



# HHS Public Access

Author manuscript

*Neuroimage*. Author manuscript; available in PMC 2023 September 28.

Published in final edited form as:

*Neuroimage*. 2022 June ; 253: 119063. doi:10.1016/j.neuroimage.2022.119063.

## Higher-order multi-shell diffusion measures complement tensor metrics and volume in gray matter when predicting age and cognition

Hamsanandini Radhakrishnan<sup>1</sup>, Ilana J Bennett<sup>2</sup>, Craig EL Stark<sup>\*,1,3</sup>

<sup>1</sup>Mathematical, Computational and Systems Biology, University of California, Irvine, California, USA - 92697.

<sup>2</sup>Department of Psychology, University of California Riverside, Riverside, California, USA.

<sup>3</sup>Department of Neurobiology and Behavior, University of California, Irvine, California, USA - 92697.

### Abstract

Recent advances in diffusion-weighted imaging have enabled us to probe the microstructure of even gray matter non-invasively. However, these advanced multi-shell protocols are often not included in large-scale studies as they significantly increase scan time. In this study, we investigated whether one set of multi-shell diffusion metrics commonly used in gray matter (as derived from Neurite Orientation Dispersion and Density Imaging, NODDI) provide enough additional information over typical tensor and volume metrics to justify the increased acquisition time, using the cognitive aging framework in the human hippocampus as a testbed. We first demonstrated that NODDI metrics are robust and reliable by replicating previous findings from our lab in a larger population of 79 younger ( $20.41 \pm 1.89$  years, 46 females) and 75 older ( $73.56 \pm 6.26$  years, 45 females) adults, showing that these metrics in the hippocampal subfields are sensitive to age and memory performance. We then asked how these subfield specific hippocampal NODDI metrics compared with standard tensor metrics and volume in predicting age and memory ability. We discovered that both NODDI and tensor measures separately predicted age and cognition in comparable capacities. However, integrating these modalities together considerably increased the predictive power of our logistic models, indicating that NODDI and tensor measures may be capturing independent microstructural information. We use these findings to encourage neuroimaging data collection consortiums to include a multi-shell diffusion sequence in their protocols since existing NODDI measures (and potential future multi-shell measures) may be able to capture microstructural variance that is missed by traditional approaches, even in studies exclusively examining gray matter.

---

\*Corresponding Author: cestark@uci.edu; Postal Address: 1400 Biological Sciences III, Irvine, CA 92697.

Conflict of interest statement:

The data in this manuscript were collected in compliance with ethical standards and have received oversight from the Internal Review Board at UCR. These data have not been published previously and are not under consideration for publication elsewhere. All authors have contributed significantly to this manuscript and have approved this submission. All authors have no conflicts of interests in the conduct or reporting of this research.

## Keywords

Diffusion imaging; NODDI; hippocampus; cognition; aging; MRI

---

## 1 Introduction

Magnetic Resonance Imaging (MRI) studies have been very valuable in non-invasively detecting structural changes associated with behavior, cognition, function, and pathology. While there are a range of MRI techniques that can capture different types of information, acquisition time is always at a premium in any study. It is challenging to determine which acquisition protocols to include in a study while finding the best balance between scan time and information received. Structural Magnetic Resonance Imaging (sMRI) techniques like T1-weighted, T2-weighted, and Fluid Attenuated Inversion Recovery (FLAIR) imaging have been crucial in understanding the anatomical alterations that the brain experiences, both in research and clinical contexts (Brans et al., 2010; Dekaban and Sadowsky, 1978; Hedman et al., 2011; Ho et al., 1980; Jernigan et al., 2001; Morrison and Hof, 1997; Peter R., 1979; Raz et al., 2004; Taki et al., 2011; van Haren et al., 2008). For example, being able to trace the outline and subsequently calculate the volume of the hippocampus using T1-weighted images has been extremely valuable when studying diseases like Alzheimer's (Fox and Freeborough, 1997). However, while these techniques can detect useful macrostructural properties like regional volume and cortical thickness, dramatic changes like gross atrophy or cortical thinning likely follow a myriad of more subtle microstructural changes that cannot be detected by standard T1- and T2-weighted imaging and their analogs (Jack et al., 2018). Cytoarchitectural properties like dendritic complexity, glial density, and cellular characteristics may be more reflective of individual differences in cognitive and pathological identities. Observing these microstructural characteristics in the brain non-invasively had been a major challenge to the field until recent advances in MRI technology.

Diffusion-weighted imaging (DWI) (Stejskal and Tanner, 1965) is increasingly becoming a key solution to this problem, given its ability to provide a range of quantitative measures describing neural microstructure within only minutes of acquisition (Johansen-Berg and Behrens, 2014). Until recently, DWI was almost exclusively associated with the characterization of white matter. Diffusion tensor metrics and tractography revealed how important features like axonal integrity, myelination, and specific structural connections change through healthy and pathological conditions and how they influence cognitive performance and other behavior (Assaf and Pasternak, 2008; Sasson et al., 2010; Thomason and Thompson, 2011; Madden et al., 2012). Given this, diffusion imaging has been considered a general microstructural probe as the geometry, organization, and morphology of any tissue type influences the diffusion signal. More recently, studies have been conducted using diffusion metrics to investigate gray matter microstructure (Aggarwal et al., 2015; Assaf, 2019; Budde and Annese, 2013; Colgan et al., 2016; Köhncke et al., 2021; Radhakrishnan et al., 2020; Venkatesh et al., 2020). However, the ratio of current publications that use diffusion MRI to study white matter vs gray matter is a striking 5:1, perhaps because early modeling algorithms could not accommodate the cytoarchitectural complexity of gray matter, and DWI used to provide poor image resolution in gray matter

sub-structures (Assaf, 2019; Nazeri et al., 2020). Recent advances in both diffusion image acquisition and analysis techniques may be able to resolve some of these issues (Frank, 2001; Jones, 2004; Papadakis et al., 1999).

A promising acquisition technique has been multi-shell DWI, which acquires scans at multiple gradient strengths (b-values). These images can then be analyzed using biophysically plausible models like Neurite Orientation Dispersion and Density Imaging (NODDI) (Zhang et al., 2012), which yield microstructural metrics that correspond to “intracellular”, “extracellular” and free water sources of the diffusion signal across tissue types. While NODDI may be well-suited to characterize the cytoarchitectural properties of gray matter, it has been seldom utilized to do so, especially in large-scale studies which usually only collect single-shelled diffusion data, and any multi-shelled data is collected in a much smaller sample size (Beekly et al., 2004; Petersen et al., 2010). This is mainly because increasing the number of shells can double or triple scan time (depending on the number of shells added), which can be a valuable commodity, especially when studying sensitive populations. The counterargument for including additional DWI shells may be strengthened if the advantage of models that need multi-shelled acquisitions (specifically in studying gray matter microstructure) can be documented, as proposed here.

In this study, we asked whether one such model utilizing multi-shell diffusion protocols, NODDI, can generate metrics that provide enough additional information over traditional tensor and volume metrics to justify the added acquisition time. To simplify this question and directly determine the value of NODDI metrics over the other metrics in a test case, we observed the effect of aging in the hippocampal subfields, and the relationship of this phenomenon with two popular hippocampal-dependent memory tasks. We first reproduced prior results from our lab and others (Radhakrishnan et al., 2020; Venkatesh et al., 2020) in a larger study population showing that NODDI metrics are sensitive to aging-related microstructural differences in hippocampal subfields and that differences in these NODDI metrics may be associated with cognitive decline. Our central question beyond this replication was how the hippocampal NODDI metrics compare to traditional tensor metrics and volume when predicting age or cognitive performance.

We found that NODDI metrics and tensor/volume metrics predicted age and cognition in similar capacities. However, integrating these modalities together significantly increased the predictive power of both our age and cognition models, suggesting that these methods may be capturing different types of information. We use these results to urge neuroimaging data collection consortiums to include acquisitions in their protocol that allow the calculation of NODDI metrics- i.e., multi-shell sequences, as they might be able to capture independently valuable information complementary to the more conventional measures.

## 2 Methods

### 2.1 Participants

Participants were recruited from the University of California, Riverside, and surrounding communities. Before enrollment, participants were screened for neurological conditions (e.g., depression, stroke, etc.) and scanner-related contraindications (e.g., claustrophobia,

pregnancy, etc.). After scanning, sixteen of the 170 participants were excluded based on data segmentation issues and/or registration artifacts. The final sample consisted of 79 younger adults ( $20.41 \pm 1.89$  years, 46 females) and 75 older adults ( $73.56 \pm 6.26$  years, 45 females) (Table 3.1). All participants provided informed consent before participation in this study and were compensated for their time. All experimental procedures were approved by the University of California, Riverside Review Board.

## 2.2 Cognitive testing

All participants completed a battery of neuropsychological tests to evaluate their cognitive abilities. We assessed participants' memory using the Rey Auditory Verbal Learning Test (RAVLT) (Rey, 1941) and the two-choice Mnemonic Similarity Test (MST) (Kirwan and Stark, 2007; Stark et al., 2019, 2013). The RAVLT has three components: 5 presentations of the same 15-word list with immediate recall, a second immediate recall test following an interference list of 15 new words, and a final delayed recall of the initial list after 15 minutes. The RAVLT Delay score reflects the final recall score, on a scale of 0 to 15.

The MST is a modified recognition memory task that was designed to tax "pattern separation" processes in an explicit attempt to rely on hippocampal processing and its link to the hippocampus has been validated in a wide range of domains (see Stark et al., 2019 for review). In the MST, participants viewed 128 images of everyday objects during an incidental encoding phase. During the test phase, participants were shown repeated images, novel foils, and items that were similar to, but not the same as studied images. On each trial, participants had to judge them as either "old" (repeated targets) or "new" (novel foils and similar lures) using a two-choice button press. A lure discrimination index (LDI) was calculated using signal detection theory as the discrimination  $d'$  between repeated targets and similar lures (Kirwan and Stark, 2007; Stark et al., 2015). A traditional recognition measure was calculated as the  $d'$  between repeated targets and novel foils. Participants who had >20% omitted trials or an extremely poor recognition score (greater than 2.5 standard deviations from the mean [ $REC < 0.5$ ]) were excluded. The final sample size for the MST analyses was 78 younger adults ( $20.41 \pm 1.89$  years, 46 females) and 69 older adults ( $74.04 \pm 6.35$  years, 35 females).

## 2.3 MR Image Acquisition

The participants were scanned using a Siemens Prisma 3T MRI scanner (Siemens Healthineers, Malvern, PA), fitted with a 32-channel receive-only head coil. Fitted padding was used to minimize head movements.

**T1W:** A T1-weighted magnetization-prepared rapid gradient echo (MP-RAGE) scan was acquired with the following parameters: echo time (TE)/repetition time (TR) = 2.72/2400 ms, 208 axial slices, GRAPPA acceleration factor = 2, and in 0.8mm isotropic resolution.

