Received: 28 October 2022 | Revised: 15 February 2023 | Accepted: 24 February 2023

DOI: 10.1002/hbm.26269

RESEARCH ARTICLE

WILEY

The use of hippocampal grading as a biomarker for preclinical and prodromal Alzheimer's disease

Cassandra Morrison[1] | Mahsa Dadar[2,3] | Neda Shafiee[1] | D. Louis Collins[1] | For Alzheimer's Disease Neuroimaging Initiative

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada

[2]Department of Psychiatry, McGill University, Montreal, Quebec, Canada

[3]McGill Univeristy, Douglas Mental Health University Institute, Montreal, Quebec, Canada

Correspondence
Cassandra Morrison, Montreal Neurological Institute, 3801 University Street, Montreal QC, H3A 2B4, Canada.
Email: cassandra.morrison@mail.mcgill.ca

Funding information
Canadian Institutes of Health Research, Grant/Award Numbers: MFE-176608, PJT -165921

Abstract

Hippocampal changes are associated with increased age and cognitive decline due to mild cognitive impairment (MCI) and Alzheimer's disease (AD). These associations are often observed only in the later stages of decline. This study examined if hippocampal grading, a method measuring local morphological similarity of the hippocampus to cognitively normal controls (NCs) and AD participants, is associated with cognition in NCs, subjective cognitive decline (SCD), early (eMCI), late (lMCI), and AD. A total of 1620 Alzheimer's Disease Neuroimaging Initiative participants were examined (495 NC, 262 eMCI, 545 lMCI, and 318 AD) because they had baseline MRIs and Alzheimer's disease Assessment Scale (ADAS-13) and Clinical Dementia Rating—Sum of Boxes (CDR-SB) scores. In a sub-analysis, NCs with episodic memory scores (as measured by Rey Auditory Verbal Learning Test, RAVLT) were divided into those with subjective cognitive decline (SCD+; 103) and those without (SCD−; 390). Linear regressions evaluated the influence of hippocampal grading on cognition in preclinical and prodromal AD. Lower global cognition, as measured by increased ADAS-13, was associated with hippocampal grading: NC ($p < .001$), eMCI ($p < .05$), lMCI ($p < .05$), and AD ($p = .01$). Lower global cognition as measured increased CDR-SB was associated with hippocampal grading in lMCI ($p < .05$) and AD ($p < .001$). Lower RAVLT performance was associated with hippocampal grading in SCD− ($p < .05$) and SCD+ ($p < .05$). These findings suggest that hippocampal grading is associated with global cognition in NC, eMCI, lMCI, and AD. Early changes in episodic memory during preclinical AD are associated with changes in hippocampal grading. Hippocampal grading may be sensitive to progressive changes early in the disease course.

KEYWORDS
Alzheimer's disease, cognitive decline, cognitive functioning, hippocampal grading, mild cognitive impairment, older adults, subjective cognitive decline

1 | INTRODUCTION

Increased age can be associated with cognitive decline, ranging from healthy aging to mild cognitive impairment (MCI) and, in some cases, Alzheimer's disease (AD). MCI is characterized by declines in cognitive functioning that are not severe enough to impair daily activities (Petersen, 2004), while AD is characterized by progressive declines in cognitive functioning that are severe enough to impair daily activities

(Alzheimer's Association, 2021). AD is typically defined by its underlying pathological biomarkers, β-amyloid (A), pathological tau (T), and neurodegeneration (N) [AT(N)] (Jack et al., 2018). It is well-established that AD-related pathology begins occurring before the clinical symptom presentation (i.e., measurable cognitive deficits) (Reisberg et al., 2010).

Hippocampal volume is often used as a biomarker of neurodegeneration in AD because it is affected early in the disease course (Fjell et al., 2014), with increased rates of atrophy in people with AD compared to cognitively normal older adults (Jack et al., 1998). Furthermore, 80% of people with MCI classified as A+T+(N+) progressed to dementia when the (N+) was defined using hippocampal volume (Strikwerda-Brown et al., 2022). These findings suggest that hippocampal neurodegeneration is associated with AD and may be predictive of cognitive decline progression. Given the relationship between hippocampal atrophy and AD and the early development of AD-related biomarkers, examining the hippocampus and its relationship to cognitive change in healthy older adults is essential for a detailed understanding of the progressive nature of early-stage AD.

Neurodegeneration is associated with observable cognitive deficits in both aging (Bettio et al., 2017) and AD (Jack et al., 2018). Declines in hippocampal volume are suggested to account for episodic memory declines in aging and AD (Köhncke et al., 2021). Strong associations between decreased hippocampal volume and poor episodic memory functioning support this conclusion (Gorbach et al., 2017; Persson et al., 2012; Schneider et al., 2019). In addition to episodic memory, studies have also found a relationship between the hippocampus and global cognitive functioning (e.g., Dawe et al., 2020). When separating groups based on diagnostic status, global cognition has been correlated with decreased hippocampal volume in only AD [not MCI or cognitively normal controls (NCs), Vipin et al., 2018], and in both AD and MCI (but not NCs, Peng et al., 2015). Most of these studies used the mini-mental status examination (MMSE) as a measure of global cognition (Peng et al., 2015; Vipin et al., 2018), which may limit the sensitivity of these associations.

Different measures of the hippocampal neurodegeneration could improve the observed relationship between the hippocampus and general cognitive functioning. Hippocampal grading, measured by the Scoring by Nonlocal Image Patch Estimator (SNIPE), has been shown to surpass hippocampal volume in predictive power (Coupé et al., 2015; Coupé, Eskildsen, Manjón, Fonov, Pruessner, et al., 2012). SNIPE computes the similarity of every voxel in the hippocampus of each person to a large library of manually segmented MRI datasets from both healthy cognitively intact older adults and an equal number of patients with AD. This procedure compares the local neighbourhood patch surrounding the voxel to corresponding neighbourhood patches for each image volume in the library. The SNIPE score is the average of similarity-weighted labels (i.e., −1 for AD and +1 for healthy control) from the library of subjects. When the average SNIPE score is positive, the structure is more similar to healthy control, and when negative, the structure is more similar to AD.

The SNIPE method has shown to differentiate diagnostic groups (NC vs. MCI, NC vs. AD, and MCI vs. AD) with higher classification accuracy than volumetric measures (Morrison et al., 2022). Higher accuracies were obtained for SNIPE grading compared to volumetric scores calculated by both SNIPE methods and Freesurfer across all groups. Using this method, researchers have also observed that SNIPE-based grading biomarkers are more relevant for cognitive decline prediction and conversion from MCI to AD than hippocampal volume measures (Coupé, Eskildsen, Manjón, Fonov, Pruessner, et al., 2012). SNIPE has a classification accuracy of 89%–93% at distinguishing people with AD from NC individuals in the ADNI cohort (Coupé, Eskildsen, Manjón, Fonov, & Collins, 2012; Morrison et al., 2022). SNIPE could also predict in cognitively normal older adults, who would progress to AD dementia within a 12-year follow-up period with 72.5% accuracy (Coupe et al., 2015). Coupe et al. (2012) also observed that in normal controls and people with AD, HC grading had a stronger correlation with global cognition, as measured by the MMSE ($R = 0.75$) than hippocampal volume ($R = 0.58$). Associations between SNIPE hippocampal grading and cognitive declines in normal aging and early MCI have yet to be determined.

