Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 11 (2019) 599-609

Neuroimaging

Braak neurofibrillary tangle staging prediction from in vivo MRI metrics

Caroline Dallaire-Théroux[a,b,*], Iman Beheshti[a], Olivier Potvin[a], Louis Dieumegarde[a], Stephan Saikali[b,c], Simon Duchesne[a,b], for the National Alzheimer's Coordinating Center, the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*CERVO Brain Research Center, Quebec City, Quebec, Canada*
[b]*Faculty of Medicine, Université Laval, Quebec City, Quebec, Canada*
[c]*Department of pathology, Centre Hospitalier Universitaire de Quebec, Quebec City, Quebec, Canada*

Abstract

Introduction: Alzheimer's disease diagnosis requires postmortem visualization of amyloid and tau deposits. As brain atrophy can provide assessment of consequent neurodegeneration, our objective was to predict postmortem neurofibrillary tangles (NFT) from in vivo MRI measurements.
Methods: All participants with neuroimaging and neuropathological data from the Alzheimer's Disease Neuroimaging Initiative, the National Alzheimer's Coordinating Center and the Rush Memory and Aging Project were selected (n = 186). Two hundred and thirty two variables were extracted from last MRI before death using FreeSurfer. Nonparametric correlation analysis and multivariable support vector machine classification were performed to provide a predictive model of Braak NFT staging.
Results: We demonstrated that 59 of our MRI variables, mostly temporal lobe structures, were significantly associated with Braak NFT stages ($P < .005$). We obtained a 62.4% correct classification rate for discrimination between transentorhinal, limbic, and isocortical groups.
Discussion: Structural neuroimaging may therefore be considered as a potential biomarker for early detection of Alzheimer's disease–associated neurofibrillary degeneration.

Keywords: Neurofibrillary degeneration; Tau pathology; Neuropathology; Neuroimaging; Structural MRI; Imaging biomarkers; Early diagnosis; Predictive model; Machine-learning; Alzheimer's disease; Dementia

1. Background

Dementias are more likely to be associated with multiple confounding etiologies; unsurprisingly, concordance with clinical syndromes is not perfect [1]. Thus, the definitive diagnosis of its leading cause, Alzheimer's disease (AD), can only be established following postmortem examination, which has long been and remains the gold standard. The "ABC" staging scheme is the recommended assessment method [2], and requires the neuropathological gradation of (1) diffuse beta-amyloid plaques following the scale of Thal et al [3]; (2) neurofibrillary tangles (NFT) according to the Braak & Braak [4] scale; and (3) neuritic plaques (NP) according to the CERAD scale of Mirra et al [5]. Yet, these characteristic pathological changes are thought to occur decades before initial clinical manifestations [6,7]. This highlights the evident need for the development of in vivo diagnostic tools and standardized methods that could be used to detect the presence of AD pathology in preclinical disease stages. Noninvasive neuroimaging techniques have the potential to address this issue by

Conflict of interest: S.D. is the officer and shareholder of True Positive Medical Devices Inc. Other authors report no conflict of interest.

[1]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.

*Corresponding author. Tel: +(418) 663-5000 x6714; Fax: +(418) 663-5971.

E-mail address: caroline.dallaire-theroux.1@ulaval.ca

https://doi.org/10.1016/j.dadm.2019.07.001

providing structural and functional information in the human living brain that could be correlated to one or many of the aforementioned ABC pathological signs. Among these modalities, magnetic resonance imaging (MRI) appears to be one of the safest and most accessible methods to visualize nervous tissue inside the cranial cavity with sufficient precision. Indeed, high-resolution structural MRI allows for the quantification of tissue losses, whereas recent sophisticated machine-learning tools have been applied during MRI postprocessing to analyze brain tissue. Based on these technologies, regional MRI metrics may therefore be used as a candidate biomarker to predict underlying AD pathology.

Among all pathological features of AD, we identified NFT pathology as the strongest correlate of atrophy on MRI [8], which remains the only significant feature even after accounting for amyloid pathology. Intraneuronal paired helical filaments of hyperphosphorylated tau protein aggregates are characteristics of AD and their deposition follows a well-defined, predictable trajectory (assessed by Braak staging [4,9]) that has been shown to correlate with the pattern of regional brain atrophy observed on structural MRIs of AD patients [10–19]. Emerging from the transentorhinal region (stage I), the immunochemistry-defined progression of tau pathology gradually involves the entorhinal region and hippocampus (stage II), the neocortex of the fusiform and lingual gyri (stage III), neocortical associative areas (stage IV), frontal, superolateral, and occipital territories and the peristriate region (stage V), and finally the secondary and primary sensory areas, including the striate area (stage VI) [9].

Previous structural MRI studies in pathologically confirmed AD subjects, although generally based on limited sample sizes, have shown that volumes of the entorhinal and hippocampal cortices are significantly correlated to postmortem changes in AD. However, most studies seem to have restricted their focus only on specific medial temporal lobe areas predominantly affected in AD [10,12,13,17,19–22], otherwise on measurements of whole-brain atrophy [15,23]. To our knowledge, only one study did focus on Braak stages as a regionally unbiased analysis approach using voxel-based analyses [16]. Hence, most of these previous reports have not accurately reproduced the complex spatiotemporal pattern of NFT spread as observed in AD brains. Moreover, brain volume loss at early stages of the disease is usually relatively subtle and heterogeneously distributed over many brain regions [24–29], and may therefore be confounded by complex patterns of brain structural changes related to normal aging [30].

Our work attempted to provide a more extensive regional analysis of the correlation between structural MRI to NFT staging, with the aim to examine the predictive value of antemortem MRI metrics on established postmortem measure of AD-associated neurofibrillary degeneration. Based on our current knowledge and exhaustive literature review [8], we hypothesized that machine-learning classification techniques applied to structural MRI data would allow us to predict Braak NFT staging in the living brain, accounting for the complex patterns of spatial changes related to normal aging and AD.