**DWI:** Axial diffusion-weighted echo-planar images were acquired in both anterior-posterior and posterior-anterior phase encodings with  $b = 1500 \text{ s/mm}^2$  and  $b = 3000 \text{ s/mm}^2$  applied in 64 orthogonal directions each with the following parameters: TE/TR = 102/3500 ms, FOV = 212 x 182 mm, 64 axial slices, Acquisition time = 10 minutes and 57 seconds, multi-band

acceleration factor = 4 and in 1.7 mm isotropic resolution. Twelve images with no diffusion weighting ( $b = 0$ ; half in each encoding direction) were also collected.

## 2.4 Diffusion data preprocessing

All preprocessing steps employed MRtrix3 (Tournier et al., 2012) ([www.mrtrix.org](http://www.mrtrix.org)) commands or used Mrtrix3 scripts that linked external software packages. Physiological noise arising from thermal motion of water molecules in the brain was first removed (Veraart et al., 2016), followed by removal of Gibbs ringing artifacts (Kellner et al., 2016), eddy current correction (Andersson and Sotiropoulos, 2016), motion correction (Andersson et al., 2003), susceptibility-induced distortion correction (Skare and Bammer, 2009) and bias field correction (Tustison et al., 2014). The image intensity was then normalized across subjects in the log domain (Raffelt et al., 2012). Images with no diffusion weighting ( $b=0$ ) were extracted and averaged to aid with structural registration.

## 2.5 Structural data processing

Each participant's structural image was nonlinearly co-registered to the average of their respective preprocessed  $b_0$  images using the ANTS Registration SyN algorithm with a b-spline transform (Avants et al., 2008; Tustison and Avants, 2013). Registration was manually checked to ensure accuracy, and DWI-registered T1w images were used for the rest of the analyses. The T1-weighted images were then processed using FMRIPREP version 20.2.1 (Esteban et al., 2018; Gorgolewski et al., 2011). Each volume was corrected for intensity non-uniformity using N4 Bias Field Correction from Advanced Normalization Tools (ANTS v2.3.4) (Tustison et al., 2010). The images were then skull stripped using the OASIS template. Brain surfaces were reconstructed using recon-all from FreeSurfer v7 (Dale et al., 1999) and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017). The hippocampus was previously hand-segmented on a template image into 3 subregions: a combined dentate gyrus and CA3 (combined due to resolution constraints; DG/CA3), CA1, and subiculum, based on our previous work (Stark and Stark, 2017). As the hippocampal atlas was in template space, spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (Fonov et al., 2009) was performed through nonlinear registration with ANTs, using brain-extracted versions of both T1w volume and template and using ANTs MultiLabel resampling technique. Regional volume was calculated by transforming the masks back to subject space and calculating the number of voxels encompassing each subfield multiplied with the image resolution. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM), and gray-matter (GM) was performed on the brain-extracted T1w using FAST from FSL v6.0.

## 2.6 Deriving diffusion metrics

All diffusion metrics were estimated in native space. Given the concerns about estimating tensor metrics from multi-shell data and diffusion being non-Gaussian at high  $b$  values, we calculated tensor metrics only from the  $b=1500$  shells. A weighted least squares (WLS) approach was first used to fit the diffusion tensor to the log signal, using weights based on empirical signal intensities (Basser et al., 1994). We repeated the weighted least squares

with weights determined by the signal predictions from the previous step (Veraart et al., 2013). We then generated maps of the following tensor-derived parameters: the mean apparent diffusion coefficient (ADC, sometimes also referred to as Mean Diffusivity or MD), fractional anisotropy (FA), axial diffusivity (AD, same as principal eigen value), and radial diffusivity (RD, equal to mean of the two non-principal eigen values) (Westin, 1997).

While these traditional tensor metrics are widely used, they were not originally designed to capture the complex cytoarchitectural properties of gray matter. Hence, we derived higher-order multi-compartment metrics using data from all shells using the Neurite Orientation Dispersion and Density Imaging (NODDI) (Zhang et al., 2012) model in the Microstructure Diffusion Toolbox (Harms et al., 2017). NODDI's metrics are tissue type agnostic and can readily be used in gray matter as it characterizes diffusion within each voxel as a combination of intracellular, extracellular, and CSF-based components. Here, we focus on three NODDI-derived parameters: the neurite density index (NDI), the orientation dispersion index (ODI), and the fractional isotropy (FISO). The NDI measures intracellular volume fraction and is calculated as the proportion of the voxel expressing unhindered diffusion along a given set of sticks, and restricted diffusion perpendicular to the same set of sticks. The ODI is a measure of tortuosity coupling an intracellular and extracellular space and models the extracellular space as hindered, gaussian anisotropic diffusion (very similar to, and hence highly correlated with, the tensor-derived FA). The amount of isotropic free volume within a voxel is measured by FISO and is usually proportional to the amount of CSF in a voxel. The intrinsic diffusivity was set to  $1.7 \mu\text{m}^2 \text{ms}^{-1}$ . Note that if this intrinsic diffusivity values is suboptimal for gray matter as suggested by some prior work (Guerrero et al., 2019), it would be expected to overestimate absolute NDI values but have minimal effect on the age group differences of interest. Because there was no significant age group difference in the number of voxels that may have insufficient signal to accurately estimate the NODDI metrics (i.e., voxels with  $\text{NDI} > 0.99$ ; (Emmenegger et al., 2021), no thresholding was applied. Outputs were transformed into atlas space where diffusion metrics in each of the hippocampal subfields were calculated by averaging the parameter maps using AFNI (Cox, 1996).

## 2.7 Statistical Analyses

All statistical analyses were performed in Python's Scipy (Jones et al., 2001) or GraphPad Prism 9.1.0. Age group differences in diffusion metrics were computed using Student's two-tailed t-tests (Student, 1908). Relationships between diffusion metrics and age or cognition scores were evaluated using ordinary least squares regression. Absolute subfield values are reported here unless stated otherwise. All diffusion parameters were evaluated independently given the high correlation within the diffusion metrics themselves (Supplementary figure S1). No nuisance covariates were included in the model and subfield volumes were not adjusted for head size to keep consistent with other metrics. In separate, but related work, we tested the efficacy of using an ANCOVA-based hippocampal volume correction and failed to find any effect and so did not include it here.

Statistical p-values were corrected for multiple comparisons using the Holm-Sidak test (Holm, 1979), unless running tests with *a priori* hypotheses. Receiver operating

characteristic (ROC) curves and their corresponding areas under the curve (AUC) were calculated using statsmodels (Seabold and Perktold, 2010). ROC curves were used to quantify the ability of these diffusion metrics to predict age or cognitive performance, and additional information about these tests is provided under the respective results sections.

## 2.8 Data and code availability

Raw and preprocessed data supporting the findings of this study would be available upon reasonable request, made via email to the corresponding author (CS). The code used for analysis is in a GitHub repository ([https://github.com/StarkLabUCI/NODDiffusion\\_ROCs](https://github.com/StarkLabUCI/NODDiffusion_ROCs)) which will be made available upon acceptance for publication.

## 3 Results

We had previously shown that NDI of the DG/CA3 was increased in older adults and was negatively associated with RAVLT Delay ( $n = 38$ ) (Radhakrishnan et al., 2020). We first reproduced these results in a larger population before diving into more complex analyses.

### 3.1 NDI of all hippocampal subfields is increased in older adults

With this larger sample size, we found that the NDI of all the hippocampal subfields was greater in the older group as compared to the younger group (Two-sample t-test. DG/CA3:  $t = 5.30$ ,  $p < 0.0001$ ; CA1:  $t = 4.32$ ,  $p < 0.0001$ ; Subiculum:  $t = 4.10$ ,  $p < 0.0001$ ). Moreover, the NDI of the DG/CA3 and CA1 subfields increased with age within the older subpopulation alone (Linear regression. DG/CA3:  $R^2 = 0.09$ ,  $p = 0.04$ ; CA1:  $R^2 = 0.179$ ,  $p = 0.002$ ; Subiculum:  $R^2 = 0.025$ ,  $p = 0.277$ ). Because of the lack of variance in age in the younger population and the absence of middle-aged data, we could not perform correlations with age in just the young group or across the lifespan, in a meaningful way. The relationship between all diffusion metrics studied and age can be found in Supplementary Tables S1 and S2.

We next asked whether these diffusion metrics followed a different pattern with age compared to the global average. To determine whether these relationships had any sort of selectivity towards the hippocampal subfields and were not just a consequence of age-related global gray matter decline, we modeled the subfield-specific NDI linearly against the average whole-brain gray matter NDI as described in our previous study (Radhakrishnan et al., 2020). The residuals of this highly correlated model were quantified as the “globally regressed” NDI for each subfield. Post global regression, we found that the older adults not only still had higher hippocampal NDI, but the globally regressed NDI for older adults was also largely positive, while that for the younger adults was largely negative. This suggests that younger adults have a hippocampal NDI that is well below the whole brain average, while older adults have a hippocampal NDI above the global average (Figure 1; Two-sample t-test. DG/CA3:  $t = 3.89$ ,  $p = 0.0002$ ; CA1:  $t = 4.63$ ,  $p < 0.0001$ ; Subiculum:  $t = 4.28$ ,  $p < 0.0001$ ), consistent with a focused age-related NDI increase in the hippocampal subfields. The globally regressed NDI was also positively correlated with age within the older subpopulation in the DG/CA3 and CA1, but not in the subiculum. (Figure 1; Linear regression. DG/CA3:  $R^2 = 0.081$ ,  $p = 0.019$ ; CA1:  $R^2 = 0.132$ ,  $p = 0.002$ ; Subiculum:  $R^2 =$

0.005,  $p = 0.847$ ). To ensure that nothing about our sample was driving the observed effect and that any deviation from normality might be altering our results, we performed 1000 random samplings of 70% of our data. The resulting slopes were entirely consistent with our regression-based confidence intervals.

We found no differences between hemispheric metrics for all subfields. The NDI was greater in biologically male participants compared to biologically female participants for all subfields (Two-sample t-test. DG/CA3:  $t = 2.04$ ,  $p = 0.043$ ; CA1:  $t = 2.036$ ,  $p = 0.044$ ; Subiculum:  $t = 3.74$ ,  $p = 0.0003$ ), but the previously described age effects remained significant after controlling for sex.