Given that the MMSE has limited value in differentiating MCI from NC (Mitchell, 2009), other neuropsychological measures with a wide range in scores (e.g., Alzheimer's Disease Assessment Scale–13, ADAS-13) and measures cognitive and functional status (e.g., Clinical Dementia Rating–sum of boxes, CDR-SB) may be more sensitive to cognitive changes early in the disease process (prodromal or preclinical AD). For example, the CDR-SB has been shown to have good prediction value for people without dementia who convert to dementia (Tzeng et al., 2022). These additional tests may, thus, offer stronger associations with AD-related hippocampal changes than the MMSE. Previous research has also shown relationships between the hippocampus and episodic memory in preclinical AD (subjective cognitive decline, SCD) (Caillaud et al., 2020). The MMSE is not designed to measure episodic memory and therefore it may not be sensitive to the early cognitive changes observed in SCD. For that reason, we also examine whether episodic memory, as measured by Rey's Auditory Verbal Learning Test (RAVLT), is associated with hippocampal change in those with SCD. Because hippocampal grading, as measured by SNIPE, has better predictive power than hippocampal volume, we predict that a strong association between SNIPE measures and sensitive measures of cognitive decline will be observed. Our previous paper also found that SNIPE grading had higher classification accuracy (for NC:AD, NC:eMCI, eMCI:lMCI, and lMCI:AD) than both SNIPE volume and Freesurfer volume measurements of the hippocampus (Morrison et al., 2022).

In this study, we aim to investigate the relationship between SNIPE grading scores and cognition. The current paper was designed to determine whether measures of general cognitive functioning (i.e., the ADAS-13 and CDR-SB) would have a strong association with hippocampal grading in normal controls, people with MCI, and people with AD and if performance on episodic memory tests would have associations with hippocampal grading in normal controls with and without SCD. While Coupe et al. determined that SNIPE could classify between NC and people with AD and predict which NC may progress to AD, it remains unknown if SNIPE grading is associated with

progressive changes in cognitive functioning at different stages of decline (Coupe et al., 2015; Coupé, Eskildsen, Manjón, Fonov, Pruessner, et al., 2012; Coupé, Eskildsen, Manjón, Fonov, & Collins, 2012).

The goal of this article was to characterize the relationship between SNIPE and cognition as measured by ADAS-13 or CDR-SB early in the disease. Although both ADAS-13 and CDR-SB are sensitive to cognitive changes in MCI and AD, the ADAS-13 has a larger range of scores (0–84) than the CDR-SB (0–18). We expect to see strong associations between the ADAS-13 in NC, MCI, and AD because of the large range in ADAS-13. In contrast, for the CDR-SB we expect to see strong associations only in the later stages of cognitive decline because NCs have a small range of scores. Additionally, we expect that RAVLT scores for measuring episodic memory will be sensitive to cognitive changes at the preclinical AD stage (i.e., SCD). These findings would not only have implications for future development and improvement of methods to measure neurodegeneration, but also for the early detection and characterization of incipient AD.

2 | METHODS

2.1 | Alzheimer's Disease Neuroimaging Initiative

Data used in the preparation of this article were obtained from the publicly available Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). Participants were selected from ADNI-1, ADNI-2, and the ADNI-GO cohorts. The study received ethical approval from the review boards of all participating institutions. Written informed consent was obtained from participants or their study partner.

2.2 | Participants

Full participant inclusion/exclusion is available online at www.adni-info.org. Briefly, all participants were between 55 and 90 years old at the time of recruitment and no evidence of depression as measured by the Geriatric Depression Scale. Healthy control participants had no evidence of memory decline on the Logical Memory II subscale from the Wechsler Memory Scale and no evidence of cognitive decline on either the Mini Mental Status Examination (MMSE) or Clinical Dementia Rating (with a score of 0 required for memory box). Early MCI and lMCI had to score between 24 and 30 on the MMSE, 0.5 on the CDR with a memory box score of at least 0.5, and had abnormal scores on the Logical Memory II test. Early MCI was differentiated from lMCI by degree of memory impairment on the Logical Memory II test; eMCI participants were characterized as memory impairment that was intermediate between normal controls and lMCI. Using the Logical Memory II test, Early MCI was assigned to participants who obtained a score of 9–11 (for 16+ years of education), a score of 5–9 (for 8–15 years of education), or a score of 3–6 (for 0–7 years of education). Late MCI was assigned to participants who obtained a score of ≤8 (for 16+ years of education), a score of ≤4 (for 8–15 years of education), or a score of ≤2 (for 0–7 years of education). AD participants had to show abnormal memory function on the Logical Memory II test, an MMSE score between 20 and 26, a CDR-SB 0.5 or 1.0 and probable AD according to the NINCDS/ADRDA criteria.

A total of 1634 participants from three ADNI cohorts had MRI baseline scans and were thus included (ADNI-1, 787 participants; ADNI-2, 759 participants; ADNI-GO 88 participants). Inclusion criteria also included availability of baseline CDR-SB and ADAS-13 scores. Fourteen participants were excluded for not having baseline ADAS-13 scores. A total of 1620 participants were included for our study. Of these 1620 participants, 495 were cognitively normal older adults (NC), 262 were early MCI (eMCI), 545 were late MCI (lMCI), and 318 had an AD diagnosis.

A sub-analysis was completed on the NCs who had baseline RAVLT scores ($N = 493$). The RAVLT scores analyzed were the summary scores that measure learning (RAVLT immediate) and delayed memory (RAVLT percent forgetting). These two scores were chosen because they are both essential aspects of AD (Moradi et al., 2017). RAVLT immediate is the sum of scores from the 5 first trials of the test (Trials 1–5) and RAVLT percent forgetting is the score of Trial 5 minus the score of Trial 1 divided by the score of Trial 5. This NC group was subdivided into those with and without subjective cognitive decline using cognitive change index (CCI) scores. Participants were considered SCD if they self-reported significant memory concern, quantified by a score of ≥16 on the first 12 items (representing memory changes) on the CCI. This threshold was selected based on previous research by Saykin et al. (2006) and because it is used by ADNI as a criterion to identify participants with significant memory concern (Risacher et al., 2015). A total of 103 NCs had SCD (SCD+) and 390 did not have SCD (SCD−). Table 1 summarizes demographic information for all participants.

2.3 | Structural MRI acquisition and processing

All participants were imaged using a 3T scanner with T1-weighted imaging parameters (see http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/ for the detailed MRI acquisition protocol). Baseline scans were downloaded from the ADNI public website.

Raw T1w scans for each participant were pre-processed through our standard pipeline including noise reduction (Coupé et al., 2008), intensity inhomogeneity correction (Sled et al., 1998) and intensity normalization into range [0–100]. The preprocessed images were then both linearly (9 parameters, 3 translation, 3 rotation, and 3 scaling) (Dadar et al., 2018) and nonlinearly (1 mm^3 grid) (Avants et al., 2008) registered to the MNI-ICBM152-2009c average (Fonov et al., 2011).

Full sample	NC ($n = 495$)	eMCI ($n = 262$)	lMCI ($n = 545$)	AD ($n = 318$)
Age	74.34 ± 5.76	70.84 ± 7.39	73.97 ± 7.57	75.03 ± 7.70
Education	16.34 ± 2.72	15.97 ± 2.64	15.89 ± 2.91	15.17 ± 3.02
Female Sex	260 (53%)	115 (44%)	209 (38%)	143 (45%)
ADAS-13	9.23 ± 4.32	12.57 ± 5.56	18.72 ± 6.52	29.95 ± 7.91
CDR-SB	0.04 ± 0.14	1.29 ± 0.76	1.65 ± 0.93	4.43 ± 1.64
Subanalysis		**NC/SCD− ($n = 390$)**	**SCD+ ($n = 103$)**	
Age		74.84 ± 5.73	72.40 ± 5.52	
Education		16.22 ± 2.75	16.77 ± 2.56	
Female sex		199 (51%)	60 (58%)	
RAVLT – Immediate		5.78 ± 2.34	6.06 ± 2.21	
RAVLT – Percent Forgetting		35.69 ± 27.47	37.31 ± 29.03	

TABLE 1 Demographic information for cognitively normal, early and late MCI, and AD participants.