2. Methods

2.1. Participants

Participants included in this study were obtained from three large North American data sets (i.e., the Alzheimer's Disease Neuroimaging Initiative [ADNI], the National Alzheimer's Coordinating Center [NACC], and the Rush Memory and Aging Project [MAP]). The ADNI (adni.loni.usc.edu) was launched in 2003 as a multicenter public-private partnership, led by Principal Investigator Michael W. Weiner, MD, and regrouping participants from 63 sites across the United States and Canada. The primary goal of ADNI has been to test whether serial MRI, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early AD. The NACC data set (www.alz.washington.edu) was established in 1999 and provides data from patients collected in the AD centers funded by the National Institute of Aging, U.S. NACC's Uniform Data Set used for the present study started more precisely on September 2005. The MAP (www.radc.rush.edu) is a community-based longitudinal study of aging that was carried out at Rush University Medical Center (Chicago, IL). It started in 1997 and was funded by the National Institute of Aging, U.S. The NACC and ADNI cohorts included both normal and cognitively impaired individuals at baseline. All participants in MAP had no known dementia before enrollment. All combined, ADNI, NACC, and MAP provide data from more than 30,000 participants. All participants underwent extensive clinical evaluation by both a physician and a neuropsychologist with expertise in evaluation and management of elderly patients presenting with cognitive impairment, including medical history, physical examination, and neuropsychological testing. Up-to-date information can be found at respective websites of each data set. In this study, we first included all participants with concurrent antemortem MRI scan and postmortem neuropathological examination available as of January 12, 2017 ($n = 204$). After initial MRI quality control, eighteen participants were excluded because of poor quality scans. In total, we report results based on 186 subjects from three independent data sets. Respective institutional review boards approved the three studies, and written informed consent was obtained from each participant. Approval from our local ethic committee was also obtained for the present study.

2.2. MRI acquisition and image processing

Image analyses were conducted using 3D T1-weighted antemortem anatomical images obtained from multiple centers using 1.5 T and 3.0 T MRI scanners (Siemens, Philips,

and GE Medical Systems) with MPRAGE sequences. MRI data were automatically segmented from the closest images to death (i.e., last scan acquired during the study) with *FreeSurfer* software version 5.3 (Laboratory for Computational Neuroimaging, Martinos Center for Biomedical Imaging; https://surfer.nmr.mgh.harvard.edu/) based on the complete Desikan-Killiany-Tourville atlas [31] and part of the ex vivo atlas for entorhinal and perirhinal cortices [32,33]. We performed initial visual inspection of segmentations for quality control, and performed minor manual corrections of regional segmentation, mainly concerning the dura, where needed. We then extracted surface, thickness, and volume measurements from all available cortical structures and volume measurements from subcortical structures, in both hemispheres, for a total of 232 MRI-derived variables available for cross-sectional analyses. Cortical regions included total cortex, rostral and caudal anterior cingulate, posterior cingulate, isthmus cingulate, pars opercularis, pars orbitalis, pars triangularis, rostral and caudal middle frontal, superior frontal, medial and lateral orbitofrontal, precentral, paracentral, postcentral, supramarginal, inferior parietal, superior parietal, pericalcarine, precuneus, cuneus, lingual, lateral occipital, insular, entorhinal, perirhinal, fusiform, parahippocampal, inferior temporal, middle temporal, superior temporal, and transverse temporal cortical areas. Subcortical regions included corpus callosum, accumbens nucleus, amygdala, hippocampus, caudate nucleus, pallidum, putamen, thalamus proper, ventral diencephalon, total brain stem, and lateral, third, and fourth ventricles.

2.3. Neuropathological assessment

We used available neuropathological data previously obtained during postmortem examination by neuropathologists. NFT pathology graded according to Braak staging was selected as the main dependent variable of interest. Briefly, neuropathological assessment was performed on one hemisphere (left side for most cases), which was fixed in paraformaldehyde and embedded in paraffin. Tissue blocks were then dissected from 0.5-1 cm slabs. Sections from standard blocks (i.e., mid-frontal, superior/middle temporal, inferior parietal, occipital, entorhinal and hippocampus; as described in the study by Montine et al [2]) were stained with tau or phosphorylated-tau antibodies or silver-based histochemistry and assessed for Braak NFT staging as described earlier. The other half of the brain was frozen for further biochemical studies. For more details, please refer to the respective neuropathological protocols of the three data sets.

2.4. Descriptive data analysis

A statistical approach was selected for preliminary descriptive analyses of the data. Statistical analyses were conducted using RStudio (http://www.rstudio.com/) version 1.0.143. We examined group differences between participants with Braak stages 0 to VI using Kruskal-Wallis test by ranks for continuous variables and Pearson χ^2 for categorical variables. We chose to regroup subjects into three classes: transentorhinal (I-II), limbic (III-IV), and isocortical (V-VI) stages. This simplified version of Braak stratification is presumed to minimize the impact of subjective pathological scoring and to improve inter-rater reliability [34]. The number of subjects with Braak 0 (n = 9) was insufficient to reach minimal conditions for statistical analyses and machine-learning model training. The control group was therefore incorporated to the subclinical transentorhinal class for this study. This merger is concordant with the "B" component of current "ABC" guidelines [2] in which no distinction is made between B0 (Braak 0) and B1 (Braak I and II) as it does not affect the level of AD neuropathologic change (i.e., not, low, intermediate, or high). Braak stages 0 through II are also considered as pathological findings consistent with typical aging [35]. Nonparametric analyses were further performed to assess the correlation between regional MRI metrics and Braak NFT stages using Spearman's rank correlation test. A 0.005 significant threshold (*P*-value) was fixed to correct for multiple comparisons.

2.5. Classification-based machine-learning algorithms and validation

In the experiments, we conducted three binary classification tasks (i.e., Braak 0-I-II/Braak III-IV, Braak 0-I-II/Braak V-VI and Braak III-IV/Braak V-VI) as well as a multiclass classification task (i.e., Braak 0-I-II/Braak III-IV/Braak V-VI). To assess the reliability of proposed method, we used a 10-fold cross-validation strategy, such that randomly 10% of data is considered a test set and the rest data (i.e., 90%) as a training set. In the training set of each iteration, the features (including sex, age at death, time interval between last MRI and death, apoE4 allele status and MRI regional metrics) were ranked on the basis on a mutual information [36] strategy, as it has previously shown a robust performance in comparison with alternative feature ranking methods [37]. The features with higher mutual information weights are more informative. The number of top-ranked features for all experiments was specified by using 20% of total variance in the training set, to limit the number of features and avoid overdetermination. It is noted that the optimal number of selected features may be different in each iteration. The optimal subset derived from a training set was applied to the respective test set. The classification results were averaged over the rounds of validation by the means of accuracy (ACC), sensitivity (SEN), specificity (SPE), and area under the receiver operating characteristics curve (AUC). To implement the classification, a standard support vector machine algorithm with linear kernel implemented in MATLAB 2016 was used. Fig. 1 illustrates the outline of the proposed ranking-based classification method.

