was associated with *lesser* iron content after seven years (−0.30, $p < 0.01$, $\alpha' = 0.03$; BS 95% CI: −0.70/−0.11)—evidence contradicting the alternative hypothesis of a relative increase in iron concentration as a consequence of shrinkage. The reverse model of effects in the caudate nucleus also provided no evidence for shrinkage preceding and predicting individual differences in iron content (all $p \geq 0.07$) and had poorer fit than the final models in that region: $\chi^2 = 75.07$, $p < 0.001$, RMSEA = 0.14, CFI = 0.87, SRMR = 0.14.

The final models of effects in the putamen and caudate nucleus were reassessed excluding the nine participants who had developed health conditions over the course of study that would have been exclusionary at baseline enrollment. In the smaller sample (n = 23), the models of the interaction between putamen iron content and volume, and of age-related differences in caudate iron content had acceptable fit (indices reported respectively): $\chi^2 = 57.46$ and 13.07, $p = 0.04$ and 0.29, RMSEA = 0.14 and 0.09, CFI = 0.92 and 0.96, SRMR = 0.15 and 0.11; and all effects remained significant (all $p < 0.03$). The nominally significant baseline age-related differences in seven-year caudate volume became a non-significant trend ($p = 0.08$) and this model had poor fit: $\chi^2 = 32.25$, $p < 0.01$, RMSEA= 0.25, CFI = 0.88, SRMR = 0.19.

4.0 Discussion

Long ago, ROS-related damage promoted by the accumulation of non-heme iron was proposed as a mechanism of brain aging (Harman 1956). The cumulative record of extant MRI studies has reinforced that hypothesis and lent support to the notion of brain iron as a biomarker of impending decline (Schenck and Zimmerman 2004; Walsh et al. 2013; Ward et al. 2014; Daugherty and Raz 2015). Here, we tested this proposition for the first time with

multiple longitudinal assessments over an extended period in a sample of healthy middle-aged and older adults. We observed that over approximately seven years of study, iron content increased in the putamen, and greater iron content at baseline predicted faster shrinkage and smaller volumes at the last follow-up.

This is a partial replication and extension of our previous findings from a two-occasion longitudinal study that showed similar effects after two years in a lifespan sample of healthy adults (Daugherty et al. 2015). In that study, iron content increased in both neostriatal regions and accounted for shrinkage. The combined evidence from these two longitudinal studies points to iron accumulation as a typical aspect of normal aging that contributes to characteristic loss of striatal volume.

Whereas extant cross-sectional studies of adult age differences demonstrate a moderate negative association between iron content and volumes (e.g., Cherubini et al. 2009; Rodrigue et al. 2013), only longitudinal studies with several measurements can test the temporal precedence of iron accumulation to shrinkage. The lead-lag parallel change LGCMs demonstrated that iron accumulation preceded shrinkage and there was no evidence supporting the reverse order. The cognitive consequences of iron accumulation and its associated shrinkage are unclear, but previous studies show greater iron content is associated with lesser repeated testing gains in verbal working memory (Daugherty et al. 2015), poorer memory performance (Bartzokis et al. 2011; Rodrigue et al. 2013; Ghadery et al. 2015), lower general cognitive aptitude (Penke et al. 2012), mental slowing (Pujol et al. 1992; Sullivan et al. 2009), and poorer cognitive and motor control (Adamo et al. 2014). Unfortunately, without cognitive variables in this study, we could not evaluate the temporal order of iron accumulation, shrinkage and cognitive change.

An important discrepancy between this and the previous longitudinal study of striatal volume and iron accumulation should be noted. Whereas we replicated effects in the putamen, the stability of iron content in the caudate nucleus contradicts our previous report conducted on an adult lifespan sample. In this sample, limited to middle-aged and older adults, caudate iron content showed no linear change in the span of seven years. Two studies on middle-aged and older healthy controls also found no significant change in caudate iron content after two years (Ulla et al. 2013; Walsh et al. 2014). One of those studies found increase in iron content of the putamen (Ulla et al. 2013) whereas the other did not (Walsh et al. 2014). Although caudate iron content was stable in adults who were age 49 and older at baseline, individuals differed in the rate of change. Thus, it is plausible that iron accumulation across the lifespan may follow a nonlinear course, with slowing after middle age. Extant cross-sectional studies of neostriatal iron content across nearly the entire lifespan suggest this non-linear pattern (e.g., Hallgren and Sourander 1958; Thomas et al. 1993; Aquino et al. 2009; Li et al. 2014). Indeed, some of these studies showed attenuated age differences in caudate iron content in adults 50 years and older, in contrast to a near linear age association with putamen iron content throughout the adult lifespan (Aquino et al. 2009; Li et al. 2014). The shape of the putative non-linear relationship is unclear but it may fit a threshold model. Accumulation of ROS induced by iron that escapes sequestration may have to reach a certain critical level before expressing itself in structural damage and functional decline. Even without significant change in caudate iron content after the fifth decade,

elevations may still explain individual differences in the rate of brain and cognitive declines in later life (Daugherty et al. 2015). Testing the threshold hypothesis of non-linear change will require a study with several more measurements and a sufficiently large sample (Hertzog et al. 2006; Serroyen et al. 2009).

Given the individual differences observed in both neostriatal regions, it is plausible that multiple health-related factors may modify the magnitude of change. Older age at baseline only partially accounted for these differences, with older participants showing *lesser* declines. Although attrition in this study was not substantial, pre-baseline selection according to strict health criteria might have introduced a bias favoring older participants. Greater variability observed among the older participants suggests that additional factors may modify the trajectories of brain aging. Identifying such modifiers in a selected sample of healthy adults is challenging, as the effects are expected to be small. Uncomplicated arterial hypertension has been shown before to explain individual differences in subcortical iron content (Raz et al. 2007; Rodrigue et al. 2011; also see Berry et al. 2001), but we did not find the same in this sample. Other factors, such as obesity (Blasco et al. 2014), elevation in metabolic risk (Daugherty et al. 2015), and chronic inflammation (Zecca et al. 2004; see Wessling-Resnick 2010; Urrutia et al. 2014 for reviews) may also exacerbate iron accumulation and its consequences. In addition to age-related risk, genetic polymorphisms associated with less efficient iron homeostasis (Lehmann et al. 2006; Bartzokis et al. 2011; Nandar and Connor, 2011) may further explain individual differences in iron accumulation. The sample size of this study could not accommodate tests of these effects that warrant further investigation.