Note: Values are expressed as mean ± standard deviation, or number (percentage %). Female sex is represented as total number of sample and percentage of sample.
Abbreviations: AD, Alzheimer's disease; ADAS-13, Alzheimer's Disease Assessment Scale–Cognitive Subscale; CDRSB, Clinical Dementia Rating Scale – Sum of Boxes; eMCI, early mild cognitive impairment; lMCI, late mild cognitive impairment; NC, cognitively normal controls; NC/SCD−, normal controls without subjective cognitive decline; RAVLT, Rey Auditory Verbal Learning Test; SCD+, subjective cognitive decline.

The quality of the linear and nonlinear registrations was visually verified by an experienced rater (author Mahsa Dadar), blinded to diagnostic group. Only seven datasets did not pass this quality control step and were discarded.

2.4 | SNIPE

Scoring by Nonlocal Image Patch Estimator (SNIPE) was used to measure the extent of AD-related change in the hippocampus using the linearly registered preprocessed T1-weighted images (Coupé, Eskildsen, Manjón, Fonov, Pruessner, et al., 2012; Coupé, Eskildsen, Manjón, Fonov, & Collins, 2012). In short, this technique uses a set of MRI volumes with manually segmented hippocampi as training library from both healthy aging subjects (CN) and patients with dementia due to AD. For each voxel from the subject under study that falls within a bounding box containing the medial temporal lobe region, a 3D $7 \times 7 \times 7$ patch centered around that voxel is compared with corresponding patches from the $N = 100$ MRI volumes (50 CN and 50 AD) in the training library. An intensity-based similarity metric (or "weight") between the patch under study and the training patch was then computed. These estimated weights were used to perform grading of the hippocampus based on the clinical label (CN vs. AD) of the training subjects:

$$g(x_i) = \frac{\sum_{s=1}^{N} \sum_{j \in \Omega} w(x_i, x_{s_j}).P_s}{\sum_{s=1}^{N} \sum_{j \in \Omega} w(x_i, x_{s_j})}$$

where x_i is the target voxel, and $g(x_i)$ is the corresponding grading value, and Ω is the search area. $w(x_i, x_{s_j})$ Shows the similarity metric between surrounding patches of target voxel i and voxel j from training subject s. P_s is the clinical label of the training subject: we set it to -1 for AD patients and $+1$ for normal healthy subjects. This means that when a patch resembles CN anatomical characteristics more than AD, the grading score will be positive, conversely, if the patch is more similar to AD anatomy, the grading score is negative. The final SNIPE hippocampal grading score is an average of all the voxels within this structure in each hemisphere.

In this method, volumes are calculated by counting voxels in pseudo-Talairach stereotaxic space (ICBM152 template), to avoid bias toward subject's head size. SNIPE segmentations were visually verified by an experience rater (author Neda Shafiee) The SNIPE procedure used has been previously described in detail (Dadar, Gee, et al., 2020).

2.5 | Statistical analysis

Analyses were performed using "R" software version 4.0.5. *t*-Tests and χ^2 analyses were completed on demographic information. Linear regression models were conducted to determine whether hippocampal grading was associated with cognitive scores (CDR-SB and ADAS-13). The model examined the association between *CognitiveScore* (ADAS-13 or CDR-SB) and Hippocampal *Grading* (Right and Left). Diagnosis was the categorical variable of interest, indicated by NC, eMCI, lMCI, or AD status with NC serving as the baseline, *Grading: Diagnosis* denotes an interaction term between Grading and Diagnosis, reflecting differences in the slope of Grading between the diagnostic groups. The models also included age, sex, and years of education as covariates, with the regression centered on SNIPE score = 0.0 as follows:

$$CognitiveScore \sim Grading\_HC + Diagnosis + Grading : Diagnosis + Age + Sex + Education \quad (1)$$

As an additional analysis to demonstrate early sensitivity of SNIPE we examined the association between HC grading and cognition in SCD, the earliest stage of preclinical AD. Previous research has observed an association between HC volume and episodic memory in those with SCD (Caillaud et al., 2020). Thus, in this additional analysis we opted to only examine episodic memory using RAVLT scores. The same linear regression model (1) was used to determine whether hippocampal grading was associated *CognitiveScore* (RAVLT).

Correction of multiple comparisons was completed using false discovery rate (FDR); *p*-values are reported as raw values with significance determined by FDR correction marked in Table 2.

3 | RESULTS

Table 1 presents the demographic information for each group. There was no difference in age between the NC and lMCI ($t = 0.9$, $p = .4$) or between NC and AD ($t = 1.36$, $p = .2$), but the eMCI group were 3.5 years younger that NC ($t = 6.67$, $p < .001$). There was no difference in education between NC and eMCI ($t = 1.83$, $p = .07$), but NC had higher education than lMCI ($t = 2.58$, $p = .01$) and AD ($t = 5.64$, $p < .001$). As expected, the average ADAS-13 score increased from NC to AD. Differences in ADAS-13 scores were observed between each successive stage of decline, NC < eMCI < lMCI < AD (NC:eMCI, $t = -8.45$, $p < .001$; eMCI:lMCI, $t = -13.89$, $p < .001$; and lMCI:AD $t = -21.43$, $p < .001$). Similarly, CDR-SB increased with NC < eMCI < lMCI < AD with statistically significant differences between each successive group (NC:eMCI, $t = -26.47$, $p < .001$; eMCI:lMCI, $t = -5.88$, $p < .001$; and lMCI:AD $t = -27.72$, $p < .001$). In the subanalysis, SCD− were younger than the SCD+ group ($t = -3.96$, $p < .001$), but had no difference in education or in any of the RAVLT scores.

3.1 | ADAS-13 and hippocampal grading

Table 2 summarizes the results of the linear regression models for both the ADAS-13 and CDR-SB analyses. Figure 1 shows a scatterplot of individual cognitive scores and hippocampal grading values for both the ADAS-13 and CDR-SB.

For the left hippocampus (lHC) analysis, the overall effect of lHC grading on ADAS-13 scores in the NC group was significant ($t = -3.95$, $p < .001$), demonstrating that decreases in lHC grading were associated with increases in ADAS-13 scores. In addition, all patient groups had greater intercepts for ADAS-13 than the NCs, and ADAS-13 scores progressively increased (i.e., lower performance) from NC to eMCI (4.19 points more than NC, $t = 5.83$, $p < .001$), lMCI (8.13 points more than NC, $t = 15.09$, $p < .001$), and AD (17.61 points more than NC, $t = 28.80$, $p < .001$) at the model center (where SNIPE grading = 0.0). Furthermore, the interaction between lHC grading and ADAS-13 was significant for eMCI ($t = -2.22$, $p = .03$) and lMCI ($t = -2.74$, $p = .006$), and marginally significant for AD ($t = -1.85$, $p = .06$); that is, the slopes of changes in ADAS-13 scores associated with changes in lHC grading were significantly steeper in all disease cohorts in comparison with the NCs. The slope for eMCI and lMCI was almost twice that for NC. When examining the covariates, age was not associated with change in ADAS-13 scores ($t = -1.11$, $p = .27$). Male sex was associated with almost 1 point increase in ADAS-13 scores ($t = 2.98$, $p = .003$), whereas increased education was associated with slightly lower (−0.21 points) ADAS-13 scores ($t = -4.06$, $p < .001$). Overall, 65% of the variance in ADAS-13 scores (adjusted $R^2 = 0.65$) can be explained by the variables included in this model. Note that the same model *without grading* has an adjusted $R^2 = 0.61$, indicating that grading explains an additional 4% of variance.