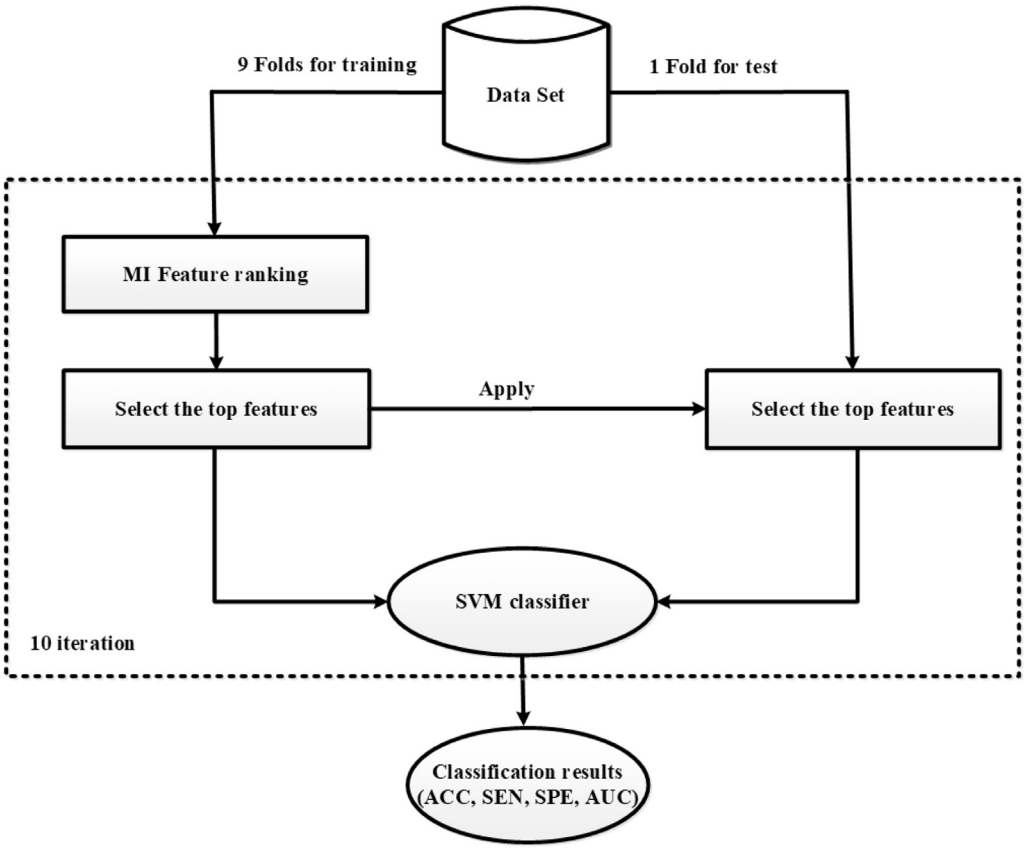

Fig. 1. Features ranking classification. This outline illustrates the proposed ranking-based classification method from a mutual information strategy.

Furthermore, to correct for morphometric differences between scanners, we tested whether scanner manufacturer and magnetic field strength variables could improve the prediction of the models.

3. Results

3.1. Participants characteristics

Among all 186 subjects included in this study, 24% come from ADNI (n = 44), 45% from NACC (n = 84), and 31% from MAP (n = 58). Caucasians (n = 175) represent 94.1% of our sample. The apoE4 status was available for 177 of our 186 participants. Two alleles were present in 5.4% of our sample, whereas one allele was found in 32.8%. The apoE4 status of the remaining subjects was either negative (57.0%) or unknown (4.8%). Distribution of Braak NFT groups across samples from the three data sets can be visualized in Supplementary Material 1. Participants were given a final clinical diagnosis of intact cognition (n = 44), mild cognitive impairment (n = 42), probable AD (n = 86), or other dementia (n = 14) before death. Participants were aged 45 to 102 years at time of death (mean age = 84.3 years; SD = 10.0) and underwent 1.5 T (n = 157) or 3.0 T (n = 29) brain MRI within a mean interval of 2.6 years (SD = 2.0) before death. The manufacturer of the scanner used for MRI acquisition was either Siemens (n = 40), Philips (n = 10), or GE (n = 136). More details on participant characteristics are summarized in Table 1.

3.2. Correlations between MRI metrics and NFT pathology

The correlation analysis between regional metrics and Braak NFT stages (0 to VI) showed significant associations ($P < .005$) for 59 of the 232 MRI variables. We found a relationship between several antemortem thickness and volume measurements and postmortem neurofibrillary degeneration within the temporal lobe, including entorhinal, perirhinal, fusiform, parahippocampal, inferior, middle, and superior temporal regions and the hippocampus. Multiple limbic (i.e., insula, amygdala, accumbens nucleus, posterior cingulate, and isthmus of cingulate gyrus), parietal (i.e., supramarginal gyrus, precuneus, and inferior and superior parietal cortices) and occipital regions (i.e., lingual gyrus and lateral occipital cortex) were also significantly correlated with NFT pathology. Other associated variables which reached statistical significance included volumes from the rostral middle frontal cortex and the ventricular system. No significant correlation was found between NFT pathology and any surface measurement across the brain. Absolute values of correlation coefficients for significant variables varied between

Table 1
Participants demographics

Characteristic	Controls (n = 44)	MCI (n = 42)	AD (n = 86)	Others (n = 14)
No. (%) of females	26 (59)	22 (52)	33 (38)	6 (43)
Median (range) education, years	15 (8, 23)	15 (0, 22)	16 (4, 20)	12 (6, 16)
Median (range) last MMSE, /30	28 (14, 30)	25.5 (15, 30)	16 (0, 28)	20 (3, 28)
Median (range) age at last MRI, years	87 (71, 102)	85.5 (50, 97)	82 (50, 94)	70.5 (43, 89)
Median (range) age at death, years	89 (72, 102)	86 (53, 99)	85 (52, 98)	76 (45, 92)
Median (range) interval MRI-death, years	2.0 (0, 5.9)	2.1 (0, 7.8)	2.0 (0, 9.8)	3.4 (0.7, 8.3)
Median (range) Braak NFT stage, /6	3 (1, 5)	4 (0, 6)	5 (0, 6)	2 (0, 6)
Median (range) CERAD NP score, /3	1 (0, 3)	2 (0, 3)	3 (0, 3)	0.5 (0, 3)

Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment; MMSE, Mini–Mental State Examination; MRI, magnetic resonance imaging; NFT, neurofibrillary tangles; NP, neuritic plaques.