### 3.2 Hippocampal NDI is negatively associated with RAVLT Delay, and more weakly with the LDI

We were also able to replicate our previous finding that the hippocampal subfield NDI was negatively correlated with RAVLT performance, even after factoring in age as a regressor. While previously (Radhakrishnan et al., 2020), we observed reliable correlations in only the DG/CA3, here the NDI of all hippocampal subfields was negatively correlated with RAVLT delay, both before and after global regression, suggesting that this relationship in the hippocampus may not be a general brain-wide phenomenon (Figure 2). Within the older subpopulation, age was not significantly correlated with RAVLT Delay. However, the NDI of all hippocampal subfields still trended towards a negative relationship with the RAVLT delay in both the younger age group (Linear regression. DG/CA3:  $R^2 = 0.059$ ,  $p = 0.007$ ; CA1:  $R^2 = 0.075$ ,  $p = 0.001$ ; Subiculum:  $R^2 = 0.143$ ,  $p < 0.0001$ ) as well as the older age group (Linear regression. DG/CA3:  $R^2 = 0.054$ ,  $p = 0.05$ ; CA1:  $R^2 = 0.080$ ,  $p = 0.021$ ; Subiculum:  $R^2 = 0.195$ ,  $p = 0.0002$ ), with the subiculum NDI having the strongest relationship in both groups, suggesting that this relationship was beyond just an effect of age, and that hippocampal subfield NDI might be capable of capturing individual differences associated with cognition. The relationship between all diffusion metrics studied and RAVLT performance can be found in Supplementary Table S3.

A more selective and weaker relationship was found between performance in the MST, as measured by the LDI, and the NDI of the hippocampal subfields. While the raw NDI of both the DG/CA3 and subiculum were negatively associated with LDI (Linear regression. DG/CA3:  $R^2 = 0.030$ ,  $p = 0.043$ ; CA1:  $R^2 = 0.021$ ,  $p = 0.087$ ; Subiculum:  $R^2 = 0.1$ ,  $p = 0.0002$ ), this relationship was only weakly significant in the subiculum after global regression (Linear regression. DG/CA3:  $R^2 = 0.002$ ,  $p = 0.557$ ; CA1:  $R^2 = 0.007$ ,  $p = 0.346$ ; Subiculum:  $R^2 = 0.03$ ,  $p = 0.0434$ ). As in the previous analysis, we performed 1000 samplings of 70% of our data for both RAVLT delay and LDI, and all resulting slopes were consistent with our confidence intervals. Interestingly, none of the raw diffusion metrics were significantly correlated with LDI (Supplementary Table S4).

### 3.3 Tensor, NODDI, and volumetric measures of hippocampal subfields can all successfully predict age group

The results in sections 3.1 and 3.2 reproduce previously reported findings in a larger dataset and establish that the relationships between age, NDI, and cognition are reliable



and consistent. We next wanted to assess how NDI and other NODDI metrics compare to tensor metrics and coarse volume in their relationships with age and cognition. Note that some hippocampal diffusion metrics like AD and ADC were not correlated with their global averages, so we used only *raw* diffusion metrics (not globally regressed) for the rest of the study to be consistent across metrics.

As an initial test, we sought to determine how well the various diffusion and volumetric measures from the hippocampus could classify participants into their age groups. To do this, we used ROC curves to evaluate the performance of each metric in accurately distinguishing between young and old groups. We calculated the AUC for each measure, by passing in the metric for each subfield split by hemisphere (for a total of 6 input features) and fitting a logistic regression model to derive the ROC and consequently predict the age group. We used hemisphere-specific metrics instead of bilateral ones simply because they generated the greatest AUC for most metrics (Supplementary Table S5). Since the input features were highly collinear (Supplementary Figure S1), we used the newton conjugate gradient method (Buckley, 1978; Knoll and Keyes, 2004) for optimization to prevent a non-Hessian matrix error. To calculate the null distribution and estimate the p-value, we conducted a permutation analysis by randomly shuffling the old/young labels and estimating the AUC over 5000 iterations using the volumetric input features (since the null distributions would essentially remain equivalent across all metrics). We found that most metrics studied could successfully distinguish between young and old age groups well above chance with just the six inputs from that metric in each hippocampal subfield alone (Figure 3), with the ODI and volume yielding the highest AUCs. All NODDI metrics and volume outperformed the tensor metrics. Given the difference in age between the groups and considering that aging causes very dramatic changes in the brain, it is unsurprising that all metrics, even volume, are sensitive to these changes.

We then asked whether we could achieve higher prediction accuracy inputting a specific combination of these metrics, instead of a single metric from all subfields, and whether having the NODDI metrics posed any advantage over just the tensor metrics and volume in predicting age group. To answer this, we calculated the AUC for each combination of 6 input features from either a) a composite of tensor metrics and volume [ ${}^3C = 593,775$  combinations] b) NODDI metrics [ ${}^18C = 18,564$  combinations] or c) a composite of all metrics studied i.e., volume, tensor, and NODDI metrics [ ${}^{48}C = 12,271,512$  combinations] (Figure 4 a-c). We found that the 99<sup>th</sup> percentile of the AUC distribution from permuting over just the NODDI metrics (0.90) was comparable with that sampled from the tensor metrics and volume (0.91). However, the 99<sup>th</sup> percentile of the AUC distribution of combinations of all metrics was significantly higher at 0.98. More specifically, the 99<sup>th</sup> percentile of the Tensor metrics + Volume distribution was equal to the 57.57<sup>th</sup> percentile of the All metrics distribution, while the 99<sup>th</sup> percentile of the just NODDI distribution was equal to 57.38<sup>th</sup> percentile of the All metrics distribution. To isolate this and understand it more completely, we then asked which individual features were most often contributing to the highest AUC values (> 99<sup>th</sup> percentile) over all the combinations (Figure 4d). Within these best-performing combinations, we found all metrics and all subfields were consistently represented. Note that the distributions were non-Gaussian when volume was included in

the selection features, due to the lack of collinearity of volume with the other metrics while using the newton conjugate gradient optimization algorithm (Supplementary Figure S1). The results remained consistent even after removing volume from the combinations and considering only the diffusion metrics (Figure 4e).

### 3.4 A combination of NODDI, tensor metrics, and volume predict RAVLT performance better than any of them alone

We then repeated the same analysis as the previous section, this time for predicting RAVLT delay. We binarized the RAVLT score at a threshold of 9, such that all those who scored higher than 9 were considered “high performing”, and those who scored 9 or below were considered “low performing”, consistent with previous findings for these age groups (Stark et al., 2013, 2010). Binarizing the RAVLT in this way also helped reduce noise that might arise from screener differences or individual participants having “off” days (it is more likely that a binarized RAVLT score would remain consistent over multiple testing days, compared to the continuous score, speaking to its robustness). While calculating the AUC for predicting high/low RAVLT delay this way in *the entire subpopulation*, using each metric from all 6 subfield features, we found that ODI, NDI, and FA were the best predictors of RAVLT performance (Figure 5).

We next performed the same  $\chi^2$  analysis of all possible combinations of metrics. Notably, when examining AUCs resulting from combinations of only NODDI metrics and AUCs resulting from only tensor+volume metrics, we observed similar distributions with AUCs at the 99<sup>th</sup> percentile of 0.79 and 0.78 respectively. However, when combinations of NODDI and Tensor metrics + volume metrics were included, we observed a distribution with the AUC at the 99<sup>th</sup> percentile = 0.95 (Figure 6), and the AUCs of the separate distributions were both equal to about the ~57<sup>th</sup> percentile of the all metrics distribution. This indicates that NODDI and Tensor metrics + volume, while correlated, contain independent aspects of variance that are useful in modeling RAVLT performance. As expected from these results, the input features that most frequently contributed to the top AUCs spanned all metrics and all subfields. This effect remained even when observing the diffusion metrics alone, without volume.

Given the correlation between age and RAVLT performance, it is certainly possible that our prediction of RAVLT status is driven by our ability to predict age group. To assess whether these metrics were capturing cognitive differences beyond just a function of age, we repeated this analysis separately in the younger and older subpopulations. In the older group, we found that these results were just as reliable (99<sup>th</sup> percentile of AUC distribution with Tensor metrics + Volume: 0.90; NODDI metrics: 0.85; All metrics: 0.98). The peak AUCs of all conditions were greater when predicting RAVLT delay in just the older group compared to the entire study sample. Notably, the features that contributed to the highest AUCs here were not in the same order as with the entire sample size but still appeared to span across all metrics and subfields evenly. In the younger group, we found that combining the metrics did not result in a significantly higher AUC distribution over just the tensor metrics or NODDI metrics (99<sup>th</sup> percentile of AUC distribution with Tensor metrics + Volume: 0.72; NODDI metrics: 0.71; All metrics: 0.70), possibly due to the low number of poor performers (75% of

the high-performing RAVLT participants belong to the younger age group), or because these metrics weren't sensitive to microstructural properties that contributed to low RAVLT delay at a young age.

### 3.5 A combination of NODDI and tensor metrics predicts MST performance better than either of them alone

We also found similar results when predicting MST performance, as relayed by the LDI. Here, we binarized the LDI at a threshold of 1.25, to match the proportion used in our RAVLT analysis. We found that the AUCs for predicting high/low LDI were lower overall than the corresponding high/low RAVLT or age group (Figure 7) when restricting ourselves to individual metrics from all 6 subfields. Interestingly, only volume predicted MST scores better than chance.

Turning to the n-choose-6 permutations, we observed a familiar pattern. The 99th percentile of the AUC distribution obtained from using just the NODDI metrics or just the Tensor metrics + Volume were 0.67 and 0.80, respectively, but using All metrics resulted in a bimodal distribution with a 99th percentile AUC of 0.94 (Figure 8), again with the separate distributions having their 99th percentile equal to below the 60th percentile in the complete distribution. As with RAVLT delay, the most informative input features spanned all metrics across all subfields. We were also able to reproduce these results in just the older subgroup (99th percentile of AUC distribution with Tensor metrics + Volume: 0.86; NODDI metrics: 0.75; All metrics: 0.97). As with RAVLT performance, combining the metrics did not help the AUC in the younger subgroup (99th percentile of AUC distribution with Tensor metrics + Volume: 0.73; NODDI metrics: 0.77; All metrics: 0.69). These results again indicate that while NODDI and traditional tensor metrics can be correlated with each other, they provide unique information and contributions to the variance that can be used to model memory performance.

## 4 Discussion

In this study, we show that at least one model of multi-shelled diffusion-weighted imaging (NODDI) can provide significant benefits in determining aging-associated microstructural differences in the hippocampal subfields and their cognitive consequences, when used in conjunction with tensor and volume measures. We first reproduced our previous studies in a larger sample size (about three-fold of that reported in our prior publications) showing that the NODDI metric NDI is increased in older adults, and this increase may be partially driving aging-associated memory decline. The ability to reproduce these effects across independent populations and study centers speaks to the robustness and reliability of these relationships, which is a major concern in neuroimaging research (Poldrack et al., 2017).