For the right hippocampus (rHC) analysis, the overall effect of rHC grading on ADAS-13 scores in the NC group was significant ($t = -4.28$, $p < .001$), demonstrating that increases in ADAS-13 scores were associated with decreases in rHC grading. In addition, all patient groups had greater intercepts for ADAS-13 than the NCs, and ADAS-13 scores progressively increased (i.e., lower performance) from NC to eMCI (4.10 points more than NC, $t = 5.35$, $p < .001$), lMCI (7.78 points more than NC, $t = 13.45$, $p < .001$), and AD (16.65 points more than NC, $t = 25.49$, $p < .001$) at the model center (where SNIPE grading = 0.0). Furthermore, the interaction between rHC grading and ADAS-13 was significant for eMCI ($t = -1.96$, $p = .05$), to lMCI ($t = -2.23$, $p = .03$), and to AD ($t = -2.51$, $p = .01$); i.e., the slopes of changes in ADAS-13 scores associated with changes in rHC grading were significantly steeper in all disease cohorts in comparison with the NCs. When examining the covariates, increased age was associated with slight increases in ADAS-13 scores (0.05 points, $t = 2.02$, $p = .04$). Male sex was associated with almost 1 point increase in ADAS-13 scores ($t = 3.00$, $p = .003$), whereas increased education was associated with slightly lower ADAS-13 scores (−0.24 points, $t = -4.80$, $p < .001$). Overall, 66% of the variance in ADAS-13 scores (adjusted $R^2 = 0.66$) can be explained by the variables included in this model. Note that this model explains 5% more of the variance the same model *without grading*.

3.2 | CDR-SB and hippocampal grading

For the left hippocampus model, the effect of lHC grading on CDR-SB scores was not significant ($t = -0.82$, $p = .41$) for the NC group. All patient groups had greater intercepts for CDR-SB than the NC group at model center, and CDR-SB scores progressively increased (i.e., lower performance) from NC to eMCI (1.28 points more than NC, $t = 10.96$, $p < .001$), to lMCI (1.58 points more than NC, $t = 18.07$, $p < .001$), and to AD (4.08 points more than NC, $t = 40.90$, $p < .001$). The interaction between lHC grading and CDR-SB was not significant for eMCI ($t = -0.87$, $p = .38$), but was significant for lMCI ($t = -2.43$, $p = .02$) and AD ($t = -5.28$, $p < .001$); that is, the slopes of changes in CDR-SB scores associated with changes in HC grading were

TABLE 2 Linear regression model results showing association between grading and global cognition.

	ADAS-13: Left HC	ADAS-13: Right HC	ADAS-13: Average HC	CDR-SB: Left HC	CDR-SB: Right HC	CDR-SB: Average HC
Intercept	**ß = 15.69**	**ß = 18.17**	**ß = 19.66**	**ß = 1.06**	**ß = 1.14**	**ß = 1.32**
	t = 7.57	**t = 8.73**	**t = 9.29**	**t = 3.12**	**t = 3.31**	**t = 3.77**
	p < .001*	**p < .001***	**p < .001***	**p = .002***	**p < .001***	**p < .001***
Grading	**ß = −6.16**	**ß = −7.04**	**ß = −8.58**	ß = −0.21	ß = −0.21	ß = −0.31
	t = −3.95	**t = −4.28**	**t = −4.89**	t = −0.82	t = −0.76	t = −1.06
	p < .001*	**p < .001***	**p < .001***	p = .41	p = .45	p = .29
Age	ß = −0.03	**ß = −0.05**	**ß = −0.07**	**ß = −0.01**	**ß = −0.01**	**ß = −0.01**
	t = −1.11	**t = −2.02**	**t = −2.83**	**t = −2.49**	**t = −2.53**	**t = −3.16**
	p = .27	**p = .04**	**p = .005***	**p = .01***	**p = .01***	**p = .002***
Male sex	**ß = 0.88**	**ß = 0.87**	**ß = 0.82**	ß = −0.03	ß = −0.03	ß = −0.04
	t = 2.98	**t = 3.00**	**t = 2.83**	t = −0.71	t = −0.63	t = −0.77
	p = .003*	**p = .003***	**p = .005***	p = .48	p = .53	p = .44
Education	**ß = −0.21**	**ß = −0.24**	**ß = −0.23**	ß = −0.02	**ß = −0.02**	**ß = −0.02**
	t = −4.06	**t = −4.80**	**t = −4.49**	t = −1.85	**t = −2.36**	**t = −2.09**
	p < .001*	**p < .001***	**p < .001***	p = .06	**p = .02***	**p = .037**
eMCI	**ß = 4.19**	**ß = 4.10**	**ß = 4.05**	**ß = 1.28**	**ß = 1.30**	**ß = 1.30**
	t = 5.83	**t = 5.35**	**t = 5.24**	**t = 10.96**	**t = 10.25**	**t = 10.09**
	p < .001*	**p < .001***	**p < .001***	**p < .001***	**p < .001***	**p < .001***
lMCI	**ß = 8.13**	**ß = 7.78**	**ß = 7.51**	**ß = 1.58**	**ß = 1.58**	**ß = 1.57**
	t = 15.09	**t = 13.45**	**t = 12.81**	**t = 18.07**	**t = 16.50**	**t = 16.13**
	p < .001*	**p < .001***	**p < .001***	**p < .001***	**p < .001***	**p < .001***
AD	**ß = 17.61**	**ß = 16.65**	**ß = 16.29**	**ß = 4.08**	**ß = 4.06**	**ß = 3.99**
	t = 28.80	**t = 25.49**	**t = 24.53**	**t = 40.96**	**t = 37.48**	**t = 33.30**
	p < .001*	**p < .001***	**p < .001***	**p < .001***	**p < .001***	**p < .001***
*Grading*eMCI*	**ß = −5.06**	**ß = −4.54**	**ß = −4.92**	ß = −0.32	ß = −0.36	ß = −0.37
	t = −2.22	**t = −1.96**	**t = −2.20**	t = −0.87	t = −0.94	t = −0.92
	p = .027*	**p = .05**	**p = .033***	p = .38	p = .35	p = .36
*Grading*lMCI*	**ß = −5.04**	**ß = −4.12**	**ß = −4.55**	**ß = −0.73**	**ß = −0.66**	**ß = −0.76**
	t = −2.74	**t = −2.23**	**t = −2.29**	**t = −2.43**	**t = −2.22**	**t = −2.31**
	p = .006*	**p = .03***	**p = .02***	**p = .02***	**p = .03***	**p = .02***
*Grading*AD*	ß = −3.94	**ß = −5.31**	**ß = −5.31**	**ß = −1.84**	**ß = −1.58**	**ß = −1.99**
	t = −1.85	**t = −2.51**	**t = −2.32**	**t = −5.28**	**t = −4.49**	**t = −5.25**
	p = .06	**p = .01***	**p = .021***	**p < .001***	**p < .001***	**p < .001***
Adjusted R^2	0.65	0.66	0.66	0.73	0.73	0.74

Note: Bolded values are statistically significant.

Abbreviations: AD, Alzheimer's Disease; ADAS-13, Alzheimer's Disease Assessment Scale-13; CDR-SB, Clinical Dementia Rating Scale – Sum of Boxes. Average, average of left and right SNIPE grading score; eMCI, early mild cognitive impairment; HC, hippocampus; lMCI, late mild cognitive impairment; HC, hippocampus.

*Represents results that were significant after false discovery rate correction for multiple comparisons.

significantly steeper in the later disease cohorts (lMCI and AD) compared to NCs. When examining the covariates, increased age was associated with increases in CDR-SB scores (0.01 points, $t = 2.49$, $p = .01$), male sex was not associated with change in CDR-SB scores ($t = -0.71$, $p = .48$), and education was only marginally associated with lower CDR-SB scores (−0.02 points, $t = -1.85$, $p = .06$). Overall, 73% of the variance in CDR-SB scores (adjusted $R^2 = 0.73$) can be explained by the variables included in this model. Note that the same model without grading has an adjusted $R^2 = 0.71$, indicating that grading explains an additional 2% of variance.