0.21 and 0.43. The strongest association was found with the right hippocampus volume. The results are shown in Fig. 2.

3.3. Classification performance

Based on the best predictors for classification between each two Braak groups, three binary classifiers were developed. The 20 features selected with the highest frequency across training sets to be included in our models are presented in Supplementary Material 2. Individuals with Braak 0-I-II were distinguished from Braak III-IV with 70.2% accuracy (SEN 50.0%, SPE 82.9%). Participants with Braak III-IV were differentiated from Braak V-VI with 69.0% accuracy (SEN 71.4%, SPE 66.7%). Finally, subjects with Braak 0-I-II were separated from Braak V-VI with 71.6% accuracy (SEN 52.3%, SPE 83.3%). The resulting AUC for the three classifiers were 0.77, 0.64, and 0.69, respectively (Fig. 3). Performance parameters are summarized in Table 2. Our multiclass classification model applied to our test set provided 62.4% accuracy for discrimination between the three NFT pathological groups. The confusion matrix of predicted versus observed pathological groups is presented in Fig. 4. A similar model including clinical features only (i.e., age, sex, apoE4 status, and time interval between MRI and death) was tested and provided lower prediction accuracy, with only 45.2% of correctly classified cases (see Supplementary Material 3). Because adding scanner manufacturer and magnetic field strength in the models did not improve the prediction accuracy (Braak 0-II vs. III-IV: 69.0%; Braak III-IV vs. V-VI: 74.0%; Braak 0-II vs. V-VI: 69.6%; multiclass: 60.0), these variables were not retained for prediction.

4. Discussion

The objective of the present study was to predict postmortem Braak NFT staging based on antemortem anatomical MRI of the brain. Using *FreeSurfer* segmentation and machine-learning techniques, the results showed that anatomical MRI could predict with a 62.4% accuracy NFT stages categorized into three groups (either 0-I-II, III-IV, or V-VI).

4.1. Anatomical considerations

Concordant with our expectations, several regional MRI metrics were significantly negatively correlated to NFT pathology, with lower volumes associated with higher Braak stages. Interestingly, these structures spatially correspond to those where tau is deposited during the course of AD, a pattern consistent with previous findings of regional atrophy reported in a voxel-based morphometry study from Whitwell et al. [16]. Effectively, several temporal, limbic, and parietal structures, as well as a few occipital areas (also known to be affected in late stages of AD) were identified, consistent with well-described Braak staging topography. Less expectedly, the right rostral middle frontal volume also showed significant association with NFT pathology. A similar finding was also reported for the caudal region of the middle frontal gyrus in another study examining atrophic patterns of the frontal-subcortical circuits in mild cognitive impairment and AD subjects [38]. While the left middle frontal gyrus is known to take part in language production, the right middle frontal gyrus is believed to be involved in attention and working memory processes [39], as well as in episodic memory retrieval [40], which is typically impaired in AD. As already described in the literature, ventricular volumes also showed significant positive correlation with neurofibrillary degeneration, with larger lateral ventricles associated with higher Braak stages. This nonspecific MRI finding is observed in several neurodegenerative conditions and most probably reflects the severity of subcortical atrophy.

Cortical surface measurements did not show any significant correlation with NFT deposition in the present study, suggesting that cortical volume loss in AD is mainly due to decreased cortical thickness instead of surface area shrinking. Brain MRI–extracted surfaces are therefore unlikely to be good predictors of underlying AD pathology.

Despite a large number of structural brain metrics identified as being significantly correlated with Braak staging, we found surprisingly low correlation coefficients when compared with widely reported strong associations between volumetric MRI measurements and NFT pathology in the current literature, reaching absolute correlation values of up to 0.80 [8]. Possible explanations include large