Several limitations qualify the findings of this study. First, iron content was estimated via R2* relaxometry, which is a reliable and valid, but not specific index. Other sources of susceptibility can affect R2* (Haacke et al. 2005) and confound estimates of iron content. One example of an age-related confounding factor is calcification of subcortical regions (Naderi et al. 1993) that can only be distinguished from iron with phase-based MRI methods (Haacke et al. 2005, 2015). Another threat to validity is the presence of myelin and myelinated fibers that contribute to susceptibility (Fukunaga et al. 2010; Liu et al. 2011; Langkammer et al. 2012; Lodygensky et al. 2012) and increase R2* (Haacke et al. 2005; Glasser and Van Essen 2011). However, the influence of myelin in the neostriatum is likely negligible (see Daugherty and Raz, 2013 for a review). Another potential confound stems from the inability of R2* to differentiate between heme and non-heme forms of iron. Heme iron binds oxygen in circulating blood and is found where blood flows or accumulates (Janaway et al. 2014; see Loitfelder et al. 2012; Yates et al. 2014 for reviews). Because the incidence of vascular pathology in this sample was low and we took care to exclude vascular objects during processing, we believe the influence of heme iron on R2* in the regions reported here was negligible. It must be noted that the validity of R2* as a proxy for the oxidative stress-driven decline is limited by its dependence on sources of non-heme iron besides that accumulated outside of binding complexes. Ferritin-bound concentrations are the largest source of magnetic susceptibility observed on MRI. Although concentrations of ferritin-bound non-heme iron do not directly contribute to oxidative stress, they are proportional to soluble non-heme iron that does (Morita et al. 1981; Salgado et al. 2010).

Thus, change in iron-related MRI susceptibility reflects an overall shift in iron homeostasis that will produce insidious oxidative stress and is a plausible explanation of the effects observed here (see Daugherty and Raz 2015 for a review). Development of methods that are differentially sensitive to free iron that is not sequestered by ferritin would add specificity to the in vivo assessment of potential markers of oxidative stress agents.

A second limitation is the small sample size, which restricts the number of variables that can be reasonably included in the models and the number of parameters that the models can contain. Due to this limitation, we averaged hemispheric measures prior to modeling and only tested three covariates to explain individual differences. Lack of statistical power can also be a factor in latent change and change-change associations failing to reach significance (Hertzog et al. 2006). This is especially true when measurement is less consistent over time (Hertzog et al. 2006) as was seen here in R2*, and removal of this bias in the models, though necessary for valid estimates of change, dilutes effects. Longitudinal studies that do not explicitly test for metric invariance, or that use methods incapable of extracting it from estimates of change, produce ambiguous results (see Horn and McArdle 1992). Thus, the minimizing effect from the constraints we imposed on the models was a necessary compromise. Measurement error due to age-related factors that can affect R2* and volume (e.g., image inhomogeneity from CSF in enlarged ventricles) may account for some of the age-related variability in longitudinal measures, and the extraction of error from latent estimates may differentially diminish effects across the lifespan, particularly in the oldest persons. We could better account for this possibility with additional covariates, but the necessary restriction on the number of model parameters coupled with the small sample size, limits the feasibility to test a large number of covariates with acceptable statistical power. Further, the limited statistical power could not support tests of nonlinear effects with only four measurements. Therefore, our conclusions are limited to linear change and we can only speculate about the source of individual differences therein and its possible cognitive consequences.

Third, although we made accommodations for the small sample size, attrition (34% by the last assessment) needs to be taken into account. The drop-outs were similar to participants who remained in several demographic variables, and the FIML estimation of data as missing at random minimized the bias. Yet, given the relatively small scale of this study, we cannot completely dismiss the possibility that selective attrition affected the results reported here and we hope to address it further in other longitudinal studies with larger samples. Finally, we were able to test the temporal relationship between changes in iron content and volume, but without experimental manipulation one cannot definitively demonstrate causal relations amongst these variables. Experimental studies in animal models suggest that iron-related oxidative stress causes degeneration (e.g., Zhang et al. 2009) and our models of iron accumulation preceding shrinkage replicated the observed data well, but our conclusions about a causal mechanism are limited (see Winship and Morgan 1999; Rubin 2005; Pearl 2014a, 2014b).

4.1 Conclusions

In middle-aged and older adults, iron accumulation in the putamen precedes shrinkage, and individuals with greater baseline iron content in the putamen experienced faster volume reduction and greater loss. Regional brain iron content may be a promising biomarker of impending structural decline in the aging brain.

Acknowledgments

This work was supported by National Institute on Aging at the National Institutes of Health grant R37 AG011230 to N.R. and Blue Cross Blue Shield of Michigan Foundation student awards program grant 1893.SAP to A.M.D. We thank Drs. Cheryl Dahle, Andrew Bender, Peng Yuan, Yiqin Yang, Karen Rodrigue and Kristen Kennedy for their assistance in data collection, and Dr. E. Mark Haacke for advice on the imaging protocol.