For the right hippocampus model, the effect of rHC grading on CDR-SB scores was not significant ($t = -0.76$, $p = .45$) for the NC group. All patient groups had greater intercepts for CDR-SB than the NCs, and CDR-SB scores progressively increased (i.e., lower performance) from NC to eMCI (1.30 points more than NC, $t = 10.25$, $p < .001$), to lMCI (1.58 points more than NC $t = 16.50$, $p < .001$), and to AD (4.06 points more than NC, $t = 37.48$, $p < .001$). The interaction between rHC grading and CDR-SB was not significant for eMCI ($t = -0.94$, $p = .35$), but was significant for lMCI ($t = -2.22$ $p = .03$), and AD ($t = -4.49$, $p < .001$); that is, the slopes of changes in CDR-

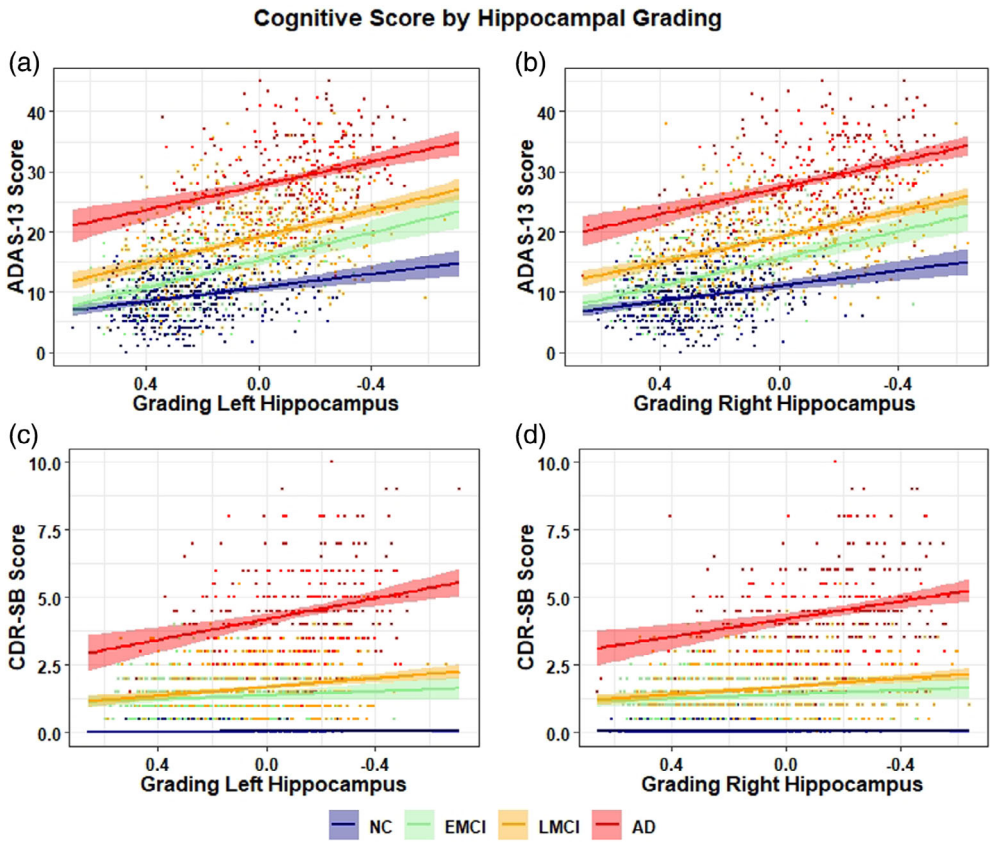

FIGURE 1 Scatterplots of Cognitive Score by Hippocampal Grading. All images show individual points grouped by color as well as the regression line for each group. Dark blue lines = cognitively normal controls (NC); Light green lines = early mild cognitive impairment (eMCI); orange lines = late mild cognitive impairment (lMCI); and red lines = Alzheimer's disease (AD). ADAS-13 = Alzheimer's Disease Assessment Scale–Cognitive Subscale. CDR-SB = Clinical Dementia Rating Scale – Sum of Boxes. Negative grading scores indicate greater similarity to the Alzheimer's anatomy, while positive scores indicate similarity to healthy controls. From left to right (A, B) Higher ADAS-13 scores were associated with decreases in hippocampal grading in all groups. (C, D) Higher CDR-SB scores were associated with decreases in hippocampal grading in lMCI and AD.

SB scores associated with changes in HC grading were significantly steeper in the later disease cohorts (lMCI and AD) in contrast with the NCs. When examining the covariates, increased age was associated with slight increases in CDR-SB scores (0.01 points, $t = 2.53$, $p = .01$), male sex was not associated with change in CDR-SB scores ($t = -0.63$, $p = .53$), whereas increased education was associated with slightly lower CDR-SB scores (0.02 points, $t = -2.36$, $p = .02$). Overall, 73% of the variance in CDR-SB scores (adjusted $R^2 = 0.73$) can be explained by the variables included in this model. Note that this model explains 2% more of the variance the same model *without grading*.

To ensure that the significant associations observed were not influenced by vascular pathology typically observed in MCI and AD, all models were rerun including total white matter hyperintensity burden as a covariate. All results remained almost exactly the same in terms of effect size and statistical significance indicating that white matter hyperintensity burden does not influence the association between hippocampal grading and cognition.

3.3 | RAVLT and hippocampal grading in preclinical AD

Table 3 summarizes the results of the linear regression models for the RAVLT analysis. Figure 2 shows a scatterplot of individual cognitive scores and hippocampal grading values for RAVLT percent forgetting and immediate scores. It should be noted that lower memory performance is associated with higher scores in RAVLT percent forgetting and lower scores in RAVLT immediate.

For the left HC, the effect of grading was not significant for SCD− for percent forgetting ($t = -1.80$, $p = .07$) or immediate ($t = -2.07$ $p = .038$) scores after correction for multiple comparisons. SCD+ had greater intercepts for RAVLT percent forgetting (15.87 more points, $p = .008$), but not for immediate. The interaction between grading and RAVLT was significant for percent forgetting ($t = -2.24$ $p = .026$) but not immediate. That is, the slopes of the RAVLT scores were steeper in SCD+ compared to SCD− for percent forgetting only. Overall, 6% of the variance in RAVLT percent forgetting scores (adjusted $R^2 = 0.06$) can be explained by the variables included in this model. Note that the same model without grading has an adjusted $R^2 = 0.04$, indicating that grading explains an additional 2% of variance. For RAVLT immediate, 18% of the variance (adjusted $R^2 = 0.18$) can be explained by the variables included in this model. Note that the same model without grading has an adjusted $R^2 = 0.15$, indicating that grading explains an additional 3% of variance.

For the right HC in the SCD− group, the effect of grading was not significant for either RAVLT percent forgetting ($t = -1.14$, $p = .26$) or immediate ($t = 1.83$ $p = .07$) scores. The SCD+ group had greater intercepts for RAVLT percent forgetting (24.12 more points, $p < .002$), but not for immediate. The interaction between grading and RAVLT was significant for percent forgetting ($t = -3.05$ $p = .002$) but again, not immediate. That is, similar to the lHC, the slopes of the RAVLT scores were steeper in SCD+ compared to SCD− for percent forgetting only. Overall, 7% of the variance in RAVLT percent

TABLE 3 Linear regression model results showing association between grading and episodic memory.