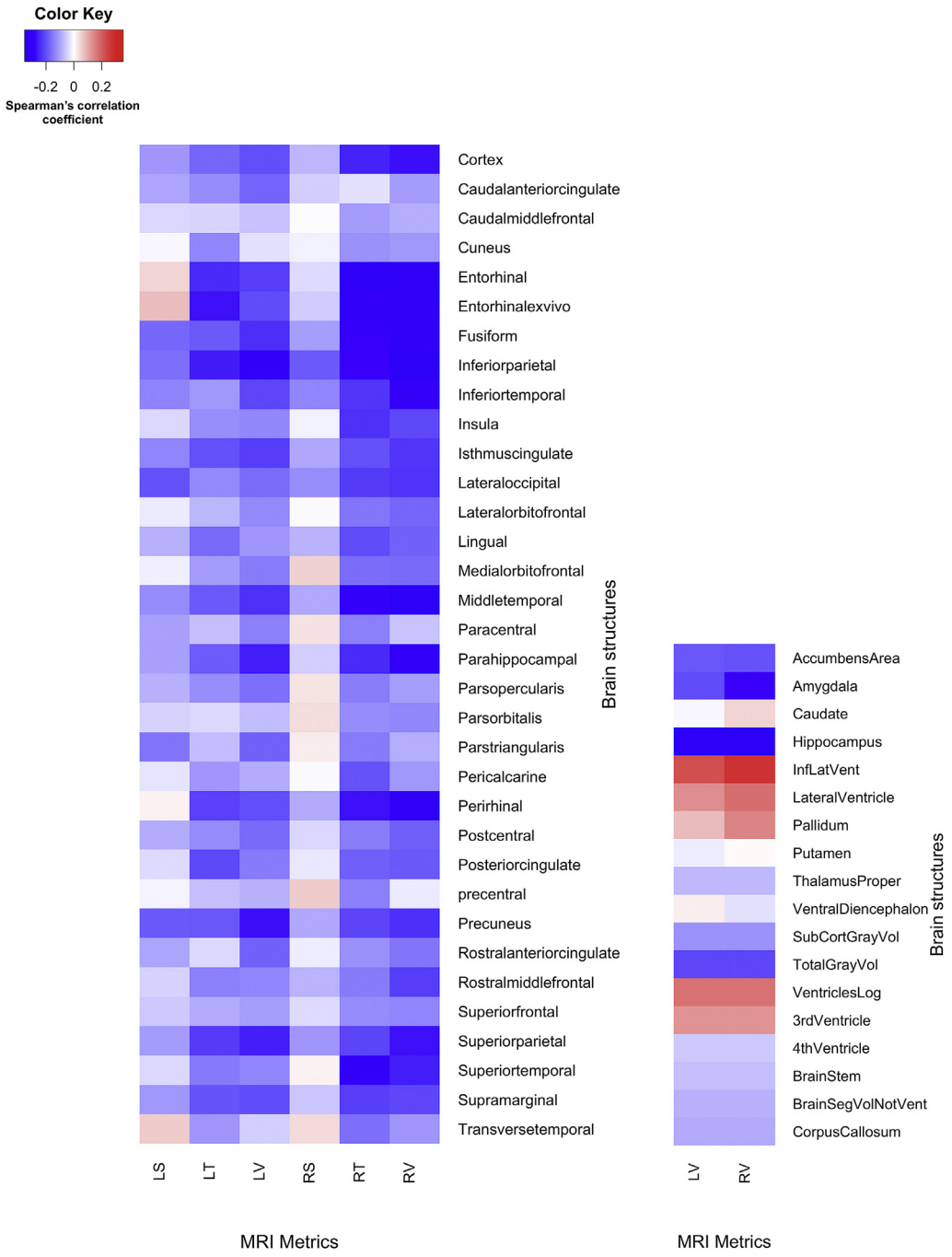

Fig. 2. Spearman's rank correlations heat map.

discrepancies between the three multicenter studies and their respective MRI acquisition and pathological examination protocols, as well as the variability in participants' characteristics from the three study samples.

4.2. *Value of MRI for distinguishing Braak NFT groups*

Accumulating evidence suggests that MRI is an increasingly relevant modality from which to extract biomarkers for the diagnosis of AD on a neurobiological basis in living persons [41]. Based on these assumptions, we developed a support vector machine classifier to predict Braak NFT staging from *in vivo* brain structural MRI metrics. Our classification model provides a global accuracy of 62.4% to distinguish between transentorhinal, limbic, and isocortical Braak stages, corresponding to current recommendations from the worldwide-used "ABC" neuropathological guidelines [2] for tau pathology assessment. This modest predictive

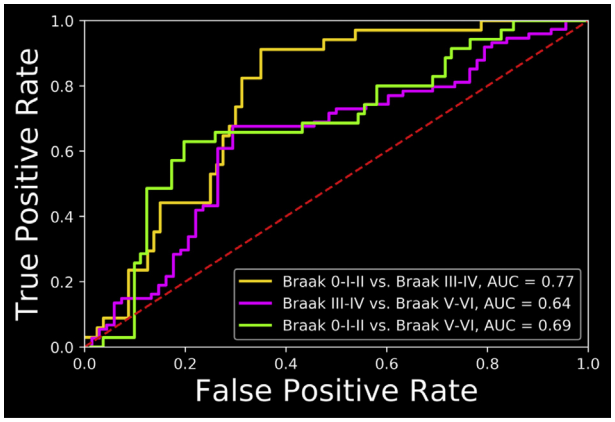

Fig. 3. Receiver operating characteristics curves of the three binary classifiers.

power was expected and can in part be explained by the presence of concomitant pathologies that also result in regional brain atrophy, such as TDP-43 [42,43] and vascular changes [19,23,44,45], and that were not taken into consideration. Indeed, AD is not a one-protein entity and is most often found to present with mixed pathology. Our study is, however, demonstrative of the possibility to predict the neuropathological staging of AD-related neurofibrillary degeneration from cross-sectional MRI scans acquired during life. The present work is therefore in line with recent collaborative, although debated, effort toward a biological definition of AD [46]. Importantly, this achievement was made possible through simple automated segmentation and machine-learning methods for analysis of individual brain MRI scans. With the growing interest in learning-based strategies, our results support the hypothesis that more advanced machine-learning approaches hold promise of higher prediction performance and thus great potential for early detection of AD pathology. In vivo assessment of NFT pathology, the best correlate of both neurodegeneration [8] and cognitive decline [47] in AD, is of particular relevance with regard to the ongoing development of tau-targeted therapies, which could not only benefit in AD, but also in amyloid-negative tauopathies (e.g., primary age-related tauopathy [48]) and mixed pathologies.

4.3. Advantages and clinical significance of in vivo pathological staging of AD

With the rapid aging of populations worldwide, AD is now considered a growing public health and socioeconomic issues, justifying current effort toward the development of quantitative and standardized biomarkers for potential clinical utilization in early detection of AD. Herein, we proposed a new approach to the in vivo automatic stratification of AD pathological changes based on MRI data. Similar imaging-based techniques could provide additional reliable information to standard neurological and cognitive examination in the clinical workup of dementia.

Table 2
Performance parameters of the three binary classifiers

Classifier	Sensitivity (%)	Specificity (%)	Accuracy (%)	AUC
Braak 0-I-II versus Braak III-IV	50.0	82.9	70.2	0.77
Braak III-IV versus Braak V-VI	71.4	66.7	69.0	0.64
Braak 0-I-II versus Braak V-VI	52.3	83.3	71.6	0.69

Abbreviation: AUC, area under the receiver operating characteristics curve.

The avenue of accurate diagnosis tools raises the possibility to ensure earlier and more adequate patient care and management, including initiation of current symptomatic pharmacological treatments (i.e., acetylcholinesterase inhibitors and NMDA receptors antagonists) in their therapeutic window of optimal efficiency, discussing care planning and medico-legal considerations with the patient while they retain the cognitive ability to do so, and refer them and their caregivers to dedicated support groups. Notably, automatic imaging diagnosis tools could allow for a better selection of participants eligible for clinical trials of new disease-modifying targeted therapies before they have crossed the step of irreversible brain tissue damage. The use of in vivo MRI, already recommended as an investigative method of choice to rule out other conditions associated with cognitive decline, could potentially provide additional standardized inclusion information indicating the presence of AD neuropathological changes, as demonstrated in the present study.