We also found that NDI increases with age within the older population alone (ages 65 – 92) in the DG-CA3 and CA1, but not in the subiculum. However, the negative relationship between NDI and RAVLT performance was strongest in the subiculum in both the entire study population as well as the older subgroup alone, potentially suggesting a subfield-specific pattern of NDI differentially associating with cognition at the outset of aging. The other hippocampal diffusion metrics derived from both NODDI and the tensor models

had selective relationships with age and RAVLT delay and LDI across the subfields (see supplementary material) that are consistent with prior reports in the literature (Fukutomi et al., 2019; Mortimer et al., 2004; Müller et al., 2005; Nazeri et al., 2015; Nobis et al., 2019; Yassa et al., 2011). These similar results across subfields may, at least in part, be due to inherent smoothing when transforming between high-resolution ROI segmentations ( $1\text{mm}^3$ ) and lower resolution diffusion images ( $1.7\text{mm}^3$ ), performing the segmentations using only T1-weighted images (Wisse et al., 2021), or using diffusion data that is not ideally suited to subfield work but is in higher resolution than what is often used. Moreover, our prediction models rely on highly correlated input features, across both metrics and subfields, which puts us at risk of overfitting data. While we theorize that differences in segmentation would not change the main findings of this study, future work examining differences as a function of methodological details would be valuable. Our finding of reliable whole hippocampal effects but little in the way of subfield-specific effects should not be taken as evidence of absence of subfield differentiation, although it does suggest that many studies can be effectively conducted without specific attention to the subfields (or without meticulous attention to the methodological details). Moreover, because we did not observe anything that would suggest a cancelation of effects via opposing patterns across subfields, future work using lower resolutions and even more blending of signals across subfields can still make clear contributions at the level of the hippocampus itself.

Also consistent with several previous studies (Aggarwal et al., 2015; Assaf, 2019; Leuze et al., 2014; Truong et al., 2014), we found that even the most basic diffusion metrics like FA and RD were sensitive to age and cognitive measures. They could detect specific microstructural properties of gray matter and were mainly better than volume in detecting individual cognitive differences. Interestingly, relationships between these diffusion metrics and aging/cognition were largely independent of hippocampal volume. Moreover, despite being highly correlated with each other, none of the diffusion metrics were associated with volume in any of the subfields. These observations suggest that hippocampal microstructure as relayed by the diffusion metrics, and macrostructure measured by subfield volume, could represent independent processes in normal aging.

Not only could diffusion metrics be reliably used in gray matter to make inferences about aging and cognition, we also found that including NODDI metrics greatly improved the ability of our statistical models to predict age and cognition. We find that a combination of NODDI and tensor measures consistently result in better age group predictions than either modality by itself. Interestingly, when using ROC analysis to predict memory performance, we observed that while combinations of just tensor metrics and volume, or just NODDI metrics yielded laudable AUCs, combining the tensor metrics, volume, and NODDI metrics could result in a near-perfect prediction model. Additionally, when predicting both RAVLT and MST performance, this effect only got stronger when looking within just the older subpopulation (resulting in AUCs of over 0.98). This enhancement could be because the neurobiological properties that help define cognitive performance are more likely to be consistent under the common effect of aging, as compared to a much wider age range.

The increase in the predictive power of the model in the older group also demonstrates that these metrics could be capturing neurobiological properties that go beyond just the

dramatic effect of age and might be sensitive to individual differences within this age group as well. This could be attributed to the idea that NODDI and tensor metrics may be capturing related but neurobiologically distinct properties associated with aging. The ability of NODDI to separate intracellular and extracellular spaces might result in sensitivity to properties like inflammation and cellularity (Garcia-Hernandez et al., 2020 ; Yi et al., 2019) that complement tensor measures of anisotropy and diffusivity. While the biological implications of these diffusion metrics require much more examination, the inability of just volume to be a relatively strong predictor of cognition in both the entire population as well as just the older adults emphasizes that aging-associated cognitive decline is a function of more microstructural changes.

The ability of the combined diffusion metrics to successfully predict LDI (Figure 8) despite none of them being correlated directly with the measure also raises an important question about how we approach relationships between neuroimaging measures and cognition. Different patterns of cytoarchitecture could result in the exact same value for certain diffusion metrics but might not result in the exact same value for all diffusion metrics.

Most prior works attempting to tie implicit neurobiological properties with diffusion imaging have tended to look at these relationships separately within each metric. However, as demonstrated in this paper, treating these varied diffusion metrics as a unique “signature” for each participant can generate more specific models that can predict behavior successfully. However, even with such a method, extreme caution must be taken when speculating on the neurobiological implications of these diffusion metrics.

The results in this paper collectively present a strong case in support of including multi-shelled diffusion sequences that allow analysis like NODDI modelling in study protocols, despite the increase in acquisition time. The full diffusion sequence here required ~16 minutes to acquire both phase-encoding directions. However, excellent results are still possible with diffusion data from one phase encoding and only b0 images from the reverse (or a separate phase map), reducing the acquisition time to ~9 minutes. Adding just a single shell and a modest number of directions over typical DTI protocols allowed for NODDI analyses that generated metrics that immensely improved the predictive power of our model, suggesting that combining higher-order models with tensor metrics could uncover large effects not discernible by either analysis technique alone. Of course, including multi-shell acquisition protocols also opens up analysis to other complex models (even those yet to be created) that could be more sensitive to specific microstructural properties in gray matter. We posit that these results observed in hippocampal aging likely extend to other domains of interest. Researchers studying the microstructural mechanisms of both healthy developmental stages, as well as various neuropsychiatric illnesses, could benefit from implementing NODDI in their protocols. In such studies, however, it is important to identify the “domain of validity” where both the assumptions and interpretations of these diffusion metrics are accurate (Lampinen et al., 2019), as diffusion indices tend to be very context-specific. Seemingly similar changes in diffusion metrics may have different anatomical implications across pathologies, age groups, and even across brain regions. For example, we report consistent and reproducible increases in hippocampal NDI with age. However, the NDI of other regions, especially white matter, have been shown to decrease

with age with similar associated cognitive decline (Merluzzi et al., 2016). Moreover, conditions like schizophrenia and bipolar disorder present decreased levels of hippocampal NDI, with potentially very different neurobiological implications (Nazeri et al., 2017). More research examining gray matter microstructure using these methods across the whole brain, in both the healthy lifespan and disease, can help in developing powerful signatures for both developmental milestones as well as neurological illnesses. This is another reason we advocate for the inclusion of multi-shell acquisition schemes, that allow for analysis like NODDI modelling, in large-scale data collection efforts.

Of note, NODDI is not the only analysis technique made possible with multiple shells. We use NODDI here as one example of a multi-shell technique and demonstrate that it can improve sensitivity to specific microstructural properties related to aging and cognition. An interesting question for future research is the extent to which other multi-shell (e.g., diffusion kurtosis imaging (Steven et al., 2014), multiple Q-shell imaging (Descoteaux et al., 2011)) or even single-shell (e.g., bi-tensor models (Pasternak et al., 2009), Q-ball imaging (Tuch, 2004), constrained spherical deconvolution (Nath et al., 2020)) models perform similarly to NODDI. However, this will first require validation of these approaches in gray matter as they have been designed to improve sensitivity to white matter microstructure. We speculate that the advantages of NODDI observed here (and potentially observed with related multi-shell or single-shell bi-tensor models) may be because it separately estimates free water diffusion. In line with this view, our group has previously shown that single tensor metrics are worse at predicting age after accounting for free diffusion by excluding voxels with low tissue content (i.e., those with high free diffusion) (Venkatesh et al, 2020). This suggests that single-tensor metrics are significantly influenced by free diffusion, which may reflect partial volume effects or neurodegeneration in aging. Of note, NODDI outperformed both the tensor metrics even after accounting for free water when predicting age and memory performance, supporting an advantage for multi-shell approaches as shown in the current study. Studies that compare metrics derived from such techniques to the metrics used in this paper may reveal valuable information on the right analysis technique to use for a given context.

While our spatial resolution is close to its optimum for a 3T scanner, one could argue that the extra time required to scan the additional shell could instead be utilized elsewhere-like in increasing angular resolution. Previous studies have explored the effects of different acquisition schemes on both single-shell and multi-shell metrics in white matter and have consistently shown that the value of adding extra directions begins to taper off at around  $n(\text{directions}) = \sim 60$  for a single shell (Li et al., 2020; Schilling et al., 2018; Tournier et al., 2013). The acquisition scheme used in this study was optimized for these observations and included 64 directions. Increasing scan times, even by small amounts, result in many issues in studies involving children or participants with neurological conditions. To accommodate for research where long scan times are not feasible, more work that examines the benefits of additional shells over additional directions (angular resolution) or higher spatial resolution is required as advantages of these acquisition schemes have been demonstrated in white matter. Moreover, since the data presented here are cross-sectional, we cannot conclude that NODDI metrics would display the same advantage when studying the temporal dynamics of phenomena like aging. Longitudinal studies and studies monitoring interventions (Eaton-

Rosen et al., 2015; Kamiya et al., 2020; Radhakrishnan et al., 2021) could greatly enhance our understanding of the benefits of these metrics.

Another limitation is that our models binarize age and continuous cognitive measures, so we cannot evaluate whether NODDI metrics would complement tensors and volume so clearly when predicting absolute age and cognition. However, these binarized cognitive scores are far more robust than the continuous scores, and much less prone to noise or sample bias. Though this study is fairly well-balanced by sex, it is also important to acknowledge the role of other demographic differences (like race and socioeconomic status) often arising from volunteer and selection biases in most human neuroimaging studies.

In conclusion, we have demonstrated that NODDI measures complement tensor metrics and volume when capturing age and cognition related information. The extra diffusion-weighted shell needed for NODDI provided predictive and neurobiologically relevant information that are ancillary to but seemingly distinct from what the tensor measures provide. NODDI metrics offer obvious value beyond and/or in conjunction with traditional tensor and volumetric measures when studying gray matter microstructure and might even be able to explain individual differences in cognition or behavior.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Funding:

This study was funded by P30 AG066519 and R01 AG034613 given to CS and R00 AG047334 and R21 AG054804 given to IJB.