	RAVLT percent forgetting: Left HC	RAVLT percent forgetting: Right HC	RAVLT percent forgetting: Average HC	RAVLT immediate: Left HC	RAVLT immediate: Right HC	RAVLT immediate: Average HC
Grading	ß = -15.97	ß = -10.67	ß = -18.21	**ß = 6.15**	ß = 5.7	**ß = 7.88**
	$t = -1.80$	$t = -1.14$	$t = -1.76$	**$t = 2.07$**	$t = 1.83$	**$t = 2.27$**
	$p = .07$	$p = .26$	$p = .08$	**$p = .038$**	$p = .07$	**$p = .023$***
SCD+	**ß = 15.87**	**ß = 24.12**	**ß = 21.77**	ß = -2.65	ß = -2.69	ß = -3.02
	$t = 2.65$	**$t = 3.43$**	**$t = 3.17$**	$t = -1.33$	$t = -1.14$	$t = -1.31$
	$p = .008$*	**$p < .001$***	**$p = .002$***	$p = .18$	$p = .25$	$p = .19$
*Grading*SCD+*	**ß = -42.33**	**ß = -63.02**	**ß = -59.27**	ß = 9.48	ß = 8.05	ß = 9.89
	$t = -2.24$	**$t = -3.05$**	**$t = -2.78$**	$t = 1.51$	$t = 1.16$	$t = 1.39$
	$p = .026$*	**$p = .002$***	**$p = .006$***	$p = .13$	$p = .25$	$p = .17$

Note: Bolded values are statistically significant. It should be noted that lower memory performance is associated with higher scores in RAVLT percent forgetting and lower scores in RAVLT immediate. Average = average of left and right SNIPE grading score.
Abbreviations: RAVLT, Rey Auditory Verbal Learning test; SCD+, subjective cognitive decline; HC, hippocampus.
*Represents results that were significant after false discovery rate correction for multiple comparisons.

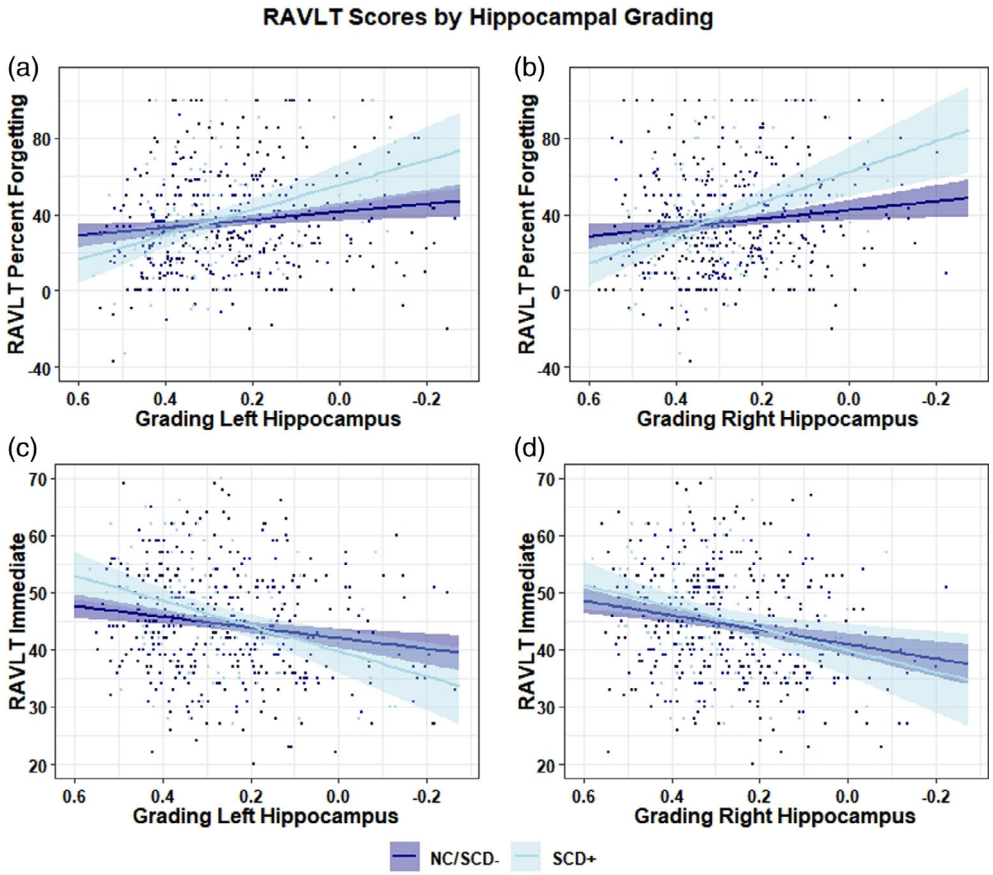

FIGURE 2 Scatterplots of RAVLT scores by Hippocampal Grading. All images show individual points grouped by color as well as the regression line for each group. Dark blue lines = cognitively normal controls without subjective cognitive decline (NC/SCD−); Light blue lines = cognitively normal controls with subjective cognitive decline (SCD+); RAVLT = Rey Auditory Verbal Learning Test. Negative grading scores indicate greater similarity to the Alzheimer's anatomy, while positive scores indicate similarity to healthy controls. From left to right (A, B): Increased RAVLT percent forgetting scores were associated with decreases in hippocampal grading in SCD+. (C, D) Lower RAVLT immediate scores were associated with decreases in only left hippocampal grading in NCs/SCD−.

forgetting scores (adjusted $R^2 = 0.07$) can be explained by the variables included in this model. Note that the same model without grading has an adjusted $R^2 = 0.04$, indicating that grading explains an additional 3% of variance. For RAVLT immediate, 17% of the variance (adjusted $R^2 = 0.17$) can be explained by the variables included in this model. Note that the same model without grading has an adjusted $R^2 = 0.17$, indicating that grading explains an additional 2% of variance.

4 | DISCUSSION

The current study used SNIPE to measure hippocampal grading in NCs, eMCI, lMCI, and AD, and compared these grading scores to global cognitive functioning as measured by the ADAS-13 and CDR-SB. In contrast to NCs, both lMCI and AD required less hippocampal change to have decreased ADAS-13 and CDR-SB scores (reflected by their significantly steeper slopes in the regression models). The eMCI

group required less hippocampal change compared to normal controls to have decreases in only the ADAS-13; whereas this group's association between hippocampal change and CDR-SB scores did not differ from NC. These findings indicate a relationship between lower cognitive scores in eMCI, lMCI, and AD with decreases in hippocampal grading. Importantly, in contrast to SCD− NCs, the SCD+ group required less hippocampal change to exhibit lower RAVLT performance. Hippocampal grading is thus also sensitive to changes that occur in episodic memory early in the preclinical AD phase.

Previous associations between global cognition in people with MCI and hippocampal volume have been mixed. Vipin et al. (2018) reported no relationship between hippocampal volume and global cognition in MCI (Vipin et al., 2018), while Peng et al. (2015) found an association between hippocampal volume and global cognition in people with MCI . These conflicting findings could be associated with the sensitivity of the MMSE to early cognitive decline as well as the use of hippocampal volume.

In this study, we observed an association between hippocampal grading and global cognition, as measured by the ADAS-13 score in all groups—NC, eMCI, lMCI, and AD. This finding suggests that SNIPE hippocampal grading is sensitive to global cognitive declines due to aging (in the NC group), early in the disease process (i.e., eMCI), and to the progressive changes that occur later in AD-related pathology.

Taken together with the previous research on SNIPE (Coupé et al., 2015; Coupé, Eskildsen, Manjón, Fonov, Pruessner, et al., 2012; Coupé, Eskildsen, Manjón, Fonov, & Collins, 2012), these results suggest that this method could be useful in the future prediction of cognitive decline and diagnostic status early in the disease trajectory.