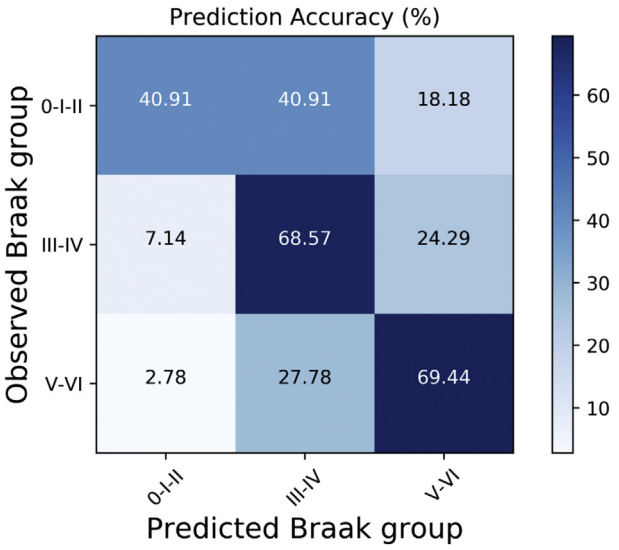

Fig. 4. Normalized confusion matrix. This classification table shows the performance of the multiclass model to discriminate between NFT pathological groups within our test set. Abbreviations: AUC, area under the receiver operating characteristics curve; NFT, neurofibrillary tangles.

4.4. Strengths and potential limitations of the study

Our study design presents multiple strengths. First, our sample was selected through multiple independent data sets from prospective multicentric studies with extensive examination of clinical, radiological, and neuropathological parameters. Considering that individuals with postmortem examination and premortem MRI data are very scare, we presented here one of the largest sample size in radiological-pathological studies to our knowledge. Although recruitment of patients from memory clinics can result in some selection biases, the MAP data set was collected from a population-based study cohort. Moreover, the entire cognitive spectrum (i.e., from cognitively normal to clinical AD) was covered in our study, allowing to generalize the prediction to any cognitive state.

The use of MRI modality represents a safe, well tolerate, precise, and widely available imaging technique. We provide here an extensive brain regional analysis which is necessary to increase accuracy of diagnostic tools at the individual level. Moreover, we propose a method based on automatic segmentation with a well-developed worldwide-used validated tool (i.e., FreeSurfer), which was demonstrated to afford more reproducible brain mapping with decreased intra- and inter-rater variability when compared with manual brain segmentation and hand-drawn regions of interest [49,50]. Furthermore, this study was designed using nonlinear model classification to provide a more comprehensive neuroimaging analysis approach which best captures the subtleties of the brain abnormalities for neuropathological stratification of NFTs associated with AD. We also trained our model using 10-fold cross-validation to ensure that a correct classification rate was evaluated on MRI scans not previously used by the classifier and feature selector. This way, we provide more reliable estimates of generalization and predictive value.

The present work also presents several limitations and potential biases. Among these, our analyses were conducted on a sample from a North American population, and thus a presumed majority of Caucasians. Our results may therefore not apply to a more ethnically diversified population. Regarding postmortem assessment, there is a part of subjectivity existing despite well-defined international guidelines [51,52]. A possible asymmetry of pathological lesions distribution and brain dysfunction across the two cerebral hemispheres has also been suggested by some authors [53–57]. Because the examined hemisphere lateralization was not always the same in the three cohorts, aleatory pathological grading may have affected our predictive ability. MRI volumetry can also be influenced by magnetic field strength and scanner manufacturer. However, our results showed that including these variables in the models did not improved prediction. This is consistent with our previous results showing that these variables explain a significant, but minor amount of variance: only 1%, 3%, and 7% of the total variance for cortical surface, volume, and thickness, respectively [58]. Furthermore, we did not examine the ability of our MRI-based SVM classifier to distinguish between AD subtypes [18] or atypical dementias. Notably, some subjects included in this study may also have tau-associated conditions other than typical AD (e.g., AD variants, primary progressive aphasia, primary age-related tauopathy) or comorbid pathology (e.g., mixed dementia), which are known to be associated with different patterns of regional brain atrophy. Because of the limited number of participants with Braak 0 available in our three original cohorts (n = 9), we were unable to consider a group pathologically free in our predictive model. Finally, classification accuracy may have been affected by our cross-sectional study design, with variable interval time between last MRI and death that did not take into consideration the variable rate of disease progression and could not properly catch the dynamic process of brain atrophy observed in neurodegenerative diseases.

5. Conclusions

In summary, our findings suggest that Braak NFT staging can be predicted to some extent from in vivo cross-sectional MRI metrics using automatic brain segmentation and classification-based machine-learning techniques. These results are encouraging regarding the development of a supportive diagnostic tool for AD on an individual patient basis and a potential read-out for tau-focused interventions. The present study and recent advances will hopefully motivate future research on the use of structural MRI markers for identification of dementia-related pathological changes in the living brain.

Acknowledgments

Authors' contributions: C.D.-T. conceptualized the study, analyzed the data, interpreted the results, and wrote the manuscript. I.B. provided his expertise for machine-learning analyses of the data and contributed to redaction. L.D. contributed to data collection and image processing. S.S. provided manuscript revision for important intellectual content regarding neuropathological considerations. O.P. and S.D. conceptualized the study, supervised the study process, and revised the manuscript for important intellectual content. The authors gratefully thank the participants in the Alzheimer's Disease Neuroimaging Initiative, the National Alzheimer's Coordinating Center, and the Rush Memory and Aging Project.

This research project was financially supported by a grant to S.D. from the Alzheimer's Society of Canada (#13-32). C.D.-T. is supported by a Frederick Banting and Charles Best Canada Graduate Scholarship Doctoral Award from the Canadian Institutes of Health Research (#406235) and by a Scholarship from the Fonds de recherche et de développement sur la maladie d'Alzheimer et les maladies apparentées–Fondation de la famille Lemaire. S.D. is a

Research Scholar from the Fonds de recherche du Québec–Santé (#30801).
ADNI 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.
The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI M. Marsel Mesulam, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).
The MAP and its imaging data were funded by NIH grants R01 AG17917, R01 AG10039 and K23 AG40652 awarded to D.A. Bennett. The MAP data can be requested online from the following link: www.radc.rush.edu.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.dadm.2019.07.001.

RESEARCH IN CONTEXT

1. Systematic review: The authors reviewed the literature using PubMed database. As the definitive diagnosis of Alzheimer's disease (AD) classically requires postmortem assessment of key neuropathological features, our work is in line with current collective effort toward the development of reliable biomarkers to detect pathological changes in preclinical stages of the disease. Early diagnosis of AD is a strongly represented research field in the literature, as cited in the present work.
2. Interpretation: Our findings suggest that structural MRI can potentially be used as a supportive biomarker for early diagnosis of AD in the living brain by predicting neurofibrillary degeneration.
3. Future directions: This article highlights the need for additional studies focusing on predictive modeling with more extensive training to provide higher classification performance. Similar studies based on longitudinal data could potentially catch the spatiotemporal patterns of brain atrophy associated with neurofibrillary tangles deposition. Association with molecular imaging classifiers could also be of extreme relevance for future clinical application.

References

[1] Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA. Religious Orders Study and Rush Memory and Aging Project. J Alzheimers Dis 2018;64:S161–89.
[2] Montine TJ, Phelps CH, Beach TG, Bigio EH, Cairns NJ, Dickson DW, et al. National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. Acta Neuropathol 2012; 123:1–11.

[3] Thal DR, Rub U, Orantes M, Braak H. Phases of A beta-deposition in the human brain and its relevance for the development of AD. Neurology 2002;58:1791–800.
[4] Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta Neuropathol 1991;82:239–59.
[5] Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology 1991;41:479–86.
[6] Hof PR, Glannakopoulos P, Bouras C. The neuropathological changes associated with normal brain aging. Histol Histopathol 1996; 11:1075–88.
[7] Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012;367:795–804.
[8] Dallaire-Theroux C, Callahan BL, Potvin O, Saikali S, Duchesne S. Radiological-pathological correlation in Alzheimer's disease: systematic review of antemortem magnetic resonance imaging findings. J Alzheimers Dis 2017;57:575–601.
[9] Braak H, Alafuzoff I, Arzberger T, Kretzschmar H, Del Tredici K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. Acta Neuropathol 2006;112:389–404.
[10] Jack CR Jr, Dickson DW, Parisi JE, Xu YC, Cha RH, O'Brien PC, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. Neurology 2002;58:750–7.
[11] Silbert LC, Quinn JF, Moore MM, Corbridge E, Ball MJ, Murdoch G, et al. Changes in premorbid brain volume predict Alzheimer's disease pathology. Neurology 2003;61:487–92.
[12] Csernansky JG, Hamstra J, Wang L, McKeel D, Price JL, Gado M, et al. Correlations between antemortem hippocampal volume and postmortem neuropathology in AD subjects. Alzheimer Dis Assoc Disord 2004;18:190–5.
[13] Zarow C, Vinters HV, Ellis WG, Weiner MW, Mungas D, White L, et al. Correlates of hippocampal neuron number in Alzheimer's disease and ischemic vascular dementia. Ann Neurol 2005;57:896–903.
[14] Josephs KA, Whitwell JL, Ahmed Z, Shiung MM, Weigand SD, Knopman DS, et al. Beta-amyloid burden is not associated with rates of brain atrophy. Ann Neurol 2008;63:204–12.
[15] Vemuri P, Whitwell JL, Kantarci K, Josephs KA, Parisi JE, Shiung MS, et al. Antemortem MRI based STructural Abnormality iNDex (STAND)-scores correlate with postmortem Braak neurofibrillary tangle stage. Neuroimage 2008;42:559–67.
[16] Whitwell JL, Josephs KA, Murray ME, Kantarci K, Przybelski SA, Weigand SD, et al. MRI correlates of neurofibrillary tangle pathology at autopsy: a voxel-based morphometry study. Neurology 2008; 71:743–9.
[17] Burton EJ, Barber R, Mukaetova-Ladinska EB, Robson J, Perry RH, Jaros E, et al. Medial temporal lobe atrophy on MRI differentiates Alzheimer's disease from dementia with Lewy bodies and vascular cognitive impairment: a prospective study with pathological verification of diagnosis. Brain 2009;132:195–203.
[18] Whitwell JL, Dickson DW, Murray ME, Weigand SD, Tosakulwong N, Senjem ML, et al. Neuroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study. Lancet Neurol 2012;11:868–77.
[19] Kaur B, Himali JJ, Seshadri S, Beiser AS, Au R, McKee AC, et al. Association between neuropathology and brain volume in the Framingham Heart Study. Alzheimer Dis Assoc Disord 2014;28:219–25.
[20] Erten-Lyons D, Woltjer RL, Dodge H, Nixon R, Vorobik R, Calvert JF, et al. Factors associated with resistance to dementia despite high Alzheimer disease pathology. Neurology 2009;72:354–60.
[21] Kantarci K, Lowe VJ, Boeve BF, Weigand SD, Senjem ML, Przybelski SA, et al. Multimodality imaging characteristics of dementia with Lewy bodies. Neurobiol Aging 2012;33:2091–105.
[22] Thaker AA, Weinberg BD, Dillon WP, Hess CP, Cabral HJ, Fleischman DA, et al. Entorhinal cortex: antemortem cortical thickness and postmortem neurofibrillary tangles and amyloid pathology. AJNR Am J Neuroradiol 2017;38:961–5.
[23] Jagust WJ, Zheng L, Harvey DJ, Mack WJ, Vinters HV, Weiner MW, et al. Neuropathological basis of magnetic resonance images in aging and dementia. Ann Neurol 2008;63:72–80.
[24] Kaye JA, Swihart T, Howieson D, Dame A, Moore MM, Karnos T, et al. Volume loss of the hippocampus and temporal lobe in healthy elderly persons destined to develop dementia. Neurology 1997; 48:1297–304.
[25] Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. Neurobiol Aging 2000;21:19–26.
[26] Killiany RJ, Gomez-Isla T, Moss M, Kikinis R, Sandor T, Jolesz F, et al. Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease. Ann Neurol 2000;47:430–9.
[27] Dickerson BC, Goncharova I, Sullivan MP, Forchetti C, Wilson RS, Bennett DA, et al. MRI-derived entorhinal and hippocampal atrophy in incipient and very mild Alzheimer's disease. Neurobiol Aging 2001;22:747–54.
[28] Chetelat G, Desgranges B, De La Sayette V, Viader F, Eustache F, Baron JC. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. Neuroreport 2002;13:1939–43.
[29] Medina D, DeToledo-Morrell L, Urresta F, Gabrieli JD, Moseley M, Fleischman D, et al. White matter changes in mild cognitive impairment and AD: a diffusion tensor imaging study. Neurobiol Aging 2006;27:663–72.
[30] Resnick SM, Goldszal AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, et al. One-year age changes in MRI brain volumes in older adults. Cereb Cortex 2000;10:464–72.
[31] Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling protocol. Front Neurosci 2012;6:171.
[32] Fischl B, Stevens AA, Rajendran N, Yeo BT, Greve DN, Van Leemput K, et al. Predicting the location of entorhinal cortex from MRI. Neuroimage 2009;47:8–17.
[33] Augustinack JC, Huber KE, Stevens AA, Roy M, Frosch MP, van der Kouwe AJ, et al. Predicting the location of human perirhinal cortex, Brodmann's area 35, from MRI. Neuroimage 2013;64:32–42.
[34] Nagy Z, Yilmazer-Hanke DM, Braak H, Braak E, Schultz C, Hanke J. Assessment of the pathological stages of Alzheimer's disease in thin paraffin sections: a comparative study. Dement Geriatr Cogn Disord 1998;9:140–4.
[35] Mesulam M. Principles of behavioral and cognitive neurology. 2nd ed. New York, NY, US: Oxford University Press; 2000.
[36] Pohjalainen J RO, Kadioglu S. Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits. Computer Speech Lang 2015; 29:145–71.
[37] Beheshti I, Demirel H, Farokhian F, Yang C, Matsuda H, Alzheimer's Disease Neuroimaging Initiative. Structural MRI-based detection of Alzheimer's disease using feature ranking and classification error. Comput Methods Programs Biomed 2016;137:177–93.
[38] Zhao H, Li X, Wu W, Li Z, Qian L, Li S, et al. Atrophic patterns of the frontal-subcortical circuits in patients with mild cognitive impairment and Alzheimer's disease. PLoS One 2015;10:e0130017.
[39] Japee S, Holiday K, Satyshur MD, Mukai I, Ungerleider LG. A role of right middle frontal gyrus in reorienting of attention: a case study. Front Syst Neurosci 2015;9:23.
[40] Rajah MN, Languay R, Grady CL. Age-related changes in right middle frontal gyrus volume correlate with altered episodic retrieval activity. J Neurosci 2011;31:17941–54.
[41] Chandra A, Dervenoulas G, Politis M, Alzheimer's Disease Neuroimaging Initiative. Magnetic resonance imaging in Alzheimer's disease and mild cognitive impairment. J Neurol 2018;266:1293–302.
[42] Josephs KA, Murray ME, Whitwell JL, Parisi JE, Petrucelli L, Jack CR, et al. Staging TDP-43 pathology in Alzheimer's disease. Acta Neuropathol 2014;127:441–50.