## 1. References

- Aggarwal M, Nauen DW, Troncoso JC, Mori S, 2015. Probing region-specific microstructure of human cortical areas using high angular and spatial resolution diffusion MRI. *NeuroImage* 105, 198–207. 10.1016/j.neuroimage.2014.10.053 [PubMed: 25449747]
- Andersson JLR, Skare S, Ashburner J, 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20, 870–888. 10.1016/S1053-8119(03)00336-7 [PubMed: 14568458]
- Andersson JLR, Sotiropoulos SN, 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* 125, 1063–1078. 10.1016/j.neuroimage.2015.10.019 [PubMed: 26481672]
- Assaf Y, 2019. Imaging laminar structures in the gray matter with diffusion MRI. *NeuroImage* 197, 677–688. 10.1016/j.neuroimage.2017.12.096 [PubMed: 29309898]
- Assaf Y, Pasternak O, 2008. Diffusion Tensor Imaging (DTI)-based White Matter Mapping in Brain Research: A Review. *J. Mol. Neurosci* 34, 51–61. 10.1007/s12031-007-0029-0 [PubMed: 18157658]
- Avants BB, Epstein CL, Grossman M, Gee JC, 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal* 12, 26–41. 10.1016/j.media.2007.06.004 [PubMed: 17659998]
- Basser PJ, Mattiello J, Lebihan D, 1994. Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo. *J. Magn. Reson. B* 103, 247–254. 10.1006/jmrb.1994.1037 [PubMed: 8019776]

- Beekly DL, Ramos EM, van Belle G, Deitrich W, Clark AD, Jacka ME, Kukull WA, NIA-Alzheimer's Disease Centers, 2004. The National Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. *Alzheimer Dis. Assoc. Disord* 18, 270–277. [PubMed: 15592144]
- Brans RGH, Kahn RS, Schnack HG, Baal G.C.M. van, Posthuma D, Haren N.E.M. van, Lepage C, Lerch JP, Collins DL, Evans AC, Boomsma DI, Pol HEH, 2010. Brain Plasticity and Intellectual Ability Are Influenced by Shared Genes. *J. Neurosci* 30, 5519–5524. 10.1523/JNEUROSCI.5841-09.2010 [PubMed: 20410105]
- Buckley AG, 1978. A combined conjugate-gradient quasi-Newton minimization algorithm. *Math. Program* 15, 200–210. 10.1007/BF01609018
- Budde M, Annese J, 2013. Quantification of anisotropy and fiber orientation in human brain histological sections. *Front. Integr. Neurosci* 7, 3. 10.3389/fnint.2013.00003 [PubMed: 23378830]
- Colgan N, Siow B, O'Callaghan JM, Harrison IF, Wells JA, Holmes HE, Ismail O, Richardson S, Alexander DC, Collins EC, Fisher EM, Johnson R, Schwarz AJ, Ahmed Z, O'Neill MJ, Murray TK, Zhang H, Lythgoe MF, 2016. Application of neurite orientation dispersion and density imaging (NODDI) to a tau pathology model of Alzheimer's disease. *NeuroImage* 125, 739–744. 10.1016/j.neuroimage.2015.10.043 [PubMed: 26505297]
- Cox RW, 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res* 29, 162–173. [PubMed: 8812068]
- Dale AM, Fischl B, Sereno MI, 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage* 9, 179–194. 10.1006/nimg.1998.0395 [PubMed: 9931268]
- Dekaban AS, Sadovsky D, 1978. Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. *Ann. Neurol* 4, 345–356. 10.1002/ana.410040410 [PubMed: 727739]
- Descoteaux M, Deriche R, Le Bihan D, Mangin J-F, Poupon C, 2011. Multiple q-shell diffusion propagator imaging. *Med. Image Anal* 15, 603–621. 10.1016/j.media.2010.07.001 [PubMed: 20685153]
- Eaton-Rosen Z, Melbourne A, Orasanu E, Cardoso MJ, Modat M, Bainbridge A, Kendall GS, Robertson NJ, Marlow N, Ourselin S, 2015. Longitudinal measurement of the developing grey matter in preterm subjects using multi-modal MRI. *NeuroImage* 111, 580–589. 10.1016/j.neuroimage.2015.02.010 [PubMed: 25681570]
- Emmenegger TM, David G, Ashtarayeh M, Fritz FJ, Ellerbrock I, Helms G, Balteau E, Freund P, Mohammadi S, 2021. The Influence of Radio-Frequency Transmit Field Inhomogeneities on the Accuracy of G-ratio Weighted Imaging. *Front. Neurosci* 15, 770. 10.3389/fnins.2021.674719
- Esteban O, Markiewicz C, Blair RW, Moodie C, Isik AI, Erramuzpe Aliaga A, Kent J, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh S, Wright J, Durnez J, Poldrack R, Gorgolewski KJ, 2018. FMRIPrep: a robust preprocessing pipeline for functional MRI. 10.1101/306951
- Fonov VS, Evans AC, McKinstry RC, Almlri CR, Collins DL, 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47, S102. 10.1016/S1053-8119(09)70884-5
- Fox NC, Freeborough PA, 1997. Brain atrophy progression measured from registered serial MRI: validation and application to Alzheimer's disease. *J. Magn. Reson. Imaging JMRI* 7, 1069–1075. 10.1002/jmri.1880070620 [PubMed: 9400851]
- Frank LR, 2001. Anisotropy in high angular resolution diffusion-weighted MRI. *Magn. Reson. Med* 45, 935–939. 10.1002/mrm.1125 [PubMed: 11378869]
- Fukutomi H, Glasser MF, Murata K, Akasaka T, Fujimoto K, Yamamoto T, Autio JA, Okada T, Togashi K, Zhang H, Van Essen DC, Hayashi T, 2019. Diffusion Tensor Model links to Neurite Orientation Dispersion and Density Imaging at high b-value in Cerebral Cortical Gray Matter. *Sci. Rep* 9, 12246. 10.1038/s41598-019-48671-7 [PubMed: 31439874]
- Garcia-Hernandez R, Carpena AT, Drakesmith M, Koller K, Jones DK, Canals S, Santis SD, 2020. Imaging Microglia and Astrocytes non-invasively using Diffusion MRI (preprint). *Neuroscience*. 10.1101/2020.02.07.938910
- Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS, 2011. Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinformatics* 5. 10.3389/fninf.2011.00013



- Guerrero JM, Adluru N, Bendlin BB, Goldsmith HH, Schaefer SM, Davidson RJ, Kecksemeti SR, Zhang H, Alexander AL, 2019. Optimizing the intrinsic parallel diffusivity in NODDI: An extensive empirical evaluation. *PLOS ONE* 14, e0217118. 10.1371/journal.pone.0217118 [PubMed: 31553719]
- Harms RL, Fritz FJ, Tobisch A, Goebel R, Roebroek A, 2017. Robust and fast nonlinear optimization of diffusion MRI microstructure models. *NeuroImage* 155, 82–96. 10.1016/j.neuroimage.2017.04.064 [PubMed: 28457975]
- Hedman AM, van Haren NEM, Schnack HG, Kahn RS, Hulshoff Pol HE, 2011. Human brain changes across the life span: A review of 56 longitudinal magnetic resonance imaging studies. *Hum. Brain Mapp* 33, 1987–2002. 10.1002/hbm.21334 [PubMed: 21915942]
- Ho KC, Roessmann U, Straumfjord JV, Monroe G, 1980. Analysis of brain weight. I. Adult brain weight in relation to sex, race, and age. *Arch. Pathol. Lab. Med* 104, 635–639. [PubMed: 6893659]
- Holm S, 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat* 6, 65–70.
- Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, Liu E, Molinuevo JL, Montine T, Phelps C, Rankin KP, Rowe CC, Scheltens P, Siemers E, Snyder HM, Sperling R, 2018. NIA-AA Research Framework: Toward a biological definition of Alzheimer’s disease. *Alzheimers Dement. J. Alzheimers Assoc* 14, 535–562. 10.1016/j.jalz.2018.02.018
- Jernigan TL, Archibald SL, Fennema-Notestine C, Gamst AC, Stout JC, Bonner J, Hesselink JR, 2001. Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol. Aging* 22, 581–594. [PubMed: 11445259]
- Johansen-Berg H, Behrens TEJ (Eds.), 2014. *Diffusion MRI: from quantitative measurement to in-vivo neuroanatomy*, 2nd ed. ed. Elsevier/Academic Press, London, UK ; Waltham, MA.
- Jones DK, 2004. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: A Monte Carlo study†. *Magn. Reson. Med* 51, 807–815. 10.1002/mrm.20033 [PubMed: 15065255]
- Jones E, Oliphant T, Peterson P, 2001. *SciPy: Open Source Scientific Tools for Python* [WWW Document]. *SciPy.org*. URL <https://www.scipy.org> (accessed 8.12.19).
- Kamiya K, Hori M, Aoki S, 2020. NODDI in clinical research. *J. Neurosci. Methods* 346, 108908. 10.1016/j.jneumeth.2020.108908 [PubMed: 32814118]
- Kellner E, Dhital B, Kiselev VG, Reiser M, 2016. Gibbs-ringing artifact removal based on local subvoxel-shifts: Gibbs-Ringing Artifact Removal. *Magn. Reson. Med* 76, 1574–1581. 10.1002/mrm.26054 [PubMed: 26745823]
- Kirwan CB, Stark CEL, 2007. Overcoming interference: An fMRI investigation of pattern separation in the medial temporal lobe. *Learn. Mem* 14, 625–633. [PubMed: 17848502]
- Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M, Chaibub Neto E, Keshavan A, 2017. Mindboggling morphometry of human brains. *PLOS Comput. Biol* 13, e1005350. 10.1371/journal.pcbi.1005350 [PubMed: 28231282]
- Knoll DA, Keyes DE, 2004. Jacobian-free Newton–Krylov methods: a survey of approaches and applications. *J. Comput. Phys* 193, 357–397. 10.1016/j.jcp.2003.08.010
- Köhncke Y, Düzel S, Sander MC, Lindenberger U, Kühn S, Brandmaier AM, 2021. Hippocampal and Parahippocampal Gray Matter Structural Integrity Assessed by Multimodal Imaging Is Associated with Episodic Memory in Old Age. *Cereb. Cortex* 31, 1464–1477. 10.1093/cercor/bhaa287 [PubMed: 33150357]
- Lampinen B, Szczepankiewicz F, Novén M, van Westen D, Hansson O, Englund E, Mårtensson J, Westin C-F, Nilsson M, 2019. Searching for the neurite density with diffusion MRI: Challenges for biophysical modeling. *Hum. Brain Mapp* 40, 2529–2545. 10.1002/hbm.24542 [PubMed: 30802367]
- Leuze CWU, Anwander A, Bazin P-L, Dhital B, Stüber C, Reimann K, Geyer S, Turner R, 2014. Layer-Specific Intracortical Connectivity Revealed with Diffusion MRI. *Cereb. Cortex* 24, 328–339. 10.1093/cercor/bhs311 [PubMed: 23099298]
- Li C-X, Patel S, Zhang X, 2020. Evaluation of multi-shell diffusion MRI acquisition strategy on quantitative analysis using multi-compartment models. *Quant. Imaging Med. Surg* 10, 82434–82834. 10.21037/qims.2020.03.11