When examining global cognition using the CDR-SB, associations between hippocampal grading were only observed in the lMCI and AD groups. As expected, CDR-SB scores were low for the normal controls (i.e., either 0 or 0.5), resulting in a flat slope for this group. As can be seen in Figure 1, the eMCI group has a much smaller range and median (0–4; median = 1) CSR-SB values than both lMCI (0–5.5; median = 1.5) and AD (1–10; median = 5.5). The lack of association with the eMCI group may thus be related to the limited range of CDR-SB scores in people with eMCI. Furthermore, the CDR-SB score is calculated based on subjective judgment by an interviewer and shows the most sensitivity across a wide range of symptom variation. Therefore, interpretation of results in individuals with minimal (or no) cognitive impairment such as NC and eMCI, should be completed cautiously. The relationship became stronger with lMCI and AD, showing that the relationship between hippocampal grading and CDR-SB is stronger in those with more severe declines. Consistent with previous studies, associations between CDR-SB and hippocampal grading were also stronger in the left, rather than right hippocampus (Basso et al., 2006; Peng et al., 2015).

While one would not expect a simple function with only a few variables to explain all the domains represented in the ADAS-13, CDR-SB, or RAVLT, it is interesting to note that the hippocampal grading score explains 4%–5% more variance of ADAS-13 than the simpler model. For CDR-SB, hippocampal grading explains a smaller amount of variance, only 2%, not accounted for by the model without grading. This low percentage is likely due to a lessor dependence on the hippocampus for the domains covered by CDR-SB. Finally, for RAVLT forgetting and immediate scores, adding the hippocampal grading measure explains 3% more of the variance—75% more variance explained for RAVLT forgetting and 20% more for RAVLT immediate. This greater contribution of the hippocampal grading score over that obtained for ADAS-13 or CDR-SB is likely due to the larger dependence on the hippocampus for the RAVLT measures. While both ADAS-13 and CDR-SB were associated with hippocampal grading for both lMCI and AD, only the ADAS-13 was associated with hippocampal grading in eMCI and NC. CDR-SB was more sensitive (than ADAS-13) to changes in hippocampal grading that occur later in the disease progression, as reflected by the interaction with grading and lMCI and AD.

Hippocampal grading was also observed to be sensitive to early stages of decline in people with SCD. Episodic memory has been observed to be associated with the hippocampus in SCD (Caillaud et al., 2020). Furthermore, RAVLT performance has been shown to help identify those with SCD who will progress to AD (Estévez-González et al., 2003), with high percent forgetting scores observed in people with AD (Lodha et al., 2018). In our study, we expanded on these findings by observing that hippocampal grading is associated with increased percent forgetting scores (i.e., poorer performance) in those with SCD. Given the previous associations between percent forgetting and AD (e.g., Lodha et al., 2018; Moradi et al., 2017), the associations observed here between grading and percent forgetting could be linked to AD-related degenerative changes that occur early in the disease process. These findings suggest that the relationship between grading and cognitive score is sensitive to different stages of disease progression, starting early in the preclinical period.

Regarding the normal controls, these findings suggest they would require much more hippocampal change (as measured by grading) for the same amount of cognitive decline to occur compared to the patient groups. For example, HC changes may affect the grading score, but compensation through cognitive reserve or plasticity may limit declines in cognitive functioning measured by the cognitive tests. Further research is needed to elucidate this issue.

The present study has a few limitations that should be investigated in future research. Participants in the current sample had a high education (mean = 15.90 years) and were mainly White (n = 1508 or 93% of the sample), which may limit the interpretation and generalizability of the results to other populations. Future studies should replicate these findings in another dataset that is more generalizable to diverse populations. The current study should also be replicated in longitudinal data to determine whether grading scores are predictive of clinical cognitive change at an individual level.

The preprocessing tools used in this manuscript have been extensively validated for use in multicenter and multiscanner studies that examine atrophy in aging, MCI, and AD populations (Dadar, Camicioli, et al., 2020; Dadar, Gee, et al., 2020; Manera et al., 2019). Taken together, we are confident that the results found in the current study reflect the heightened sensitivity of SNIPE grading to detect structural and cognitive changes associated with early cognitive decline.

5 | CONCLUSION

The findings from this study indicate a strong association between hippocampal grading and cognition. Importantly, the relationships were observed not only late in the disease course (AD and lMCI) but also earlier in the course of the disease in eMCI, NC, and SCD. Hippocampal grading is sensitive to global cognition scores as measured by the ADAS-13 starting in the normal aging and CDR-SB starting in prodromal AD, as well as in episodic memory in the pre-clinical AD group. Although future work is needed to determine if HC grading is predictive and associated with longitudinal changes in cognition at an individual level, the current study suggests that hippocampal grading may be a useful measure that is sensitive to cognitive changes early in the disease course.

ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

FUNDING INFORMATION

Alzheimer's Disease Neuroimaging Initiative; This research was supported by a grant from the Canadian Institutes of Health Research and a donation from the Famille Louise & André Charron. Cassandra Morrison was supported by a postdoctoral fellowship from Canadian Institutes of Health Research, Funding Reference Number: MFE-176608. Mahsa Dadar reports receiving research funding from the Healthy Brains for Healthy Lives (HBHL), Alzheimer Society Research Program (ASRP), and Douglas Research Centre (DRC). Neda Shafiee has no conflicts of interest to declare. D. Louis Collins reports receiving research funding from Canadian Institutes of Health research, the Canadian National Science and Engineering Research Council, Brain Canada, the Weston Foundation, and the Famille Louise & André Charron.

CONFLICT OF INTEREST STATEMENT

The authors report no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf. The data used for this analysis are available on request from the ADNI database (ida.loni.usc.edu).

ORCID

Cassandra Morrison https://orcid.org/0000-0002-1087-5770

REFERENCES

Alzheimer's Association. (2021). *What is Alzheimer's disease*. https://www.alz.org/alzheimers-dementia/what-is-alzheimers.

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Basso, M., Yang, J., Warren, L., MacAvoy, M. G., Varma, P., Bronen, R. A., & van Dyck, C. H. (2006). Volumetry of amygdala and hippocampus and memory performance in Alzheimer's Disease. *Psychiatry Research: Neuroimaging*, *146*(3), 251–261. https://doi.org/10.1016/j.pscychresns.2006.01.007

Bettio, L. E., Rajendran, L., & Gil-Mohapel, J. (2017). The effects of aging in the hippocampus and cognitive decline. *Neuroscience & Biobehavioral Reviews*, *79*, 66–86. https://doi.org/10.1016/j.neubiorev.2017.04.030

Caillaud, M., Hudon, C., Boller, B., Brambati, S., Duchesne, S., Lorrain, D., Gagnon, J. F., Maltezos, S., Mellah, S., Phillips, N., Consortium for the Early Identification of Alzheimer's Disease-Quebec, & Belleville, S. (2020). Evidence of a relation between hippocampal volume, white matter hyperintensities, and cognition in subjective cognitive decline and mild cognitive impairment. *The Journals of Gerontology: Series B*, *75*(7), 1382–1392. https://doi.org/10.1093/geronb/gbz120

Coupé, P., Eskildsen, S. F., Manjón, J. V., Fonov, V. S., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2012). Simultaneous segmentation and grading of anatomical structures for patient's classification: Application to Alzheimer's disease. *NeuroImage*, *59*(4), 3736–3747. https://doi.org/10.1016/j.neuroimage.2011.10.080

Coupé, P., Eskildsen, S. F., Manjón, J. V., Fonov, V. S., Pruessner, J. C., Allard, M., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2012). Scoring by nonlocal image patch estimator for early detection of Alzheimer's disease. *NeuroImage: Clinical*, *1*(1), 141–152. https://doi.org/10.1016/j.nicl.2012.10.002

Coupé, P., Fonov, V. S., Bernard, C., Zandifar, A., Eskildsen, S. F., Helmer, C., Manjón, J. V., Amieva, H., Dartigues, J. F., Allard, M., Catheline, G., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2015). Detection of Alzheimer's disease signature in MR images seven years before conversion to dementia: Toward an early individual prognosis. *Human Brain Mapping*, *36*(12), 4758–4770. https://doi.org/10.1002/hbm.22926

Coupé, P., Yger, P., Prima, S., Hellier, P., Kervrann, C., & Barillot, C. (2008). An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Transactions on Medical Imaging*, *27*(4), 425–441. https://doi.org/10.1109/TMI.2007.906087