Abbreviations

BS 95% CI	bootstrapped 95% confidence intervals
CES-D	Center for Epidemiological Study depression scale
CFI	comparative fit index
FIML	full information maximum likelihood
ICV	intracranial volume
MMSE	mini-mental state exam
RMSEA	root mean square error of approximation
ROS	reactive oxygen species
SRMR	square root mean residual
SWI	susceptibility-weighted imaging
WRMR	weighted root mean residual

References

Adamo DE, Daugherty AM, Raz N. Brain iron content and grasp force-matching ability in older women. Brain Imag Behav. 2014; 8(4):579–87.

Antonini A, Leenders KL, Meier D, Oertel MD, Boesiger P, Anliker M. T_2 relaxation tiem in patients with Parkinson's disease. Neurol. 1993; 43:697–700.

Aquino D, Bizzi A, Grisoli M, Garavaglia B, Bruzzone MG, Nardocci N, Savoiardo M, Chiapparini L. Age-related iron deposition in the basal ganglia: quantitative analysis in healthy subjects. Radiol. 2009; 252(1):165–172.

Bartzokis G, Lu P, Tingus K, Peters DG, Amar CP, Tishler TA, Finn JP, Willablanca P, Altshuler LL, Mintz J, Neely E, Connor JR. Gender and iron genes may modify associations between brain iron and memory in healthy aging. Neuropsychopharmacology. 2011; 36:1375–1384. [PubMed: 21389980]

Bartzokis G, Mintz J, Sultzer D, Marx P, Herzberg JS, Phelan CK, Marder SR. In vivo MR evaluation of age-related increases in brain iron. AJNR. 1994; 15(6):1129–1138. [PubMed: 8073983]

Berry C, Brosnan MJ, Fennel J, Hamilton CA, Dominiczak AF. Oxidative stress and vascular damage in hypertension. Current opinion in Nephrology and Hypertension. 2001; 10(2):247–255. [PubMed: 11224701]

Blasco G, Puig J, Daunis-I-Estadella J, Molina XL, Xifra G, Fernández-Aranda F, Pedraza S, Ricart W, Portero-Otín M, Fernández-Real J. Brain iron overload, insulin resistance and cognitive performance in obese subjects: A preliminary MRI case-control study. Diabetes Care. 2014; 37(11): 3076–83. [PubMed: 25125507]

Brown GC, Borutaite V. There is no evidence that mitochondria are the main source of reactive oxygen species in mammalian cells. Mitochondrion. 2012; 12(1):1–4. [PubMed: 21303703]

Cherubini A, Péran P, Caltagirone C, Sabatini U, Spalletta G. Aging of subcortical nuclei: Microstructural, mineralization and atrophy modifications measured in vivo using MRI. Neuroimage. 2009; 48:29–36. [PubMed: 19540925]

Daugherty A, Raz N. Age-related differences in iron content of subcortical nuclei observed in vivo: A meta-analysis. NeuroImage. 2013; 70:113–121. [PubMed: 23277110]

Daugherty AM, Raz N. Appraising the role of iron in brain aging and cognition: promises and limitations of MRI. Neuropsychol Rev. 2015; 25(3):272–87. [PubMed: 26248580]

Daugherty AM, Haacke EM, Raz N. Striatal iron content predicts its shrinkage and changes in working memory after two years in healthy adults. J Neurosci. 2015; 35(17):6731–6743. [PubMed: 25926451]

Dröge W, Schipper HM. Oxidative stress and aberrant signaling in aging and cognitive decline. Aging Cell. 2007; 6:361–370. [PubMed: 17517043]

Finch CE, Crimmins EM. Inflammatory exposure and historical changes in human lifespans. Science. 2004; 305:1736–1739. [PubMed: 15375259]

Finch CE, Foster JR, Mirsky AE. Ageing and the regulation of cell activities during exposure to cold. J Gen Physiol. 1969; 54:690–712. [PubMed: 4391050]

Fukunaga M, Li TQ, van Gelderen P, de Zwart JA, Shmueli K, Yao B, Lee J, Maric D, Aronova MA, Zhang G, Leapman RD, Schenck JF, Merkle H, Duyn JH. Layer-specific variation of iron content in cerebral cortex as a source of MRI contrast. Proc Natl Acad Sci U S A. 2010; 107:3834–3839. [PubMed: 20133720]

Ghadery C, Pirpamer L, Hofer E, Langkammer C, Petrovic K, Loitfelder M, Schwingenschuh P, Seiler S, Duering M, Jouvent E, Schmidt H, Fazekas F, Mangin JF, Chabriat H, Dichgans M, Ropele S, Schmidt R. R2* mapping for brain iron: associations with cognition n normal aging. Neurobiol Aging. 2015; 36:925–932. [PubMed: 25443291]

Glasser MF, Van Essen DC. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. J Neurosci. 2011; 31(32):11597–11616. [PubMed: 21832190]

Görlach A, Dimova EY, Petry A, Martínez-Ruiz A, Hernansanz-Agustín P, Rolo AP, Palmeira CM, Kietzmann T. Reactive oxygen species, nutrition, hypoxia and diseases: Problems solved? Redox Biol. 2015; 6:372–385. [PubMed: 26339717]

Grammas P. Neurovascular dysfunction, inflammation and endothelial activation: implications for the pathogenesis of Alzheimer's disease. J Neuroinflamm. 2011; 8:26.

Haacke EM, Cheng NYC, House MJ, Liu Q, Neelaavalli J, Ogg RJ, Khan A, Ayaz M, Kirsch W, Obenaus A. Imaging iron stores in the brain using magnetic resonance imaging. Mag Res Imag. 2005; 23(1):1–25.

Haacke EM, Liu S, Busch S, Zheng W, Wu D, Ye Y. Quantitative susceptibility mapping: current status and future directions. Mag Res Imag. 2015; 33:1–25.

Haacke EM, Miao Y, Liu M, Habib CA, Katkuri Y, Liu T, Yang Z, Lang Z, Hu J, Wu J. Correlation of putative iron content as represented by changes in R2* and phase with age in deep gray matter of healthy adults. J Magn Reson Imaging. 2010; 32:561–576. [PubMed: 20815053]

Hallgren B, Sourander P. The effect of age on the non-haemin iron in the human brain. J Neurochem. 1958; 3:41–51. [PubMed: 13611557]

Halliwell, B. Iron and damage to biomolecules. In: Lauffer, editor. Iron and Human Disease. Roca Baton (FL): CRC Press; 1992. p. 209-236.

Hare D, Ayton S, Bush A, Lei P. A delicate balance: Iron metabolism and diseases of the brain. Front Aging Neurosci. 2013; 5:34. [PubMed: 23874300]

Harman D. Aging: A theory based on free radical and radiation chemistry. J Gerontol. 1956; 11(3): 298–300. [PubMed: 13332224]

Hayes AF, Scharkow M. The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? Psych Sci. 2013; 24:1918–1927.

Hertzog C, Lindenberger U, Ghisletta P, von Oertzen T. On the power of multivariate latent growth curve models to detecting correlated change. Psychol Method. 2006; 11(3):244–252.

Horn JL, McArdle JJ. A practical and theoretical approach to measurement invariance in aging research. Experimental Ageing Research. 1992; 18:117–144.

Jack CR Jr, Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. Radiol. 1989; 172:549–554.

Janaway BM, Simpson JE, Hoggard N, Highley JR, Forster G, Drew D, Gebril OH, Matthews FE, Bryane C, Wharton SB, Ince PG. MRC Cognitive FunctionAgeing Neuropathology Study. Brain haemosiderin in older people: pathological evidence for an ischemic origin of magnetic resonance imaging (MRI) microbleeds. Neuropathol Appl Neurobiol. 2014; 40(3):258–69. [PubMed: 23678850]

Langkammer C, Schweser F, Krebs N, Deistung A, Goessler W, Scheurer E, Sommer K, Reishofer G, Yen K, Fazekas F, Ropele S, Reichenbach JR. Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. NeuroImage. 2012; 62:1593–1599. [PubMed: 22634862]

Larsen R. Missing data imputation versus full information maximum likelihood with second-level dependencies. Structural Equation Modeling: A Multidisciplinary Journal. 2011; 18(4):649–662.

Lehmann DJ, Worwood M, Ellis R, Wimhurst VLJ, Merryweather-Clarke AT, Warden DR, Smith AD, Robson KJH. Iron genes, iron load and risk of Alzheimer's disease. J Med Genet. 2006; 43:e52. [PubMed: 17047092]

Li W, Yu B, Batrachenko A, Bancroft-Wu V, Morey RA, Shashi V, Langkammer C, De Bellis MD, Ropele S, Song AW, Liu C. Differential developmental trajectories of magnetic susceptibility in human brain gray and white matter over the lifespan. Hum Brain Mapp. 2014; 35(6):2698–713. [PubMed: 24038837]

Liu C, Li W, Johnson A, Wu B. High-field (9.4 T) MRI of brain dysmyelination by quantitative mapping of magnetic susceptibility. NeuroImage. 2011; 56:930–938. [PubMed: 21320606]

Lindenberger U, von Oertzen T, Ghisletta P, Hertzog C. Cross-sectional age variance extraction: What's change got to do with it? Psychol Aging. 2011; 26(1):34–47. [PubMed: 21417539]

Loitfelder M, Seiler S, Schwingenschuh P, Schmidt R. Cerebral microbleeds: a review. Panimerva Med. 2012; 54(3):149–60.

Maxwell SE, Cole DA. Bias in cross-sectional analyses of longitudinal mediation. Psychol Methods. 2007; 12(1):23–44. [PubMed: 17402810]

Mills E, Dong X, Wang F, Xu H. Mechanisms of brain iron transport: Insight into neurodegeneration and CNS disorders. Future Med Chem. 2010; 2(1):51–72. [PubMed: 20161623]

Morita R, Yoshii M, Nakajima K, Kohsaka T, Miki M, Torizuka K. Clinical evaluation of serum ferritin to iron ratio in malignant diseases. European J of Nuc Med. 1981; 6(7):331–336.

Murphy MP. How mitochondria produce reactive oxygen species. Biochem Journal. 2009; 417(1):1–13. [PubMed: 19061483]

Muthén B, Kaplan D, Hollis M. On structural equation modeling with data that not missing completely at random. Psychometrika. 1987; 52:431–462.

Naderi S, Colakoglu Z, Lüleci G. Calcification of basal ganglia associated with pontine calcification in four cases: a radiologic and genetic study. Clin Neurol Neurosurg. 1993; 95:155–57. [PubMed: 8344016]

Nandar W, Connor JR. HFE gene variants affect iron in the brain. J Nutrition. 2011; 141:729S–739S. [PubMed: 21346098]

Oldfield RC. The assessment and analysis of handedness. Neuropsychologica. 1971; 9:97–113.

Pearl J. Interpretation and identification of causal mediation. Psychol Methods. 2014a; 19:459–81. [PubMed: 24885338]

Pearl J. Reply to commentary by Imai, Keele, Tingley, and Yamamoto, concerning causal mediation analysis. Psychol Methods. 2014b; 19:488–92. [PubMed: 25486117]

Penke L, Hernandéz MCV, Maniega SM, Gow AJ, Murray C, Starr JM, Bastin ME, Deary IJ, Wardlaw JM. Brain iron deposits are associated with general cognitive ability and cognitive aging. Neurobiol Aging. 2012; 33:510–517. [PubMed: 20542597]

Peran P, Cherubini A, Luccichenti G, Hagberg G, Démont J-F, Rascol O, Celsis P, Caltagirone C, Spalletta G, Sabatini U. Volume and iron content in the basal ganglia and thalamus. Hum Brain Mapp. 2009; 30:2667–2675. [PubMed: 19172651]

Pfefferbaum A, Adalsteinsson E, Rohlfing T, Sullivan EV. Diffusion tensor imaging of deep gray matter brain structures: Effects of age and iron concentration. Neurobiol Aging. 2010; 31(3):482–500. [PubMed: 18513834]

Preacher, KJ.; Wichman, AL.; MacCallum, RC.; Briggs, NE. Latent Growth Curve Modeling. Thousand Oaks (CA): SAGE Publications, Inc; 2008.