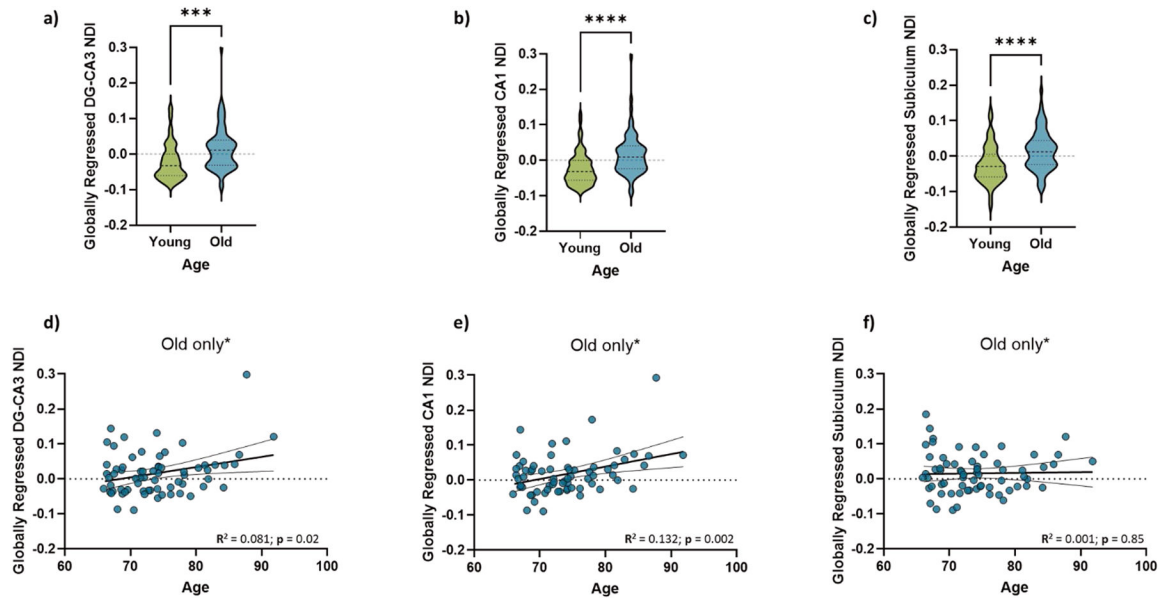
- Merluzzi AP, Dean DC, Adluru N, Suryawanshi GS, Okonkwo OC, Oh JM, Hermann BP, Sager MA, Asthana S, Zhang H, Johnson SC, Alexander AL, Bendlin BB, 2016. Age-dependent differences in brain tissue microstructure assessed with neurite orientation dispersion and density imaging. *Neurobiol. Aging* 43, 79–88. 10.1016/j.neurobiolaging.2016.03.026 [PubMed: 27255817]
- Morrison JH, Hof PR, 1997. Life and Death of Neurons in the Aging Brain. *Science* 278, 412–419. 10.1126/science.278.5337.412 [PubMed: 9334292]
- Mortimer JA, Gosche KM, Riley KP, Markesbery WR, Snowdon DA, 2004. Delayed recall, hippocampal volume and Alzheimer neuropathology: Findings from the Nun Study. *Neurology* 62, 428–432. 10.1212/01.WNL.0000106463.66966.65 [PubMed: 14872025]
- Müller MJ, Greverus D, Dellani PR, Weibrich C, Wille PR, Scheurich A, Stoeter P, Fellgiebel A, 2005. Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. *NeuroImage, Special Section: Social Cognitive Neuroscience* 28, 1033–1042. 10.1016/j.neuroimage.2005.06.029
- Nath V, Pathak SK, Schilling KG, Schneider W, Landman BA, 2020. Deep learning estimation of multi-tissue constrained spherical deconvolution with limited single shell DW-MRI, in: *Medical Imaging 2020: Image Processing*. Presented at the Medical Imaging 2020: Image Processing, SPIE, pp. 162–171. 10.1117/12.2549455
- Nazeri A, Chakravarty MM, Rotenberg DJ, Rajji TK, Rathi Y, Michailovich OV, Voineskos AN, 2015. Functional Consequences of Neurite Orientation Dispersion and Density in Humans across the Adult Lifespan. *J. Neurosci* 35, 1753–1762. 10.1523/JNEUROSCI.3979-14.2015 [PubMed: 25632148]
- Nazeri A, Mulsant BH, Rajji TK, Levesque ML, Pipitone J, Stefanik L, Shahab S, Roostaei T, Wheeler AL, Chavez S, Voineskos AN, 2017. Gray Matter Neuritic Microstructure Deficits in Schizophrenia and Bipolar Disorder. *Biol. Psychiatry, Risk Genes and the Emergence of Schizophrenia* 82, 726–736. 10.1016/j.biopsych.2016.12.005
- Nazeri A, Schifani C, Anderson JAE, Ameis SH, Voineskos AN, 2020. In Vivo Imaging of Gray Matter Microstructure in Major Psychiatric Disorders: Opportunities for Clinical Translation. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 5, 855–864. 10.1016/j.bpsc.2020.03.003 [PubMed: 32381477]
- Nobis L, Manohar SG, Smith SM, Alfaro-Almagro F, Jenkinson M, Mackay CE, Husain M, 2019. Hippocampal volume across age: Nomograms derived from over 19,700 people in UK Biobank. *NeuroImage Clin.* 23, 101904. 10.1016/j.nicl.2019.101904 [PubMed: 31254939]
- Madden DJ, Bennett IJ, Burzynska A, Potter GG, Chen N, Song AW, 2012. Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochim Biophys Acta.* 1822(3): 386–400. <https://doi.org/0.1016/j.bbadis.2011.08.003> [PubMed: 21871957]
- Papadakis NG, Xing D, Huang CL-H, Hall LD, Carpenter TA, 1999. A Comparative Study of Acquisition Schemes for Diffusion Tensor Imaging Using MRI. *J. Magn. Reson* 137, 67–82. 10.1006/jmre.1998.1673 [PubMed: 10053134]
- Pasternak O, Sochen N, Gur Y, Intrator N, Assaf Y, 2009. Free water elimination and mapping from diffusion MRI. *Magn. Reson. Med* 62, 717–730. 10.1002/mrm.22055 [PubMed: 19623619]
- Peter R, H., 1979. Synaptic density in human frontal cortex — Developmental changes and effects of aging. *Brain Res.* 163, 195–205. 10.1016/0006-8993(79)90349-4 [PubMed: 427544]
- Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, Jack CR, Jagust WJ, Shaw LM, Toga AW, Trojanowski JQ, Weiner MW, 2010. Alzheimer’s Disease Neuroimaging Initiative (ADNI). *Neurology* 74, 201–209. 10.1212/WNL.0b013e3181cb3e25 [PubMed: 20042704]
- Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline J-B, Vul E, Yarkoni T, 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci* 18, 115–126. 10.1038/nrn.2016.167 [PubMed: 28053326]
- Radhakrishnan H, Stark SM, Stark CEL, 2020. Microstructural Alterations in Hippocampal Subfields Mediate Age-Related Memory Decline in Humans. *Front. Aging Neurosci* 12, 94. 10.3389/fnagi.2020.00094 [PubMed: 32327992]
- Radhakrishnan H, Ubele MF, Krumholz SM, Boaz K, Mefford JL, Jones ED, Meacham B, Smiley J, Puskás LG, Powell DK, Norris CM, Stark CEL, Head E, 2021. Tacrolimus Protects against Age-

- Associated Microstructural Changes in the Beagle Brain. *J. Neurosci* 41, 5124–5133. 10.1523/JNEUROSCI.0361-21.2021 [PubMed: 33952632]
- Raffelt D, Tournier J-D, Rose S, Ridgway GR, Henderson R, Crozier S, Salvado O, Connelly A, 2012. Apparent Fibre Density: a novel measure for the analysis of diffusion-weighted magnetic resonance images. *NeuroImage* 59, 3976–3994. 10.1016/j.neuroimage.2011.10.045 [PubMed: 22036682]
- Raz N, Gunning-Dixon F, Head D, Rodrigue KM, Williamson A, Acker JD, 2004. Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiol. Aging* 25, 377–396. 10.1016/S0197-4580(03)00118-0 [PubMed: 15123343]
- Rey A, 1941. L'examen psychologique dans les cas d'encephalopathie traumatique. *Arch Psychol* 28, 286–340.
- Sasson E, Doniger GM, Pasternak O, Assaf Y, 2010. Structural correlates of memory performance with diffusion tensor imaging. *Neuroimage* 50, 1231–42. 10.1016/j.neuroimage.2009.12.079 [PubMed: 20045476]
- Schilling KG, Janve V, Gao Y, Stepniewska I, Landman BA, Anderson AW, 2018. Histological validation of diffusion MRI fiber orientation distributions and dispersion. *NeuroImage* 165, 200–221. 10.1016/j.neuroimage.2017.10.046 [PubMed: 29074279]
- Seabold S, Perktold J, 2010. *Statsmodels: Econometric and Statistical Modeling with Python* 5.
- Skare S, Bammer R, 2009. Jacobian weighting of distortion corrected EPI data [WWW Document]. URL <https://www.semanticscholar.org/paper/Jacobian-weighting-of-distortion-corrected-EPI-data-Skare-Bammer/7ed01244fd7c4ab273656d40b8b76fcbdc4bee7d> (accessed 10.1.21).
- Stark SM, Kirwan CB, Stark CEL, 2019. Mnemonic Similarity Task: A Tool for Assessing Hippocampal Integrity. *Trends Cogn. Sci* 23, 938–951. 10.1016/j.tics.2019.08.003 [PubMed: 31597601]
- Stark SM, Stark CEL, 2017. Age-related deficits in the mnemonic similarity task for objects and scenes. *Behav. Brain Res.* 333, 109–117. 10.1016/j.bbr.2017.06.049 [PubMed: 28673769]
- Stark SM, Stevenson R, Wu C, Rutledge S, Stark CEL, 2015. Stability of age-related deficits in the mnemonic similarity task across task variations. *Behav. Neurosci* 129, 257–268. 10.1037/bne0000055 [PubMed: 26030427]
- Stark SM, Yassa MA, Lacy JW, Stark CE, 2013. A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment. *Neuropsychologia* 51, 2442–9. 10.1016/j.neuropsychologia.2012.12.014 [PubMed: 23313292]
- Stark SM, Yassa MA, Stark CEL, 2010. Individual differences in spatial pattern separation performance associated with healthy aging in humans. *Learn. Mem* 17, 284–288. [PubMed: 20495062]
- Stejskal EO, Tanner JE, 1965. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J. Chem. Phys* 42, 288–292. 10.1063/1.1695690
- Steven AJ, Zhuo J, Melhem ER, 2014. Diffusion Kurtosis Imaging: An Emerging Technique for Evaluating the Microstructural Environment of the Brain. *Am. J. Roentgenol* 202, W26–W33. 10.2214/AJR.13.11365 [PubMed: 24370162]
- Student, 1908. The Probable Error of a Mean. *Biometrika* 6, 1–25. 10.2307/2331554
- Taki Y, Kinomura S, Sato K, Goto R, Kawashima R, Fukuda H, 2011. A longitudinal study of gray matter volume decline with age and modifying factors. *Neurobiol. Aging* 32, 907–915. 10.1016/j.neurobiolaging.2009.05.003 [PubMed: 19497638]
- Thomason ME, Thompson PM, 2011. Diffusion Imaging, White Matter, and Psychopathology. *Annu. Rev. Clin. Psychol* 7, 63–85. 10.1146/annurev-clinpsy-032210-104507 [PubMed: 21219189]
- Tournier J-D, Calamante F, Connelly A, 2013. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR Biomed.* 26, 1775–1786. 10.1002/nbm.3017 [PubMed: 24038308]
- Tournier J-D, Calamante F, Connelly A, 2012. MRtrix: Diffusion tractography in crossing fiber regions. *Int. J. Imaging Syst. Technol* 22, 53–66. 10.1002/ima.22005