Dadar, M., Camicioli, R., Duchesne, S., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2020). The temporal relationships between white matter hyperintensities, neurodegeneration, amyloid beta, and cognition. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *12*(1), e12091. https://doi.org/10.1002/dad2.12091

Dadar, M., Fonov, V. S., Collins, D. L., & Alzheimer's Disease Neuroimaging Initiative. (2018). A comparison of publicly available linear MRI stereotaxic registration techniques. *NeuroImage*, *174*, 191–200. https://doi.org/10.1016/j.neuroimage.2018.03.025

Dadar, M., Gee, M., Shuaib, A., Duchesne, S., & Camicioli, R. (2020). Cognitive and motor correlates of grey and white matter pathology in Parkinson's disease. *NeuroImage: Clinical*, *27*, 102353. https://doi.org/10.1016/j.nicl.2020.102353

Dawe, R. J., Yu, L., Arfanakis, K., Schneider, J. A., Bennett, D. A., & Boyle, P. A. (2020). Late-life cognitive decline is associated with hippocampal volume, above and beyond its associations with traditional neuropathologic indices. *Alzheimer's & Dementia*, *16*(1), 209–218. https://doi.org/10.1002/alz.12009

Estévez-González, A., Kulisevsky, J., Boltes, A., Otermín, P., & García-Sánchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: Comparison with mild cognitive impairment and normal aging. *International Journal of Geriatric Psychiatry*, *18*(11), 1021–1028. https://doi.org/10.1002/gps.1010

Fjell, A. M., McEvoy, L., Holland, D., Dale, A. M., Walhovd, K. B., & Alzheimer's Disease Neuroimaging Initiative. (2014). What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in Neurobiology*, *117*, 20–40. https://doi.org/10.1016/j.pneurobio.2014.02.004

Fonov, V., Evans, A., Botteron, K., Almli, C., McKinstry, R., & Collins, D. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54*(1), 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033

Gorbach, T., Pudas, S., Lundquist, A., Orädd, G., Josefsson, M., Salami, A., de Luna, X., & Nyberg, L. (2017). Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiology of Aging*, *51*, 167–176. https://doi.org/10.1016/j.neurobiolaging.2016.12.002

Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1998). Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology*, *51*(4), 993–999.

Jack, C. R., Jr., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Contributors. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018.NIA-AA

Köhncke, Y., Düzel, S., Sander, M. C., Lindenberger, U., Kühn, S., & Brandmaier, A. M. (2021). Hippocampal and parahippocampal gray matter structural integrity assessed by multimodal imaging is associated with episodic memory in old age. *Cerebral Cortex*, *31*(3), 1464–1477. https://doi.org/10.1093/cercor/bhaa287

Lodha, P., Talele, A., & Degaonkar, K. (2018). Diagnosis of Alzheimer's Disease using machine learning. In *2018 fourth international conference on computing communication control and automation (ICCUBEA), 2018* (pp. 1–4). IEEE Transactions on Medical Imaging. https://doi.org/10.1109/ICCUBEA.2018.8697386

Manera, A. L., Dadar, M., Collins, D. L., Ducharme, S., & Frontotemporal Lobar Degeneration Neuroimaging Initiative. (2019). Deformation based morphometry study of longitudinal MRI changes in behavioral variant frontotemporal dementia. *Neuroimage: Clinical.*, *24*(24), 102079. https://doi.org/10.1016/j.nicl.2019.102079

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research*, *43*(4), 411–431. https://doi.org/10.1016/j.jpsychires.2008.04.014

Moradi, E., Hallikainen, I., Hänninen, T., Tohka, J., & Alzheimer's Disease Neuroimaging Initiative. (2017). Rey's auditory verbal learning test scores can be predicted from whole brain MRI in Alzheimer's disease. *NeuroImage: Clinical*, *13*, 415–427. https://doi.org/10.1016/j.nicl.2016.12.011

Morrison, C., Dadar, M., Shaifee, N., & Collins, D. L. (2022). *Hippocampal grading provides higher Alzheimer's Disease prediction accuracy than hippocampal volume*. Manuscript submitted for publication.

Peng, G. P., Feng, Z., He, F. P., Chen, Z. Q., Liu, X. Y., Liu, P., & Luo, B. Y. (2015). Correlation of hippocampal volume and cognitive performances in patients with either mild cognitive impairment or Alzheimer's disease. *CNS Neuroscience & Therapeutics*, *21*(1), 15–22. https://doi.org/10.1111/cns.12317

Persson, J., Pudas, S., Lind, J., Kauppi, K., Nilsson, L. G., & Nyberg, L. (2012). Longitudinal structure–function correlates in elderly reveal MTL dysfunction with cognitive decline. *Cerebral Cortex*, *22*(10), 2297–2304. https://doi.org/10.1093/cercor/bhr306

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Reisberg, B., Shulman, M. B., Torossian, C., Leng, L., & Zhu, W. (2010). Outcome over seven years of healthy adults with and without subjective cognitive impairment. *Alzheimer's & Dementia*, *6*(1), 11–24. https://doi.org/10.1016/j.jalz.2009.10.002

Risacher, S., Kim, S., Nho, K., Foroud, T., Shen, L., Petersen, R., Jack, C., Jr., Beckett, L., Aisen, P., Koeppe, R., & Jagust, W. (2015). APOE effect on Alzheimer's disease biomarkers in older adults with significant memory concern. *Alzheimer's & Dementia*, *11*(12), 1417–1429. https://doi.org/10.1016/j.jalz.2015.03.003

Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., McHugh, T. L., & Mamourian, A. C. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. *Neurology*, *67*(5), 834–842. https://doi.org/10.1212/01.wnl.0000234032.77541.a2

Schneider, A. L., Senjem, M. L., Wu, A., Gross, A., Knopman, D. S., Gunter, J. L., ALC, S., Senjem, M. L., Wu, A., Gross, A., Knopman, D. S., Gunter, J. L., Schwarz, C. G., Mosley, T. H., Gottesman, R. F., Sharrett, A. R., & Jack, C. R., Jr. (2019). Neural correlates of domain-specific cognitive decline: The ARIC-NCS study. *Neurology*, *92*(10), e1051–e1063. https://doi.org/10.1212/WNL.0000000000007042

Strikwerda-Brown, C., Gonneaud, Julie, Hobbs, Diana, Frederic, S.-O., Binette, A. P., Ozlen, H., Provost, K., & Villeneuve, S. (2022). AT(N) predicts near-term development of Alzheimer's disease symptoms in unimpaired older adults. *Authors*, *2858*, 1–28.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*(1), 87–97. https://doi.org/10.1109/42.668698

Tzeng, R. C., Yang, Y. W., Hsu, K. C., Chang, H. T., & Chiu, P. Y. (2022). Sum of boxes of the clinical dementia rating scale highly predicts conversion or reversion in predementia stages. *Frontiers in Aging Neuroscience*, *14*, 1021792. https://doi.org/10.3389%2Ffnagi.2022.1021792

Vipin, A., Foo, H. J. L., Lim, J. K. W., Chander, R. J., Yong, T. T., Ng, A. S. L., Hameed, S., Ting, S. K. S., Zhou, J., & Kandiah, N. (2018). Regional white matter hyperintensity influences grey matter atrophy in mild cognitive impairment. *Journal of Alzheimer's Disease*, *66*(2), 533–549. https://doi.org/10.3233/JAD-180280

How to cite this article: Morrison, C., Dadar, M., Shafiee, N., Collins, D. L., & For Alzheimer's Disease Neuroimaging Initiative (2023). The use of hippocampal grading as a biomarker for preclinical and prodromal Alzheimer's disease. *Human Brain Mapping*, 1–11. https://doi.org/10.1002/hbm.26269