Pujol J, Junque C, Vendrell P, Grau JM, Marti-Vilalta JL, Olivé C, Gili J. Biological significance of iron-related magnetic resonance imaging changes in the brain. Arch Neurol. 1992; 49(7):711–717. [PubMed: 1497497]

Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. J Struct Biol. 1977; 153:42–54.

Raykov, T.; Marcoulides, GA. A first course in structural equation modeling. 2. Mahwah (NJ): Lawrence Erlbaum; 2006.

Raz, N.; Kennedy, KM. A systems approach to the aging brain: Neuroanatomic changes, their modifiers, and cognitive correlates. In: Jagust, W.; D'Esposito, M., editors. Imaging the Aging Brain. New York (NY): Oxford University Press; 2009. p. 43-70.

Raz N, Lindenberger U. News of cognitive cure for age-related brain shrinkage is premature: a comment on Burgmans et al., (2009). Neuropsychol. 2011; 24(2):255–257.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. Regional brain changes in aging healthy adults: General trends, individual differences, and modifiers. Cer Cort. 2005; 15:1676–1689.

Raz N, Rodrigue KM, Haacke EM. Brain aging and its modifiers: Insights from in vivo neuromophometry and susceptibility weighted imaging. Annals NY Acad Sci. 2007; 1097:84–93.

Raz N, Rodrigue K, Head D, Kennedy K, Acker J. Differential aging of the medial temporal lobe: A study of a five-year change. Neurol. 2004; 62:433–439.

Rodrigue KM, Daugherty AM, Haacke EM, Raz N. The role of hippocampal iron content and hippocampal volume in age-related differences in memory. Cer Cor. 2013; 23(7):1533–41.

Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. J Am Stat Assoc. 2005; 100:322–31.

Salgado JC, Olivera-Nappa A, Gerdtzen ZP, Tapia V, Theil EC, Conca C, Nuñez MT. Mathematical modeling of the dynamic storage of iron in ferritin. BMC Sys Bio. 2010; 4:147.

Schenck JF, Zimmerman EA. High-field magnetic resonance imaging of brain iron: birth of a biomarker? NMR Biomed. 2004; 17:433–445. [PubMed: 15523705]

Serroyen J, Molenberghs G, Verbeke G, Davidian M. Non-linear models for longitudinal data. Am Stat. 2009; 63(4):378–388. [PubMed: 20160890]

Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing raters reliability. Psychol Bull. 1979; 86:420–428. [PubMed: 18839484]

Sohal RS, Orr WC. The redox stress hypothesis of aging. Free Rad Biol Med. 2012; 52(3):539–555. [PubMed: 22080087]

Sullivan EV, Adalsteinsson E, Rohlfing T, Pfefferbaum A. Relevance of iron deposition in deep gray matter brain structures of cognitive and motor performance in healthy elderly men and women: exploratory findings. Brain Imag Behav. 2009; 3:167–175.

Thomas LO, Boyko OB, Anthony DC, Burger PC. MR detection of brain iron. Am J Neurorad. 1993; 14(5):1043–1048.

Ulla M, Bonny JM, Ouchchane L, Rieu I, Claise B, Durif F. Is R2* a new MRI biomarker for the progression of Parkinson's disease? A longitudinal follow-up. PLoS One. 2013; 8(3):e57904. [PubMed: 23469252]

Urrutia PJ, Mena NP, Núñez MT. The interplay between iron accumulation, mitochondrial dysfunction, and inflammation during the execution step of neurodegenerative disorders. Front Pharmacol. 2014; 5:article 38.

Walsh AJ, Belvins G, Lebel RM, Seres P, Emery DJ, Wilman AH. Longitudinal MR imaging of iron in multiple sclerosis: An imaging marker of disease. Radiol. 2014; 270(1):186–196.

Ward RJ, Zucca FA, Duyn JH, Crichton RR, Zecca L. The role of iron in brain ageing and neurodegenerative disorders. Lancet Neurol. 2014; 13:1045–60. [PubMed: 25231526]

Wessling-Resnick M. Iron homeostasis and the inflammatory response. Ann Rev Nutr. 2010; 30:105–122. [PubMed: 20420524]

Winship C, Morgan SL. The estimation of causal effects from observational data. Ann Rev Sociol. 1999; 26:659–706.

Xu X, Wang Q, Zhang M. Age, gender, and hemispheric differences in iron deposition in the human brain : An *in vivo* MRI study. NeuroImage. 2008; 40:35–42. [PubMed: 18180169]

Yates PA, Villemagne VL, Ellis KA, Desmond PM, Masters CL, Rowe CC. Cerebral microbleeds: a review of clinical, genetic, and neuroimaging associations. Front Neurosci. 2014; 4:205.

Zecca L, Youdim MBH, Riederer P, Connor JR, Crichton RR. Iron, brain ageing and neurodegenerative disorders. Nat Rev. 2004; 5:863–873.

Zhang S, Wang J, Song N, Xie J, Hiang H. Up-regulation of divalent metal transporter 1 is involved in 1-methyl-4-phenylpyridinium (MPP+)-induced apoptosis in MES23.5 cells. Neurobiol Aging. 2009; 30:1466–76. [PubMed: 18191877]

Highlights

- Iron-related oxidative stress is a proposed mechanism of brain aging
- Neostriatal iron content and volume were assessed multiple times over seven years
- Middle-aged and older adults showed iron accumulation in putamen but not caudate
- Putamen iron accumulation preceded and predicted shrinkage
- Greater baseline iron predicted faster shrinkage and greater loss after seven years

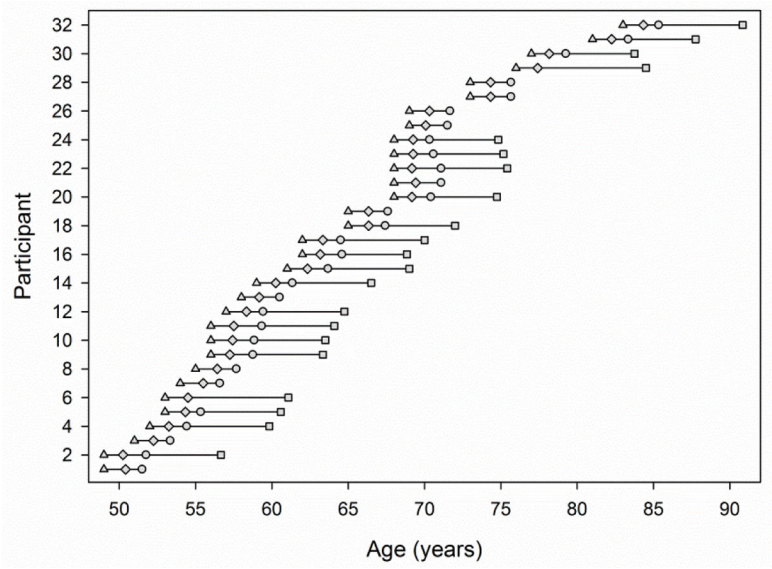

Figure 1.

Distribution of ages-at-measurement and intervals between measurement occasions for the 32 participants. The symbols represent each measurement occasion: triangle = 1st occasion; diamond = 2nd occasion; circle = 3rd occasion; square = 4th occasion. The mean interval duration between baseline and first follow up measurement was 15.69 months (*SD* = 1.28). The mean interval between the second and third visit was 15.40 months (*SD* = 2.79), and between the third and the fourth occasion 58.05 months (*SD* = 5.28).

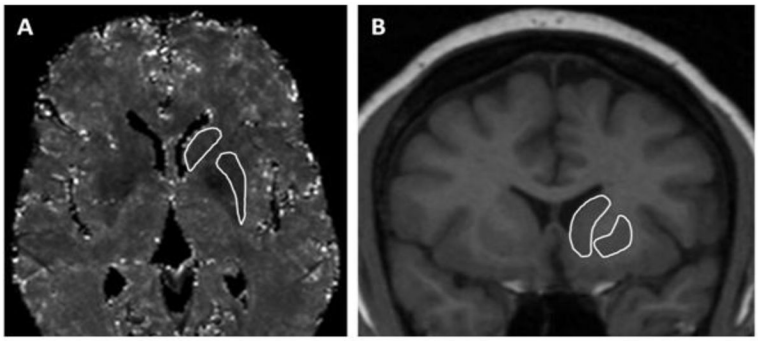

Figure 2.

Example manual segmentation of the caudate nucleus (Cd) and putamen (Pt) on images used for iron content estimation and volumetry, both acquired on the same 1.5T scanner and coil, at the same session, without participant re-positioning. **(A)** T2* maps (axial, $1 \times 1 \times 2$ mm^3 voxel) were calculated from four echoes with respect to individual image inhomogeneity, additional noise was removed with filtering prior to R2* measurement (= (1/T2*) × 1000). On T2*-weighted images, darker (lower intensity) regions have greater estimated iron content. **(B)** Volumes were measured from T1-MPRAGE images (coronal, $0.5 \times 0.5 \times 0.5$ mm^3 voxel) and were adjusted by intracranial volume via analysis of covariance at each assessment to correct for possible change in scanner-related variables over time, as well as sex differences in head size. See Methods section for image processing and data collection details.

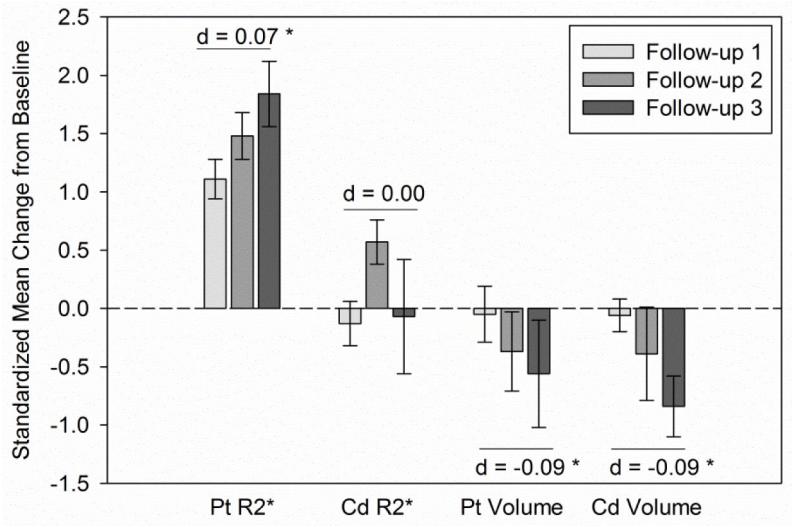

Figure 3.

Standardized mean latent change in neostriatal iron content (R2*) and volume from baseline. The dotted line demarks standardized baseline mean values, positive change from 0 indicates longitudinal gain, and negative change, longitudinal decline. Error bars represent standard error of the means. *d* is a standardized effect size of the latent rate of linear change across the four occasions (refer to Table 2); * indicates significant latent change after correction for multiple comparisons ($p < 0.01$; $\alpha' = 0.03$). On average, Follow-up 1 occurred 15 months after baseline; Follow-up 2, 30 months after baseline; and Follow-up 3, 90 months after baseline. Pt—putamen; Cd—caudate nucleus.

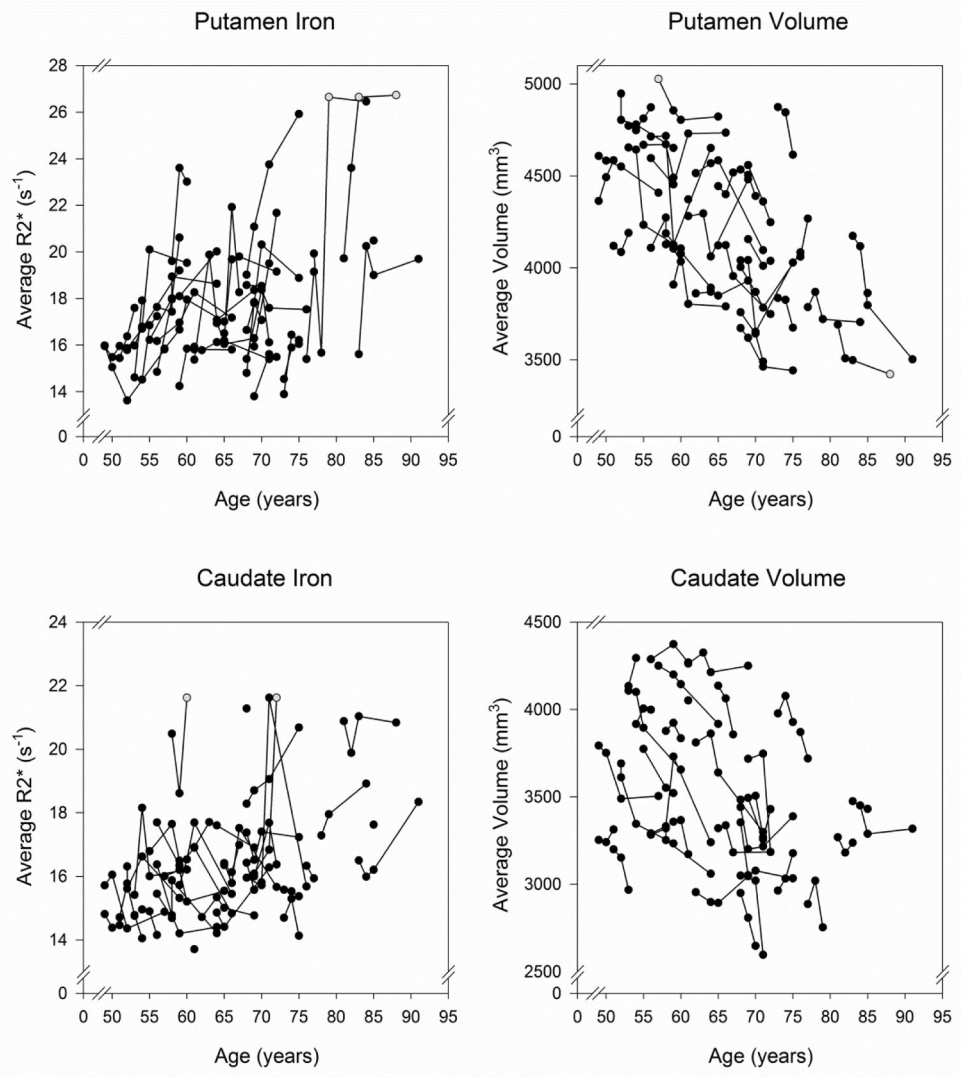

Figure 4.

Plots of individual change trajectories in iron content (R2*) and volume in the neostriatum measured four times across 7 years. Larger values of R2* correspond to greater iron content. Gray circles represent univariate outliers that were winsorized prior to analysis. The regression lines were fit with 95% confidence intervals (broken lines) and prediction intervals (dotted lines).

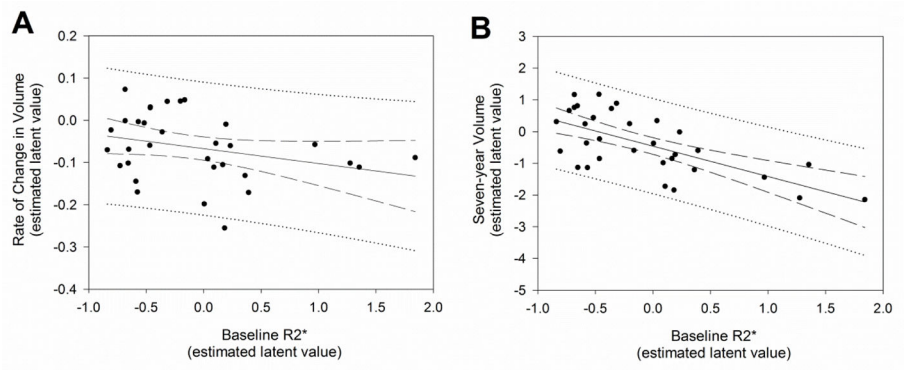

Figure 5.

Latent effects in the putamen estimated in a parallel latent growth curve model (see text for details). **A.** Greater baseline iron content in the putamen predicted faster shrinkage (−0.36, $p = 0.02$). Metrics are latent estimates derived from the model for each individual: greater values of baseline R2* correspond to greater iron content and increasing negative rate of change in volume indicates faster shrinkage. **B**. As an independent effect, greater baseline iron content predicted smaller volume at the last follow-up, approximately seven years after baseline (−0.62, $p < 0.001$). Smaller volumes are indicated by increasing negative latent estimated values.

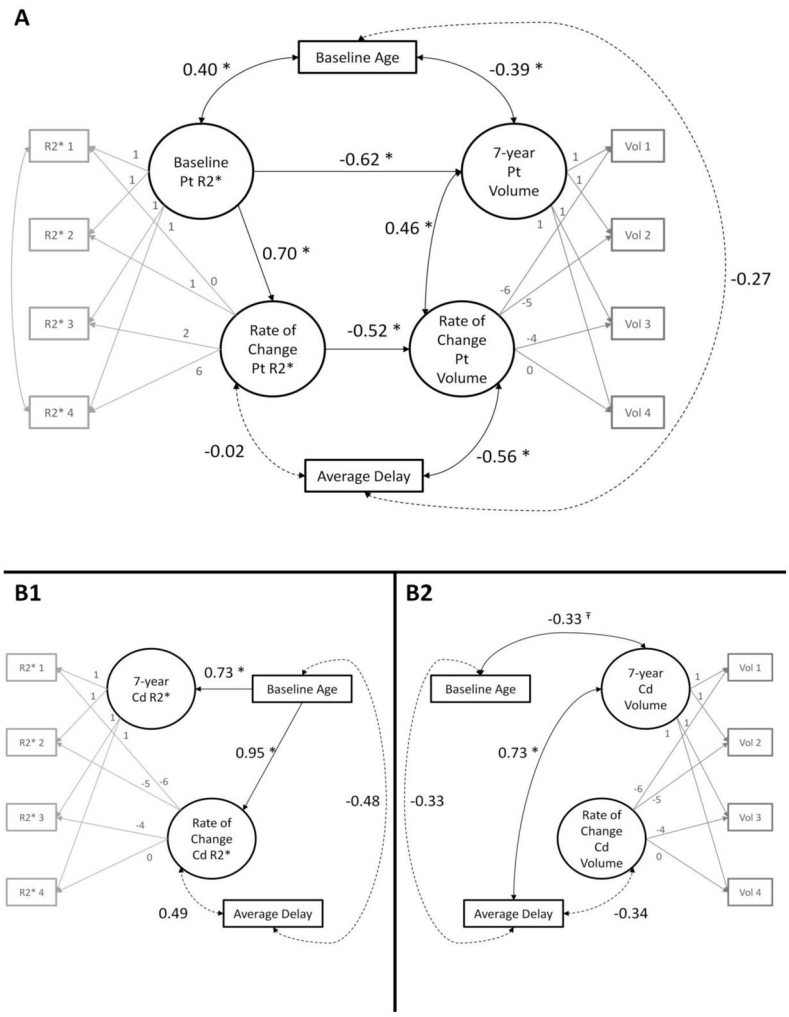

Figure 6.

Final latent growth curve models for (**A**) the putamen and (**B1-2**) the caudate nucleus. Gray lines and boxes summarize the measurement model (see methods section for details) with fixed linear slopes centered at the time point corresponding to the 0 loading. Curved, double-headed arrows indicate an estimated correlation whereas a straight arrow is a directional regression. Standardized effects are reported; * indicates significant effects after correction for multiple comparisons ($p < 0.03$), ∓ indicates effects of nominal significance ($p < 0.05$) that did not survive correction, and broken lines are non-significant effects maintained in the model for covariate control. **A.** In the putamen, a faster rate of iron accumulation was directly related to faster shrinkage and individuals with greater baseline iron content evidenced faster shrinkage (indirect effect: −0.36, $p = 0.02$, $\alpha' = 0.03$; BS 95% CI: −0.14/−0.02). Further, greater baseline iron content accounted for smaller volumes at the last follow-up. **B1–2.** In the caudate nucleus, after initial hypothesis testing, separate models for (B1) iron content and (B2) volume had the best fit (see results section for more detail). Older age at baseline was associated with faster iron accumulation, and greater iron content and smaller volumes after seven years. Pt—putamen; Cd—caudate nucleus; Vol—volume.

Table 1

Demographic profile of the sample measured four times.

	Baseline	Follow-up 1	Follow-up 2	Follow-up 3
N	32	32	30	21
Age (years)	62.94 ± 9.38	64.34 ± 9.33	65.37 ± 9.09	70.85 ± 9.91
Education (years)	16.28 ± 2.37	16.84 ± 2.67	16.90 ± 2.83	16.86 ± 2.69
MMSE	28.69 ± 1.23	28.44 ± 1.16	28.73 ± 1.02	28.52 ± 1.33
CES-D	4.16 ± 4.22	3.84 ± 4.14	3.53 ± 3.19	3.29 ± 3.90
Hypertension Freq.	11	10	11	10
Systolic (mmHg)	130.76 ± 12.67	128.16 ± 11.17	128.65 ± 12.50	127.05 ± 10.63
Diastolic (mmHg)	80.01 ± 6.42	77.64 ± 7.27	76.33 ± 6.21	77.12 ± 6.62

Note: Sample averages and standard deviations are reported. MMSE—mini-mental state exam (cut-off > 25); CES-D—center for epidemiologic study-depression scale (cut-off < 16); Hypertension was determined by clinical diagnosis or observed blood pressure > 140 mmHg systolic or 90 mmHg diastolic.

Table 2

Latent volume and iron estimates: Baseline variance, mean change and individual differences in change.

Region	Measure	Variance at Baseline	Mean Rate of Change	Variance in Change	d
Putamen	R2*	0.56 * (0.28/0.95)	0.05 * (0.01/0.09)	0.01 * (0.005/0.03)	0.07
	Volume	0.81 * (0.61/1.08)	−0.08 *(−0.12/−0.05)	0.01 * (0.004/0.01)	−0.09
Caudate	R2*	0.69 * (0.31/1.19)	−0.002 (−0.04/0.05)	0.01 (0.00/0.02)	0.00
	Volume	0.98 * (0.74/1.30)	−0.09 *(−0.12/−0.06)	0.003* (0.001/0.01)	−0.09

Note: Unstandardized coefficients are reported for data normed to baseline. 95% confidence intervals are bootstrapped with bias-correction and reported in parentheses. *d* is a standardized effect size estimation of mean change = Mean Rate of Change/√(Variance at baseline).

* $p < 0.01$; Bonferonni correction for multiple comparisons, $\alpha' = 0.03$.