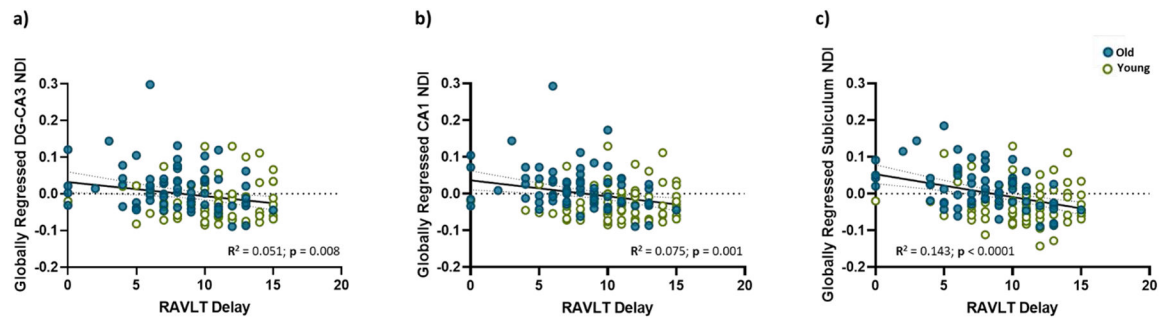
- Truong T-K, Guidon A, Song AW, 2014. Cortical Depth Dependence of the Diffusion Anisotropy in the Human Cortical Gray Matter In Vivo. *PLOS ONE* 9, e91424. 10.1371/journal.pone.0091424 [PubMed: 24608869]
- Tuch DS, 2004. Q-ball imaging. *Magn. Reson. Med* 52, 1358–1372. 10.1002/mrm.20279 [PubMed: 15562495]
- Tustison NJ, Avants BB, 2013. Explicit B-spline regularization in diffeomorphic image registration. *Front. Neuroinformatics* 7, 39. 10.3389/fninf.2013.00039
- Tustison NJ, Avants BB, Cook PA, Yuanjie Zheng, Egan A, Yushkevich PA, Gee JC, 2010. N4ITK: Improved N3 Bias Correction. *IEEE Trans. Med. Imaging* 29, 1310–1320. 10.1109/TMI.2010.2046908 [PubMed: 20378467]
- Tustison NJ, Cook PA, Klein A, Song G, Das SR, Duda JT, Kandel BM, van Strien N, Stone JR, Gee JC, Avants BB, 2014. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. *Neuroimage* 99, 166–79. 10.1016/j.neuroimage.2014.05.044 [PubMed: 24879923]
- van Haren NEM, Pol HEH, Schnack HG, Cahn W, Brans R, Carati I, Rais M, Kahn RS, 2008. Progressive Brain Volume Loss in Schizophrenia Over the Course of the Illness: Evidence of maturational abnormalities in early adulthood. *Biol. Psychiatry, Schizophrenia: From Genetics to Treatment* 63, 106–113. 10.1016/j.biopsych.2007.01.004
- Venkatesh A, Stark SM, Stark CEL, Bennett IJ, 2020. Age- and memory- related differences in hippocampal gray matter integrity are better captured by NODDI compared to single-tensor diffusion imaging. *Neurobiol. Aging* 96, 12–21. 10.1016/j.neurobiolaging.2020.08.004 [PubMed: 32905951]
- Veraart J, Fieremans E, Novikov DS, 2016. Diffusion MRI noise mapping using random matrix theory: Diffusion MRI Noise Mapping. *Magn. Reson. Med* 76, 1582–1593. 10.1002/mrm.26059 [PubMed: 26599599]
- Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B, 2013. Weighted linear least squares estimation of diffusion MRI parameters: strengths, limitations, and pitfalls. *NeuroImage* 81, 335–346. 10.1016/j.neuroimage.2013.05.028 [PubMed: 23684865]
- Westin, 1997. Geometrical diffusion measures for MRI from tensor basis analysis. *Proc ISMRM97*.
- Wisse LEM, Chételat G, Daugherty AM, de Flores R, la Joie R, Mueller SG, Stark CEL, Wang L, Yushkevich PA, Berron D, Raz N, Bakker A, Olsen RK, Carr VA, 2021. Hippocampal subfield volumetry from structural isotropic 1 mm<sup>3</sup> MRI scans: A note of caution. *Hum. Brain Mapp* 42, 539–550. 10.1002/hbm.25234 [PubMed: 33058385]
- Yassa MA, Mattfeld AT, Stark SM, Stark CEL, 2011. Age-related memory deficits linked to circuit-specific disruptions in the hippocampus. *Proc. Natl. Acad. Sci* 108, 8873–8878. 10.1073/pnas.1101567108 [PubMed: 21555581]
- Yi SY, Barnett BR, Torres-Velázquez M, Zhang Y, Hurley SA, Rowley PA, Hernando D, Yu J-PJ, 2019. Detecting Microglial Density With Quantitative Multi-Compartment Diffusion MRI. *Front. Neurosci* 13. 10.3389/fnins.2019.00081
- Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC, 2012. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 61, 1000–1016. 10.1016/j.neuroimage.2012.03.072 [PubMed: 22484410]

**Highlights:**

- Diffusion metrics from multi-shell sequences are reliable and sensitive to gray matter microstructure.
- NODDI complements traditional tensor and volume measures when predicting age and cognition.
- NODDI metrics greatly improve the predictive power of models estimating individual differences.

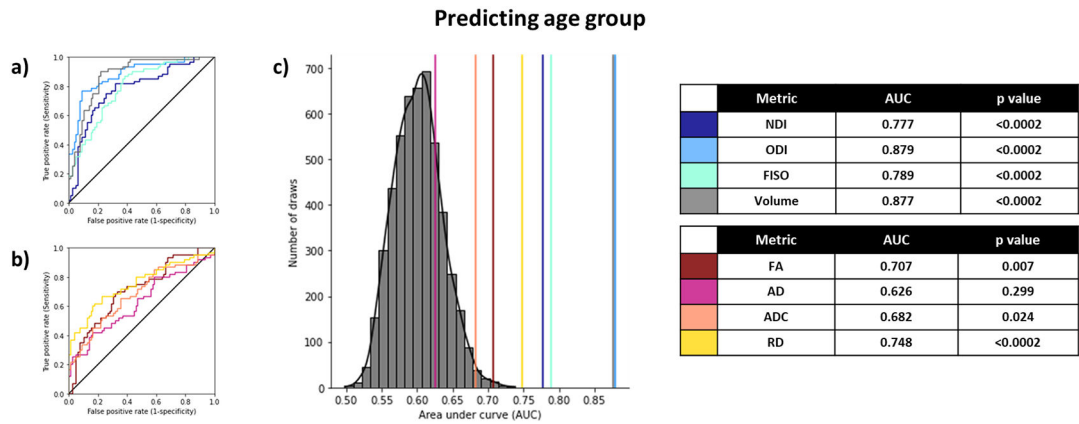
**Figure 1:**

(a-c) Hippocampal NDI is significantly greater in older adults as compared to young adults for all three hippocampal subfields. Moreover, the NDI of the DG/CA<sub>3</sub> (d) and CA<sub>1</sub> (e) linearly increases with age within the older subpopulation, while that of the subiculum (f) does not.



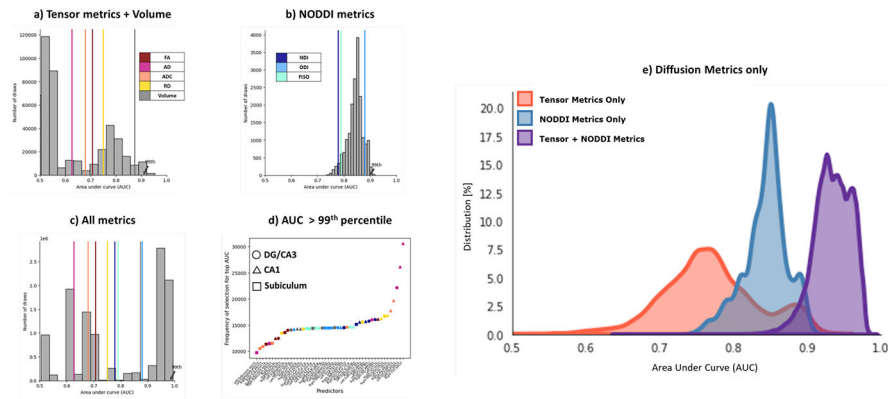
**Figure 2:**

Globally regressed NDI is negatively associated with RAVLT Delay in all three hippocampal subfields, in both the younger and older groups, with the subiculum NDI having the strongest relationship with RAVLT performance. Green dots indicate the younger adults (18 – 29 years), while blue dots represent the older adults (65 – 92 years).

**Figure 3:**

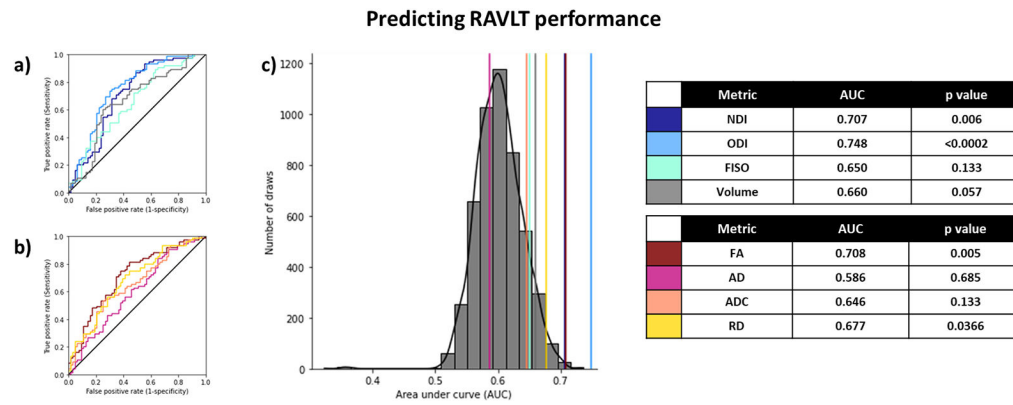
Using diffusion and structural metrics of the hippocampal subfields to predict age group. All metrics can successfully predict age group, with ODI, volume, and RD being the best predictors and AD being the worst predictor (a-b). The histogram represents the null distribution (random permutation of old/young labels), and the colored lines represent the AUC of each of the metrics from all 6 subfields (c). After Holm-Sidak correction for multiple comparisons, p-values < 0.0253 are considered statistically significant.