[43] Josephs KA, Dickson DW, Tosakulwong N, Weigand SD, Murray ME, Petrucelli L, et al. Rates of hippocampal atrophy and presence of post-mortem TDP-43 in patients with Alzheimer's disease: a longitudinal retrospective study. Lancet Neurol 2017;16:917–24.

[44] Erten-Lyons D, Dodge HH, Woltjer R, Silbert LC, Howieson DB, Kramer P, et al. Neuropathologic basis of age-associated brain atrophy. JAMA Neurol 2013;70:616–22.

[45] Raman MR, Preboske GM, Przybelski SA, Gunter JL, Senjem ML, Vemuri P, et al. Antemortem MRI findings associated with microinfarcts at autopsy. Neurology 2014;82:1951–8.

[46] Jack CR Jr, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. Alzheimers Dement 2018; 14:535–62.

[47] Nelson PT, Alafuzoff I, Bigio EH, Bouras C, Braak H, Cairns NJ, et al. Correlation of Alzheimer disease neuropathologic changes with cognitive status: a review of the literature. J Neuropathol Exp Neurol 2012; 71:362–81.

[48] Crary JF, Trojanowski JQ, Schneider JA, Abisambra JF, Abner EL, Alafuzoff I, et al. Primary age-related tauopathy (PART): a common pathology associated with human aging. Acta Neuropathol 2014; 128:755–66.

[49] Fischl B. Freesurfer. Neuroimage 2012;62:774–81.

[50] Mulder ER, de Jong RA, Knol DL, van Schijndel RA, Cover KS, Visser PJ, et al. Hippocampal volume change measurement: quantitative assessment of the reproducibility of expert manual outlining and the automated methods FreeSurfer and FIRST. Neuroimage 2014; 92:169–81.

[51] Mirra SS, Gearing M, McKeel DW Jr, Crain BJ, Hughes JP, van Belle G, et al. Interlaboratory comparison of neuropathology assessments in Alzheimer's disease: a study of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). J Neuropathol Exp Neurol 1994;53:303–15.

[52] Bancher C, Paulus W, Paukner K, Jellinger K. Neuropathologic diagnosis of Alzheimer disease: consensus between practicing neuropathologists? Alzheimer Dis Assoc Disord 1997;11:207–19.

[53] Foster NL, Chase TN, Fedio P, Patronas NJ, Brooks RA, Di Chiro G. Alzheimer's disease: focal cortical changes shown by positron emission tomography. Neurology 1983;33:961–5.

[54] Ishii K, Sasaki H, Kono AK, Miyamoto N, Fukuda T, Mori E. Comparison of gray matter and metabolic reduction in mild Alzheimer's disease using FDG-PET and voxel-based morphometric MR studies. Eur J Nucl Med Mol Imaging 2005;32:959–63.

[55] Hunt A, Schonknecht P, Henze M, Seidl U, Haberkorn U, Schroder J. Reduced cerebral glucose metabolism in patients at risk for Alzheimer's disease. Psychiatry Res 2007;155:147–54.

[56] Fan Y, Resnick SM, Wu X, Davatzikos C. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. Neuroimage 2008;41:277–85.

[57] Stefanits H, Budka H, Kovacs GG. Asymmetry of neurodegenerative disease-related pathologies: a cautionary note. Acta Neuropathol 2012;123:449–52.

[58] Potvin O, Dieumegarde L, Duchesne S, Alzheimer's Disease Neuroimaging Initiative. Normative morphometric data for cerebral cortical areas over the lifetime of the adult human brain. Neuroimage 2017; 156:315–39.