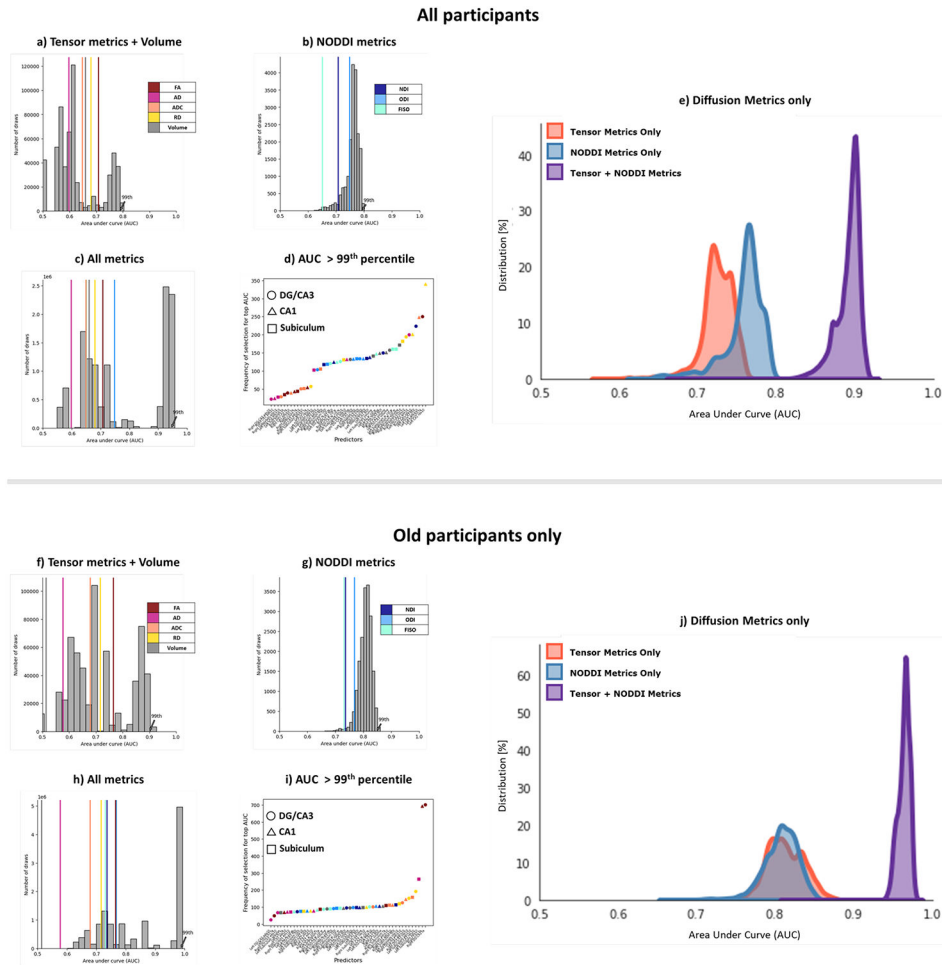


**Figure 4:**

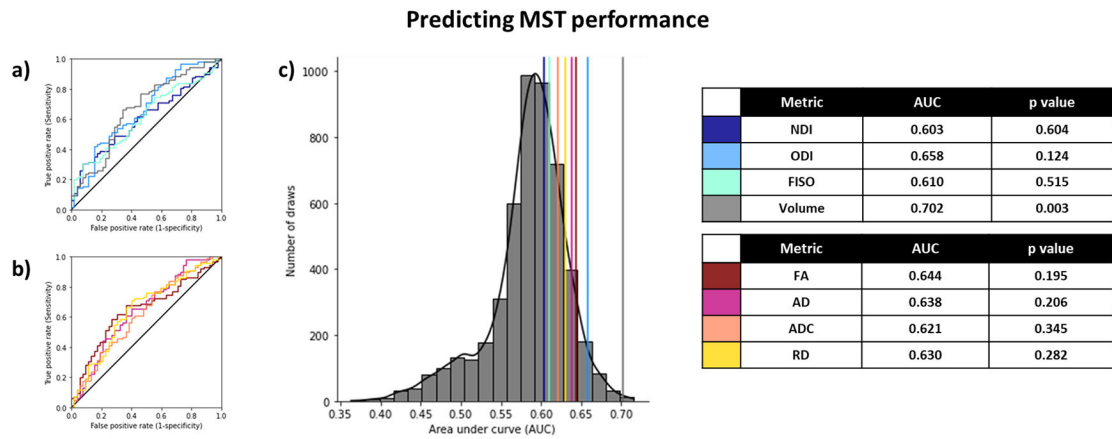
A combination of all metrics results in the highest AUC peaks, compared to just the tensor metrics and volume or the NODDI metrics. Histograms represent the AUCs generated following an exhaustive n-choose-6 analysis to determine how well various combinations of 6 Region x Metric regressors could model age group. (a) Analysis of only the traditional tensor and volume metrics; (b) Analysis of only the NODDI metrics; and (c) Analysis of all metrics combined. The colored lines represent the AUCs of each metric from all 6 subfields. (d) Plot of the frequency of selection of a given Region x Metric combination in the top AUCs (above 99<sup>th</sup> percentile) in the All metrics analysis. (e) Histogram of distribution of AUCs with only combinations of the diffusion metrics (no volume included).



**Figure 5:** Using diffusion and structural metrics of the hippocampal subfields to *predict RAVLT performance* (<9). ODI, NDI, and FA were the best predictors of RAVLT performance. The histogram (c) represents the null distribution (randomly permuting the high/low RAVLT labels), and the colored lines represent the AUC of each of the metrics from all 6 subfields. After Holm-Sidak correction for multiple comparisons, p-values < 0.0253 are considered statistically significant.

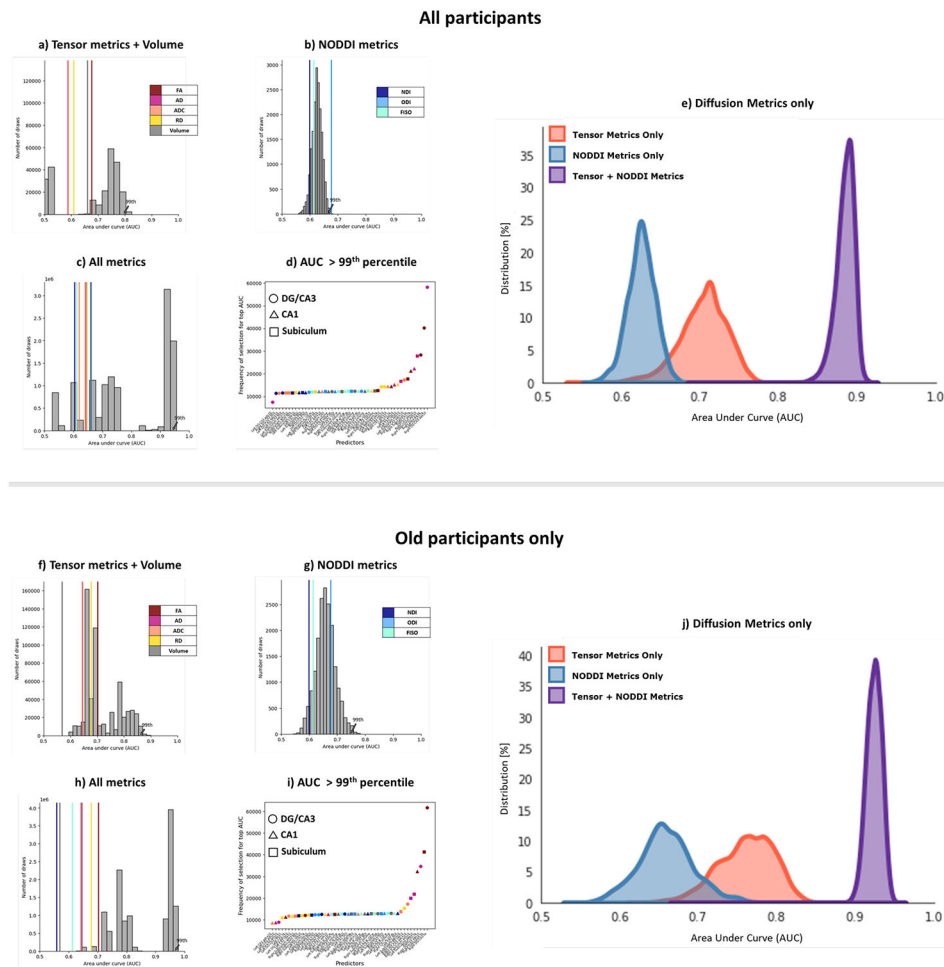


**Figure 6:** Combining NODDI and tensor metrics results in a significantly higher AUC after selecting for best-performing features in predicting high/low RAVLT performance, as compared to just traditional measures or just NODDI measures both when examining the full study population (a-e) and when examining only the older subpopulation (f-j). (a, f) Analysis of only the traditional tensor and volume metrics; (b, g) Analysis of only the NODDI metrics; and (c, h) Analysis of all metrics combined. Histograms represent the distribution of AUCs generated following an exhaustive n-choose-6 analysis to determine how well various combinations of 6 Region x Metric regressors could model high/low RAVLT group. The colored lines represent the AUCs of each metric from all 6 subfields. (d, i) Plot of the frequency of selection of a given Region x Metric combination in the top AUCs (above 99<sup>th</sup> percentile) in the All metrics analysis. e, j) Histogram of distribution of AUCs with only combinations of the diffusion metrics (no volume included).



**Figure 7:**

Using diffusion and structural metrics of the hippocampal subfields to *predict LDI* ( $<1.25$ ). Only volume could predict MST performance slightly better than chance. The histogram represents the null distribution (randomly permuting the high/low labels), and the colored lines represent the AUC of each of the metrics from all 6 subfields. After Holm-Sidak correction for multiple comparisons, p-values  $< 0.0253$  are considered statistically significant.



**Figure 8:** Combining NODDI and tensor metrics results in a significantly higher AUC after selecting for best-performing features in predicting high/low MST performance, as compared to just traditional measures or just NODDI measures, in both the full study population (a-e), as well as in the older subpopulation (f-j). (a, f) Analysis of only the traditional tensor and volume metrics; (b, g) Analysis of only the NODDI metrics; and (c, h) Analysis of all metrics combined. Histograms represent the distribution of AUCs generated following an exhaustive n-choose-6 analysis to determine how well various combinations of 6 Region x Metric regressors could model high/low RAVLT group. The colored lines represent the AUCs of each metric from all 6 subfields. (d, i) Plot of the frequency of selection of a given Region x Metric combination in the top AUCs (above 99<sup>th</sup> percentile) in the All metrics analysis. e, j) Histogram of distribution of AUCs with only combinations of the diffusion metrics (no volume included).

**Table 3.1:**

Demographics and cognitive test scores. \*Rows in blue indicate  $p < 0.05$  in a t-test between young and old groups.

	Young	Old
N	79	75
Age*	20.41 ± 1.89	73.56 ± 6.26
Female (%)	58.2%	60%
Education	13.98 ± 2.26	14.56 ± 2.30
RAVLT Immediate*	11.29 ± 2.50	8.30 ± 3.02
RAVLT Delay*	10.68 ± 3.00	7.67 ± 3.40
MST LDI*	0.83 ± 0.50	1.12 ± 0.57
MST REC	2.67 ± 0.94	2.97 ± 0.72

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript