Exploring the power of MRI and clinical measures in predicting Alzheimer's disease neuropathology

Running title: Predicting postmortem neuropathology

Farooq Kamal¹², Cassandra Morrison³, Michael D. Oliver⁴⁵, Mahsa Dadar¹²
¹Department of Psychiatry, McGill University, Montreal, Quebec, H3A 1A1, Canada
²Douglas Mental Health University Institute, Montreal, Quebec, H4H 1R3, Canada
³Department of Psychology, Carleton University, Ottawa, Ontario, K1S 5B6, Canada
⁴Department of Psychological Science and Neuroscience, Belmont University, Nashville, Tennessee, TN 37212, United States
⁵Belmont Data Collaborative, Belmont University, Nashville, Tennessee, TN 37212, United States

‡ 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

Correspondence to: Farooq Kamal and Mahsa Dadar
Douglas Mental Health University Institute, Montreal, Quebec, Canada, H4H 1R3
E-mail: farooq.kamal@mail.mcgill.ca, mahsa.dadar@mcgill.ca

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background: The ability to predict Alzheimer's disease (AD) before diagnosis is a topic of intense research. Early diagnosis would aid in improving treatment and intervention options, however, there are no current methods that can accurately predict AD years in advance. This study examines a novel machine learning approach that integrates the combined effects of vascular (white matter hyperintensities, WMHs), and structural brain changes (gray matter, GM) with clinical factors (cognitive status) to predict post-mortem neuropathological outcomes.

Methods: Healthy older adults, participants with mild cognitive impairment, and AD from the Alzheimer's Disease Neuroimaging Initiative dataset with both post-mortem neuropathology data and antemortem MRI and clinical data were included. Longitudinal data were analyzed across three intervals before death (post-mortem data): 0-4 years, 4-8 years, and 8-14 years. Additionally, cross-sectional data at the last visit or interval (within four years, 0-4 years) before death were also examined. Machine learning models including gradient boosting, bagging, support vector regression, and linear regression were implemented. These models were applied towards feature selection of the top seven MRI, clinical, and demographic data to identify the best performing set of variables that could predict postmortem neuropathology outcomes (i.e., neurofibrillary tangles, neuritic plaques, diffuse plaques, senile/amyloid plaques, and amyloid angiopathy).

Results: A total of 94 participants (55-90 years of age) were included in the study. At last visit, the best-performing model included total and temporal lobe WMHs and achieved $r=0.87$ ($RMSE=0.62$) during cross-validation for neuritic plaques. For longitudinal assessments across different intervals, the best-performing model included regional GM (i.e., hippocampus, amygdala, caudate) and frontal lobe WMH and achieved $r=0.93$ ($RMSE=0.59$) during cross-validation for neurofibrillary tangles. For MRI and clinical predictors and clinical-only predictors, $t$-tests demonstrated significant differences at all intervals before death ($t[-13.60-7.90]$, $p$-values<0.001). Overall, post-mortem neuropathology outcome were predicted up to 14 years before death with high accuracies (~90%).

Conclusions: Prediction accuracy was higher for post-mortem neuropathology outcomes that included MRI (WMHs, GM) and clinical features compared to clinical-only features. These findings highlight that MRI features are critical to successfully predict AD-related pathology years in advance which will improve participant selection for clinical trials, treatments, and intervention options.

Keywords: Alzheimer's Disease, Machine Learning, Predictive Modeling, Neuropathology, MRI, White matter hyperintensity
Background

Alzheimer’s disease (AD) is a neurodegenerative disease characterized by a progressive decline in cognitive functioning. AD is characterized by pathological accumulation of neurofibrillary tau tangles and amyloid-beta deposits [1]. In addition to these pathological changes, structural brain alterations, including gray matter (GM) and white matter (WM) atrophy, and white matter hyperintensity (WMH) burden contribute to cognitive decline in AD [2–5]. For example, GM changes (particularly volume reduction) are closely associated with cognitive impairment in AD, highlighting their importance in the neurodegeneration process [6–8]. Research has also shown a strong relationship between WM volume changes and cognitive impairment across the spectrum of AD progression [9]. Similarly, WMHs are markers of cerebrovascular disease burden and are increasingly recognized for their role in cognitive decline and the progression of neurodegenerative diseases [10–12].

People with AD often present with both structural and cerebrovascular changes such as GM and WM volume loss and WMH burden at autopsy [13–16]. These volumetric changes are commonly studied in the entorhinal cortex [6] and hippocampus [7,8] as they are two of the initial brain regions to show abnormal volume loss in AD, thus providing information about AD staging/progression. Although neurodegeneration has been included as a marker for staging AD, the Alzheimer’s Association Workgroup has recently proposed a revised criteria for diagnosing and staging of AD. The updated guidelines now also include vascular brain injury (often measured using MRI markers of cerebrovascular disease) as a common non-AD co-pathology [1], further highlighting the prevalence of vascular pathology in AD. Taking markers of both structural and cerebrovascular pathology into consideration, the present study employs a novel approach to examine whether these in vivo abnormalities are predictive of neuropathology severity.
Previous neuropathological studies have provided insights into early risk factors for AD by exploring the correlations between neuropathological and clinical or brain volume measures individually. For instance, Kaur et al. [17] explored the association between neuropathology (e.g., neurofibrillary tangle burden (NFT)) and total and regional brain volume and reported significant correlations between neuropathological measures and brain volume. Additionally, a recent study by Zhang et al. [18] further explored the relationship between clinical and neuropsychological measures and neuropathological data in cognitively healthy older adults. Neuropathological measures such as Braak-NFT stage were significantly correlated with clinical measures of memory, executive function, and visuospatial function. These studies predominantly concentrate on either clinical or MRI-based markers that correlate with neuropathology [17–20]. However, these studies did not investigate the potential of using a combination of WMH, WM, and GM volumes as features to predict neuropathology. Thus, a comprehensive approach that combines these features to predict neuropathological outcomes remains unexplored.

The integration of machine learning (ML) has opened new avenues for understanding and diagnosing neurodegenerative diseases. With their ability to handle large datasets and complex nonlinear patterns, ML models have been widely utilized for disease diagnosis and prognosis applications [21]. For example, support vector machines (SVMs) have been used to classify different stages of AD using MRI data, helping to outline disease progression [22]. Similarly, linear regression has been instrumental in identifying biomarkers and predicting disease outcomes in neurodegenerative conditions [18]. Despite these advances, accurately predicting neuropathology remains a significant challenge that has not been previously addressed. Early prediction of neuropathology remains an essential component in AD research because it would aid in improving selection of participants for clinical trials and enhance treatment options. For example, AD-related brain changes occur 10-20 years before
diagnosis occurs, and currently there are no methods that can accurately predict AD that early. If researchers develop a method that could predict who would progress to AD, it may be possible to intervene earlier in the disease course and prevent the irreversible brain changes that occur.

Current research has yet to fully incorporate vascular (WMHs) and structural brain changes (WM and GM), with clinical factors in ML models for neuropathology prediction. Given the new criteria for diagnosis and staging\(^1\), this gap underscores a critical area for exploration, as WMHs represent a potentially valuable marker for early disease detection that is yet to be fully examined. Given the limitations of the existing research, the present study aims to integrate these features with other AD-related brain measures to predict post-mortem neuropathology years in advance. Additionally, the study will compare the importance of MRI measures in addition to clinical measures. To achieve this, ML methods will be applied to determine the importance of different features (e.g., cognitive, regional and global brain volumes, and WMHs) in predicting AD-related neuropathologies. We hypothesize that in addition to typical MRI measures (i.e., brain volume) such as WM and GM, WMH burden will also be a critical feature that contributes to AD-related post-mortem neuropathology prediction.

**Methods**

**Participants**

Data were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Participants were included if they had post-mortem neuropathology information, MRIs from which regional white and gray matter and WMH volume could be extracted, completed the Clinical Dementia Rating- Sum of Boxes (CDR-SB, a measure of global cognition), digit span (as a measure of working memory), trail making test - part B (as
a measured of processing speed), everyday cognition scale (as a measure of subjective cognitive decline), and logical memory (as a measure of episodic memory). MRI, clinical, and demographic data were collected from participants over time (i.e., longitudinally) before their death (when post-mortem neuropathology information was obtained). A total of 94 participants (55-90 years of age) were included in the study. There were 59 participants with data within 4 years prior to death, 63 participants with data from 4 to 8 years before death, and 28 participants with data between 8 to 14 years before death. Summary of demographic information for participants in ADNI are presented in Table 1.

< Insert Table 1 here>

**Structural MRI acquisition and processing of WMH in ADNI**

Scans were downloaded from the ADNI website (see [http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/](http://adni.loni.usc.edu/methods/mri-tool/mri-analysis/)) for MRI acquisition protocol). T1w scans were pre-processed through our standard pipeline including noise reduction [23], intensity inhomogeneity correction [24], and intensity normalization into range [0-100]. The pre-processed images were linearly (9 parameters: 3 translation, 3 rotation, and 3 scaling) [25] registered to the MNI-ICBM152-2009c average [26].

A previously validated WMH segmentation technique that has been extensively tested in aging and neurodegenerative diseases was used to obtain WMH measurements [27]. This technique has been previously employed in other multi-center studies [28,29] and has also been validated in the ADNI cohort [30]. The automated WMH segmentation technique extracts a set of location (i.e. spatial priors) and intensity (distribution histograms) features and uses them in combination with a random forest classifier to detect the WMHs in new images [27, 28, 29]. The method has been trained based on a library of manually segmented scans from the ADNI data (50 participants independent of the ones studied here), ensuring its
accuracy when applied to the current dataset. Automatic segmentation of the WMHs was
completed using only the T1w contrasts. The decision to use only T1w images for WMH
segmentation was made since ADNI1 study only includes T1w, and T2w/PD images, whereas
ADNI2/GO includes T1w and axial 2D FLAIR images, and ADNI3 includes T1w sagittal 3D
FLAIRs. Since the resolution of the images (i.e., voxel sizes) were very different between
T2w/PD (1*1*3 mm³) and FLAIRs (0.85*0.85*5 mm³ and 1*1*1.2 mm³), to be able to
reliably combine and use WMH volumes from all these studies, we only used the T1w images
for WMH segmentation. We have previously validated the performance of our pipeline in
detecting WMHs based on T1w images both in ADNI and another independent dataset, and
have shown that the T1w-based WMH volumes (1) hold strong correlations with T2w/PD and
FLAIR based WMH volumes ($r = .89$, $p < .001$ for ADNI1, and $r = .97$, $p < .001$
for ADNI2/GO), and (2) have similar relationships with clinical/cognitive scores and risk
factors associated with WMHs. 34 Finally, the quality of all WMH segmentations was visually
assessed by M.D. (blinded to clinical diagnosis). WMH load was defined as the volume of all
voxels identified as WMH in the standard space (in mm³) and were normalized for head size.
Regional (frontal, temporal, occipital, and parietal) and total WMH volumes were calculated
based on Hammers Atlas [27,31]. All WMH volumes were also log-transformed to achieve
normal distribution.

**Freesurfer Measurements**

T1w images were processed using FreeSurfer and quality controlled by the UCSF group, and
total WM as well as total and regional GM volumes were extracted. 1.5T and 3T data were
processed with FreeSurfer versions 4.3 and 5.1, respectively, as appropriate [32]. Regional
GM features examined included the hippocampus, entorhinal cortex, parahippocampal gyrus,
fusiform gyrus, and middle temporal gyrus, amygdala, putamen, caudate, and thalamus volumes.

**Predictive modelling with multiple feature sets in ADNI**

Analyses were performed using Python software version 3.11. The predictor variables included global cognition, episodic memory, subjective cognitive decline (SCD), perceptual speed, perceptual orientation, working memory, and structural brain volumes (WM, total and regional GM and WMHs). These predictors were obtained based on the last available visit, within 4 years prior to death (0-4 years) and therefore included 1 timepoint per person. Neurofibrillary tangles (classified by Braak stage), neuritic plaques, and diffuse plaques (classified according to CERAD), amyloid plaques (classified according to Thal amyloid phase), and amyloid angiopathy were the outcomes of interest for the prediction models.

All continuous values were standardized using the `StandardScaler` from the `sklearn.preprocessing` package in Python prior to analyses, aligning with best practices in machine learning research [33–35]. We incorporated a range of linear and nonlinear machine learning models, including gradient boosting classifier and regressor, support vector machines, linear regression, and bagging techniques. These models were selected due to their robustness in handling diverse datasets. To evaluate the model performance, we implemented a 10-fold cross-validation (CV) scheme using KFold from `sklearn.model_selection`, minimizing bias and overfitting [36]. Each model within our framework was trained on a subset of the predictors, ranging from individual features to combinations of up to seven features. We evaluated models based on their performance in predicting continuous outcomes.
For continuous outcomes, we measured root mean square error (RMSE) and correlation. These metrics provided a comprehensive assessment of model performance.

We examined combinations of predictors using `itertools.combinations`, testing each possible combination to identify the most informative feature set for predicting neuropathology outcomes. To further confirm the validity of our initial results, we employed a bootstrap method, which involved resampling our dataset and performing the same k-fold evaluation on each bootstrap sample. We conducted 50 bootstrap iterations, each involving a 10-fold cross-validation process. This technique, implemented using the resample function, is increasingly recognized for its efficacy in assessing model uncertainty and variance [37,38].

Regression plots were used to visualize and evaluate the predictive performance of machine learning models by comparing neuropathology outcome values against the cross-validated predicted neuropathology values. Higher R² values indicated a better fit of the model to the data, while lower root mean squared error (RMSE) values suggested higher accuracy in predictions.

Longitudinal data at different time intervals before death (0-4 years, 4-8 years, and 8-14 years) were also examined for all predictor variables across the different neuropathological outcomes to investigate the capabilities of the ML models in predicting future pathology. Longitudinal data included multiple timepoints per person. To evaluate model performance in the longitudinal data, we utilized grouped k-fold cross-validation, which stratifies the data based on participants. This method ensures that the same participant does not appear in both training and test sets within a fold, thereby preventing data leakage and ensuring robust model evaluation. Similar modelling procedures were used to examine the longitudinal data. T-tests were used to assess whether the model performances were statistically different across time-to-death groups. The results were corrected for multiple comparisons using false discovery rate (FDR) controlling method.
Results

The most informative predictors of neuropathology outcomes across each model are presented in Table 2.

< Insert Table 2 here>

Neuropathology Prediction

For neuritic plaques, the SVM model, which included total WMH, Thalamus, Temporal Lobe WMH, Episodic memory, SCD, APOE genotype, and Sex, achieved the highest correlation of 0.87 with an RMSE of 0.62 (Fig. 1A). This model outperformed others in predictive accuracy for neuritic plaques. The Gradient Boosting model included predictors such as Amygdala, Ventricles, Global Cognition, Working memory, Episodic memory, SCD, and APOE genotype, demonstrating a high correlation of 0.76 with the lowest RMSE of 0.85. The Bagging model, incorporating predictors like Amygdala, Temporal Lobe WMH, Working memory, Episodic memory, SCD, APOE genotype, and Age at visit, also showed good predictive performance with a correlation of 0.75 and an RMSE of 0.84. In contrast, the Linear Regression model, which included the Thalamus, Frontal Lobe WMH, Parietal Lobe WMH, Episodic memory, SCD, APOE Genotype, and Sex, had a lower correlation of 0.73 and a higher RMSE of 0.86.

In predicting diffuse plaques, the Gradient Boosting model used Hippocampus, Amygdala, Ventricles, Working memory, Episodic memory, SCD, and Sex, showing the strongest performance with an RMSE of 0.63 and a correlation of 0.78 (Fig. 1B). The Bagging model included Total WMH, Amygdala, Ventricles, Processing speed, Working memory, Episodic memory, and SCD, achieving an RMSE of 0.66 and a correlation of 0.73.
The Linear Regression model incorporated the Hippocampus, Putamen, Processing speed, Episodic memory, SCD, APOE Genotype, and Sex, with an RMSE of 0.84 and a correlation of 0.49. The SVM model, with Hippocampus, Amygdala, Ventricle, Temporal Lobe WMH, Working memory, SCD, and APOE genotype, showed an RMSE of 0.66 and a correlation of 0.74.

For neurofibrillary tangles, the Gradient Boosting model included Amygdala, Temporal Lobe WMH, Processing speed, Episodic memory, SCD, APOE genotype, and Sex, achieving a correlation of 0.70 and an RMSE of 1.23. The Bagging model included the Amygdala, Parahippocampal, Temporal Lobe WMH, Processing speed, Episodic memory, SCD, and APOE genotype, with a correlation of 0.69 and an RMSE of 1.23. The Linear Regression model incorporated the Thalamus, Temporal Lobe WMH, Parietal Lobe WMH, Episodic memory, SCD, APOE genotype, and Sex, showing a correlation of 0.70 and the same RMSE of 1.23. The SVM model incorporated Hippocampus, Thalamus, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, achieving a correlation of 0.81 and an RMSE of 1.02 (Fig. 1C).

In predicting amyloid plaques, the Gradient Boosting model included the Ventricle, Thalamus, Global Cognition, Processing speed, Episodic memory, SCD, and APOE genotype, achieving a correlation of 0.78 and an RMSE of 0.98 (Fig. 1D). The Bagging model incorporated the Ventricle, Caudate, Frontal Lobe WMH, Global Cognition, Working memory, Episodic memory, SCD, achieving a correlation of 0.72 and an RMSE of 1.09. The Linear Regression model included Amygdala, Caudate, Thalamus, Episodic memory, SCD, APOE genotype, and Sex, resulting in a correlation of 0.65 and an RMSE of 1.16. The SVM model incorporated Caudate, Thalamus, Frontal Lobe WMH, Occipital Lobe WMH, Episodic memory, SCD, and APOE genotype, with a correlation of 0.76 and an RMSE of 1.03.
For amyloid angiopathy, the Gradient Boosting model included the Ventricles, Caudate, Parahippocampal, Working memory, Episodic memory, SCD, and APOE genotype, achieving the best correlation of 0.82 and the lowest RMSE of 0.52 (Fig. 1E). The Bagging model included Ventricles, Entorhinal, Parahippocampal, Global cognition, Episodic memory, SCD, and APOE genotype, with a correlation of 0.70 and an RMSE of 0.63. The Linear Regression model incorporated the Ventricles, Thalamus, Entorhinal, Processing speed, Episodic memory, SCD, and APOE genotype, showing a correlation of 0.62 and an RMSE of 0.72. The SVM model employed Ventricles, Parahippocampal, Frontal Lobe WMH, Parietal Lobe WMH, Episodic memory, SCD, and APOE genotype, with a correlation of 0.72 and an RMSE of 0.62.

**Neuropathology Prediction with clinical predictors**

To examine whether MRI features enhance predictive performance beyond that achieved with clinical features alone, we conducted a subsequent analysis focusing solely on clinical features (see Table 3). For neuritic plaques, the SVM model, with clinical predictors such as Global Cognition, Working memory, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, demonstrated the highest correlation of 0.80 and the lowest RMSE of 0.75 (Fig 1F). The Bagging model, with a similar set of predictors including Global Cognition, Processing speed, Working memory, SCD, APOE genotype, Sex, and Age at visit, also showed a high correlation of 0.79 but with a slightly higher RMSE of 0.78. The Gradient Boosting model, incorporating Processing speed, Working memory, Episodic memory, APOE genotype, Sex, and Age at visit, achieved a correlation of 0.74 and an RMSE of 0.89. The Linear Regression model, which included Global cognition among other cognitive and genetic factors, showed a lower correlation of 0.72 and a higher RMSE of 0.88.
In the prediction of diffuse plaques, the Bagging model incorporated Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE genotype, and Sex, achieving the lowest RMSE of 0.77 and a correlation of 0.62 (Fig. 1G). The SVM model included a similar set of predictors but focused more on Working memory, recording an RMSE of 0.78 and a correlation of 0.55. The Gradient Boosting model, incorporating Global cognition, Processing speed, Working memory, Episodic memory, SCD, Sex, and Age at visit, demonstrated a higher RMSE of 0.85 and a lower correlation of 0.59. The Linear Regression model, incorporating Global cognition, Processing speed, Working memory, Episodic memory, SCD, APOE genotype, and Age at visit, showed the highest RMSE of 0.89 and the lowest correlation of 0.44.

For neurofibrillary tangles, the Bagging model, including Processing speed, Working memory, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, achieved a correlation of 0.82 and an RMSE of 0.97 (Fig 1H). The SVM model, incorporating the same predictors, had the same correlation of 0.82 but a slightly better RMSE of 0.99. The Gradient Boosting model, including Global cognition, Processing speed, Working memory, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, demonstrated a lower correlation of 0.74 and the highest RMSE of 1.14. The Linear Regression model, incorporating similar predictors, recorded the lowest correlation of 0.71 and an RMSE of 1.22.

In the prediction of amyloid plaques, the Gradient Boosting model, combined Global Cognition, Processing speed, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, exhibited a correlation of 0.68 and an RMSE of 1.23. The Bagging model, including similar predictors, achieved a correlation of 0.69 and an RMSE of 1.16 (Fig 1I). The SVM model, focusing more on Working memory, showed a correlation of 0.64 and an RMSE of 1.19. The Linear Regression model, encompassed Global cognition, Processing speed, Working
memory, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, recorded the lowest correlation of 0.53 and the highest RMSE of 1.36.

In predicting amyloid angiopathy, the Gradient Boosting model, combined Global cognition, Processing speed, Working memory, Episodic memory, SCD, APOE genotype, and Sex, achieved the best correlation of 0.74 and the lowest RMSE of 0.62 (Fig. 1J). The Bagging model incorporated Global cognition, Processing speed, Working memory, SCD, APOE genotype, Sex, and Age at visit, with a correlation of 0.67 and an RMSE of 0.68. The SVM model achieved a correlation of 0.54 and an RMSE of 0.78. The Linear Regression model, including Processing speed, Working memory, Episodic memory, SCD, APOE genotype, Sex, and Age at visit, showed the lowest correlation of 0.46 and an RMSE of 0.82.
Longitudinal neuropathology prediction

To examine whether the ML models are capable of predicting future neuropathology and if MRI features enhance the prognostic predictive performance of the models beyond that achieved with clinical features alone, we conducted additional analysis focusing on longitudinal data across different timepoints prior to death (0-4 years, 4-8 years, and 8-14 years) (see Table 4). The results from the Gradient Boosting model, which was the best performing model overall at predicting neuropathology at last visit are presented for the longitudinal data. Please see the supplementary materials for the results of the other models.

For the prediction of neuritic plaques between the 0-4 years time prior to death, MRI and clinical predictors such as total WMH, hippocampus, amygdala, entorhinal region, parahippocampal, episodic memory, SCD, and APOE genotype achieved an RMSE of 0.47 and a correlation of 0.92. This model was comparable to the clinical-only model, which achieved an RMSE of 0.50 and a correlation of 0.91. In the 4-8 years interval prior to death, the model utilized total WMH, hippocampus, amygdala, ventricles, temporal lobe WMH, global cognition, and SCD as predictors. The resulting RMSE was 0.50 with a correlation of 0.92, which outperformed the clinical-only model's RMSE of 0.64 and correlation of 0.87, suggesting enhanced predictive accuracy with the inclusion of MRI data. For the longest interval, 8+ years prior to death, the model incorporated hippocampus, amygdala, ventricles, parahippocampal, episodic memory, and age at visit. The model achieved an RMSE of 0.62 and a correlation of 0.90, surpassing the clinical-only model, which achieved an RMSE of
0.73 and a correlation of 0.84. These results indicate a significant improvement in prediction with the inclusion of MRI features as the time to death increases. For MRI and clinical predictors, small but significant differences were found in the model performances between the 4-8 and 8+ years ($t = -2.44, p = 0.02$) and between 0-4 and 8+ years prior to death ($t = -2.48, p = 0.02$). For clinical-only predictors, t-tests demonstrated significant differences at all compared intervals prior to death ($t$ values ranging from -8.06 to -2.96, corrected $p$-values < 0.005).

For diffuse plaques, between the 0-4 years prior to death, MRI and clinical predictors such as amygdala, ventricles, parahippocampal, frontal lobe WMH, global cognition, processing speed, and episodic memory achieved an RMSE of 0.50 and a correlation of 0.85. This model was comparable to the clinical-only model, which achieved an RMSE of 0.53 and a correlation of 0.83. In the 4-8 years interval prior to death, the model utilized hippocampus, caudate, temporal lobe WMH, processing speed, SCD, APOE genotype, and age at visit as predictors. The resulting RMSE was 0.43 with a correlation of 0.91, which outperformed the clinical-only model's RMSE of 0.55 and correlation of 0.84, suggesting enhanced predictive accuracy with the inclusion of MRI data. For the longest interval, 8+ years, the model incorporated amygdala, putamen, thalamus, entorhinal, episodic memory, and SCD. The model achieved an RMSE of 0.45 and a correlation of 0.87, surpassing the clinical-only model, which achieved an RMSE of 0.62 and a correlation of 0.75. For MRI and clinical predictors, t-tests demonstrated significant differences at 0-4 and 4-8 years ($t = -6.23, p < 0.001$). For clinical-only predictors, $t$-tests demonstrated significant differences between 4-8 and 8+ years ($t = -7.10, p < 0.001$) and between 0-4 and 8+ years ($t = -10.07, p < 0.001$).

For neurofibrillary tangles, between the 0-4 years prior to death, MRI and clinical predictors such as the amygdala, putamen, thalamus, episodic memory, SCD, APOE genotype, and sex achieved an RMSE of 0.74 and a correlation of 0.90. This model
outperformed the clinical-only model, which achieved an RMSE of 0.82 and a correlation of 0.88. In the 4-8 years interval, the model utilized hippocampus, amygdala, caudate, frontal lobe WMH, episodic memory, SCD, and APOE genotype as predictors. The resulting RMSE was 0.59 with a correlation of 0.93, outperforming the clinical-only model's RMSE of 0.67 and correlation of 0.90, suggesting enhanced predictive accuracy with the inclusion of MRI data. For the longest interval, 8+ years, the model incorporated amygdala, parahippocampal, processing speed, episodic memory, SCD, APOE genotype, and age at visit. The model achieved an RMSE of 0.92 and a correlation of 0.82, which was identical to the clinical-only model's RMSE and correlation. This indicates a consistent predictive performance over time.

For MRI and clinical predictors and clinical only predictors, t-tests demonstrated significant differences at all compared intervals (t values ranging from 7.09 to -13.60, corrected p-values < 0.001).

For the prediction of amyloid plaques, at the 0-4 years prior to death, MRI and clinical predictors such as total WMH, parahippocampal region, frontal lobe WMH, episodic memory, SCD, APOE genotype, and sex achieved an RMSE of 0.68 and a correlation of 0.90. This model was more accurate than the clinical-only model, which achieved an RMSE of 0.82 and a correlation of 0.85. In the 4-8 years interval, the model utilized total WMH, ventricles, parahippocampal region, working memory, episodic memory, SCD, and APOE genotype as predictors. The resulting RMSE was 0.62 with a correlation of 0.93, outperforming the clinical-only model's RMSE of 0.69 and correlation of 0.91, indicating improved predictive accuracy with the inclusion of MRI data. For the longest interval, 8+ years, the model incorporated ventricles, parahippocampal region, frontal lobe WMH, occipital lobe WMH, episodic memory, SCD, and sex. The model achieved an RMSE of 0.68 and a correlation of 0.90, which was superior to the clinical-only model's RMSE of 0.80 and correlation of 0.86. This indicates the consistent predictive advantage of including MRI features. For MRI and
clinical predictors, t-tests demonstrated significant differences between 0-4 and 4-8 years ($t = 10.36, p < 0.001$) and between 4-8 and 8+ years ($t = -5.84, p < 0.001$). For clinical-only predictors, t-tests showed significant differences between 0-4 and 4-8 years ($t = 4.03, p < 0.001$) and between 4-8 and 8+ years ($t = 2.05, p = 0.05$).

For amyloid angiopathy, between 0-4 years prior to death, MRI and clinical predictors such as the ventricles, entorhinal cortex, parahippocampal region, temporal lobe WMH, occipital lobe WMH, processing speed, and SCD achieved an RMSE of 0.44 and a correlation of 0.86. This model was better than the clinical-only model, which achieved an RMSE of 0.54 and a correlation of 0.78. In the 4-8 years interval prior to death, the model utilized parahippocampal region, frontal lobe WMH, occipital lobe WMH, processing speed, episodic memory, SCD, and APOE genotype as predictors. The resulting RMSE was 0.49 with a correlation of 0.88, outperforming the clinical-only model's RMSE of 0.54 and correlation of 0.85, indicating enhanced predictive accuracy with the inclusion of MRI data. For the longest interval, 8+ years, the model incorporated amygdala, ventricles, thalamus, episodic memory, SCD, APOE genotype, and age at visit. The model achieved an RMSE of 0.62 and a correlation of 0.83, surpassing the clinical-only model's RMSE of 0.81 and correlation of 0.69, demonstrating significant improvement in prediction over time. For MRI and clinical predictors, t-tests demonstrated significant differences at all compared intervals ($t$ values ranging from 5.72 to -6.85, $p$-values < 0.001). For clinical-only predictors, t-tests showed significant differences between 0-4 and 8+ years ($t=-5.73, p<0.001$), and between 4-8 and 8+ years ($t=-6.94, p<0.001$). This finding affirms the model's enhanced predictive capability with the inclusion of MRI features, particularly as the time between MRI follow-up and measurement of pathology (i.e., time to death) increases.

**Discussion**
Previous literature has shown the importance of neurofibrillary tangles, amyloid-beta deposits, and alterations in brain morphology, such as GM atrophy and WM abnormalities, in the development of AD [1,3]. Although these studies have recognized the importance of brain changes, the integration of such features into ML models for neuropathology prediction analysis has been less explored. In this study, we investigated the predictive power of ML models in identifying neuropathology outcomes in aging individuals on the spectrum of AD, with a focus on the importance of MRI measures in addition to clinical measures. Our findings reveal that models incorporating MRI predictors, alongside clinical measures demonstrate significant predictive ability for key neuropathological measures such as neurofibrillary tangles, neuritic plaques, diffuse plaques, amyloid plaques, and amyloid angiopathy scores. The longitudinal results indicate a clear trend of increasing number of selected MRI predictors across neuropathology outcomes as the time interval from data collection to death increased, specifically for Neuritic Plaques, Diffuse Plaques, and Amyloid Angiopathy. This is in line with findings from previous studies that investigate the utility of clinical and MRI features in predicting conversion from MCI to dementia in future timepoints [39]. Notably, the models that include MRI features achieved higher performance than models using clinical measures alone at different time intervals prior to death. Gradient boosting, SVM, and bagging models achieved consistently high correlations, gradient boosting however, had superior predictive capabilities, achieving high correlations and low RMSE across most outcomes. When analyzing only clinical predictors, the performance metrics generally indicated lower predictive accuracy for AD neuropathology, reinforcing the added value of incorporating MRI features alongside clinical scores to enhance the predictive capability of ML models. These results are consistent with previous literature highlighting the role of structural and vascular brain changes in cognitive decline and AD progression [2–5].
Results from the present study demonstrate the predictive utility of integrating vascular markers, specifically WMHs, with structural brain changes through ML models.

Our findings also revealed that nonlinear ML models outperformed linear models in predicting AD neuropathology. MRI features do not necessarily influence outcomes in a linear manner; therefore, the nonlinear methods are able to better capture the variability observed in neuropathology. This finding underscores the significance for clinical settings and early prediction research, as it reveals that traditional methods, which assume linear patterns of change, may overlook crucial nuances in neuropathological progression due to the nonlinear influence of predictor variables like GM, WM, WMHs. Nonlinear models, capable of navigating these complexities, can combine information from a set of features to offer more accurate estimation of AD neuropathology and progression than linear models.

The ML models employed in our study exhibited notably high predictive accuracy in estimating neuropathological outcomes as defined by different criteria for neurofibrillary tangles (Braak stage), neuritic and diffuse plaques (CERAD), and amyloid plaques (Thal amyloid phase). In the prediction of neuritic plaques, the SVM model, which used MRI and clinical data, achieved an $R^2$ of 0.76, reflecting a strong predictive performance. Comparatively, the SVM model using clinical data only produced an $R^2$ of 0.64, highlighting the importance of including MRI data for more accurate predictions even at last visit.

The longitudinal results which incorporate MRI and clinical data from varying time intervals before death (0-4 years, 4-8 years, and 8-14 years), revealed that specific MRI predictors become more useful over time. For instance, regions such as the Amygdala, Frontal Lobe WMH, Hippocampus, Parahippocampal, and Occipital Lobe WMH played increasingly significant roles in long-term predictions. These findings are consistent with the progressive staging of AD. For example, the hippocampus is one of the first regions affected by
neurodegeneration[8], and occipital WMHs are the main WMHs associated with AD [4,40]. These findings underscore the importance of utilizing longitudinal MRI data to enhance predictive accuracy.

The comparison of prediction accuracy across time-to-death groups revealed the importance of MRI predictors for future neuropathological outcomes. MRI predictors such as the Amygdala, Hippocampus, and Ventricles remained significant across both last visit and longitudinal analyses, underscoring their critical roles in predicting neuropathological outcomes. For Neurofibrillary Tangles, both datasets identified the Amygdala as an important predictor, although the longitudinal data showed superior predictive accuracy. Additionally, the Frontal and Occipital Lobe WMHs, Ventricles, and Parahippocampal were consistently highlighted as significant predictors in both analyses for outcomes like Amyloid Plaques and Amyloid Angiopathy, respectively.

This high level of predictive performance, achieved using MRI features such as GM, and WMHs, is remarkable, particularly when integrating longitudinal data. The strong performance of these models demonstrates that ML can effectively leverage standard MRI features to provide insights into the underlying neuropathological state, complementing clinical assessments. However, the few studies that have examined neuropathology prediction have been in relatively small samples and report conflicting findings. For instance, a study by Toledo et al [41] utilized cerebrospinal markers, MRI features, and clinical data from 22 ADNI participants and found that non-MRI features were more effective than MRI features in predicting dementia with Lewy bodies in addition to Alzheimer's pathology. They further highlighted that hippocampal volume was not a reliable feature for predicting pathologies. However, Kautzky et al. [20] applied ML with MRI data in 49 participants to predict neuropathological changes in AD, achieving an accuracy of 77%. Important MRI features highlighted in their study included the entorhinal cortex and anterior cingulate cortex.
Building on previous studies, our results underline the potential of integrating advanced machine learning techniques with clinical and MRI data to enhance the predictive accuracy and offer a more comprehensive approach to understanding Alzheimer’s disease progression.

Despite the promising findings of our study, there are limitations that warrant attention. Our study relied on data from the closest MRI timepoint available (under four years from time of death) to maximize sample size. The time difference between the MRI and neuropathology assessments might have impacted the performance of the models. Future studies with closely matched MRI and neuropathology information and longitudinal settings could improve the accuracy of the prediction and assess the change in the prediction accuracy as a function of time between MRI and neuropathology information. The clinical measures and biomarkers used in this study, though carefully selected, do not encompass all possible factors influencing AD neuropathology. Emerging biomarkers and novel imaging techniques continue to be identified, suggesting that integrating these new data sources could potentially enhance the predictive power of future models. Future studies should aim to address these challenges by incorporating longitudinal data and including a broader range of biomarkers to improve prediction accuracy and clinical utility. Lastly, future studies should validate the results in other datasets to ensure generalizability of the selected features. The ability to maintain high performance levels, across multiple dataset’s demographic and clinical diversity, would further emphasize the model's robust generalizability. This consistency in feature importance across datasets reinforces the reliability and relevance of these features in predicting AD neuropathology, aligning with, and expanding on previous ML research [17,18].

This study combined vascular, structural, and clinical features with machine learning techniques to predict post-mortem neuropathology. We successfully predicted post-mortem neuropathology (~90% accuracy) in these individuals at up to 14 years before death. Models
that included MRI features consistently had the highest accuracy for all AD neuropathology outcomes compared to models with clinical features only. That is, clinical features are successful at predicting AD, but the highest accuracy when predicting AD-related pathology includes MRI features. Not only will these findings improve AD research and patient care in clinical settings, but they could be implemented in clinical trials to select older adults who are at increased risk for AD and who would most benefit from clinical intervention. Our study demonstrates the feasibility and efficacy of using ML models, equipped with a carefully selected set of MRI features, to predict AD neuropathology across diverse cognitive status.
Abbreviations

AD: Alzheimer’s disease

GM: gray matter

WM: white matter

WMH: white matter hyperintensity

MRI: Magnetic resonance imaging

NFT: neurofibrillary tangle burden

ML: machine learning

SVM: support vector machine

ADNI: Alzheimer’s Disease Neuroimaging Initiative

CDR-SB: Clinical Dementia Rating- Sum of Boxes

SCD: subjective cognitive decline

RMSE: root mean square error

FDR: false discovery rate
References


https://doi.org/10.1016/j.neuroimage.2018.03.025.


Table 1: Demographic Information in the ADNI dataset across time of death intervals prior to post-mortem

<table>
<thead>
<tr>
<th></th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-4 Years</td>
</tr>
<tr>
<td>Sample Size</td>
<td>59</td>
</tr>
<tr>
<td>Age</td>
<td>81.2 ± 6.7</td>
</tr>
<tr>
<td>Sex (Female)</td>
<td>15 (25%)</td>
</tr>
<tr>
<td>Education</td>
<td>16.3 ± 2.8</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>7 (11%)</td>
</tr>
<tr>
<td>MCI</td>
<td>8 (14%)</td>
</tr>
<tr>
<td>AD</td>
<td>44 (75%)</td>
</tr>
</tbody>
</table>

Numbers are presented as total (percentage of group) or mean ± standard deviation. NC = cognitively healthy older adults. MCI = mild cognitive impairment. AD = Alzheimer’s Disease.
Table 2: MRI and clinical predictors of neuropathology outcomes across each model at last visit (within 4 years prior to death)

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>RMSE</th>
<th>CORR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuritic Plaques</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Amygdala, Temporal Lobe WMH, Working memory, Episodic memory, SCD, APOE Genotype, Age at visit</td>
<td>0.84</td>
<td>0.75</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Amygdala, Ventricles, Global Cognition, Working memory, Episodic memory, SCD, APOE Genotype</td>
<td>0.85</td>
<td>0.76</td>
</tr>
<tr>
<td>Linear Regression Svm</td>
<td>Thalamus, Frontal Lobe WMH, Parietal Lobe WMH, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>0.86</td>
<td>0.73</td>
</tr>
<tr>
<td>SVM</td>
<td>Total WMH, Thalamus, Temporal Lobe WMH, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>0.62</td>
<td>0.87</td>
</tr>
<tr>
<td>Diffuse Plaques</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Total WMH, Amygdala, Ventricles, Processing speed, Working memory, Episodic memory, SCD</td>
<td>0.66</td>
<td>0.73</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Hippocampus, Amygdala, Ventricles, Working memory, Episodic memory, SCD, Sex</td>
<td>0.63</td>
<td>0.78</td>
</tr>
<tr>
<td>Linear Regression Svm</td>
<td>Hippocampus, Putamen, Processing speed, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>0.84</td>
<td>0.49</td>
</tr>
<tr>
<td>SVM</td>
<td>Hippocampus, Amygdala, Ventricles, Temporal Lobe WMH, Working memory, SCD, APOE Genotype</td>
<td>0.66</td>
<td>0.74</td>
</tr>
<tr>
<td>Neurofibrillary Tangles</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Amygdala, Para-hippocampal, Temporal Lobe WMH, Processing speed, Episodic memory, SCD, APOE Genotype</td>
<td>1.23</td>
<td>0.69</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Amygdala, Temporal Lobe WMH, Processing speed, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>1.23</td>
<td>0.70</td>
</tr>
<tr>
<td>Linear Regression Svm</td>
<td>Thalamus, Temporal Lobe WMH, Parietal Lobe WMH, Episodic memory, SCD, APOE Genotype</td>
<td>1.23</td>
<td>0.70</td>
</tr>
<tr>
<td>SVM</td>
<td>Hippocampus, Thalamus, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>1.02</td>
<td>0.81</td>
</tr>
<tr>
<td>Amyloid Plaques</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Ventricles, Caudate, Frontal Lobe WMH, Global Cognition, Working memory, Episodic memory, SCD</td>
<td>1.09</td>
<td>0.72</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Ventricles, Thalamus, Global Cognition, Processing speed, Episodic memory, SCD, APOE Genotype</td>
<td>0.98</td>
<td>0.78</td>
</tr>
<tr>
<td>Linear Regression Svm</td>
<td>Amygdala, Caudate, Thalamus, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>1.16</td>
<td>0.65</td>
</tr>
<tr>
<td>SVM</td>
<td>Caudate, Thalamus, Frontal Lobe WMH, Occipital Lobe WMH, Episodic memory, SCD, APOE Genotype</td>
<td>1.03</td>
<td>0.76</td>
</tr>
<tr>
<td>Amyloid Angiopathy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Ventricles, Entorhinal, Para-hippocampal, Global Cognition, Episodic memory, SCD, APOE Genotype</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>Gradient</td>
<td>Ventricles, Caudate, Para-hippocampal, Working memory, Episodic memory, SCD, APOE Genotype</td>
<td>0.52</td>
<td>0.82</td>
</tr>
<tr>
<td>Model</td>
<td>Predictors</td>
<td>RMSE</td>
<td>CORR</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td><strong>Neuritic Plaques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Global Cognition, Processing speed, Working memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, APOE Genotype, Sex, Age at visit</td>
<td>0.89</td>
<td>0.74</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Global Cognition, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.88</td>
<td>0.72</td>
</tr>
<tr>
<td>SVM</td>
<td>Global Cognition, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.75</td>
<td>0.80</td>
</tr>
<tr>
<td><strong>Diffuse Plaques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>0.77</td>
<td>0.62</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, Sex, Age at visit</td>
<td>0.85</td>
<td>0.59</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.89</td>
<td>0.44</td>
</tr>
<tr>
<td>SVM</td>
<td>Global Cognition, Processing speed, Working memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.78</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>Neurofibrillary Tangles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagging</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.97</td>
<td>0.82</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>1.14</td>
<td>0.74</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>1.22</td>
<td>0.71</td>
</tr>
<tr>
<td>SVM</td>
<td>Global Cognition, Processing speed, Working memory, SCD, APOE Genotype, Sex, Age at visit</td>
<td>0.99</td>
<td>0.82</td>
</tr>
<tr>
<td><strong>Amyloid Plaques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: WMH = White matter hyperintensity, SCD = Subjective cognitive decline
<table>
<thead>
<tr>
<th>Method</th>
<th>Features</th>
<th>Visit 1</th>
<th>Visit 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagging</td>
<td>Global Cognition, Processing speed, Episodic memory, SCD, APOE Genotype, Sex, Age</td>
<td>1.16</td>
<td>0.69</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Global Cognition, Processing speed, Episodic memory, SCD, APOE Genotype, Sex, Age</td>
<td>1.23</td>
<td>0.68</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>1.36</td>
<td>0.53</td>
</tr>
<tr>
<td>SVM</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age</td>
<td>1.19</td>
<td>0.64</td>
</tr>
</tbody>
</table>

**Amyloid Angiopathy**

<table>
<thead>
<tr>
<th>Method</th>
<th>Features</th>
<th>Visit 1</th>
<th>Visit 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagging</td>
<td>Global Cognition, Processing speed, Working memory, SCD, APOE Genotype, Sex, Age</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>Gradient Boosting</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex</td>
<td>0.62</td>
<td>0.74</td>
</tr>
<tr>
<td>Linear Regression</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age</td>
<td>0.82</td>
<td>0.46</td>
</tr>
<tr>
<td>SVM</td>
<td>Global Cognition, Working memory, Episodic memory, SCD, APOE Genotype, Sex, Age</td>
<td>0.78</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Abbreviations: WMH = White matter hyperintensity, SCD = Subjective cognitive decline.
<table>
<thead>
<tr>
<th>Neurpathology</th>
<th>Time</th>
<th>MRI and Clinical Predictors</th>
<th>RMSE</th>
<th>CORR</th>
<th>Clinical Predictors</th>
<th>RMSE</th>
<th>CORR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffuse Plaques</td>
<td>0-4 Years</td>
<td>Amygdala, Ventricles, Parahippocampal, Frontal Lobe WMH, Global Cognition, Processing speed, Episodic memory¹</td>
<td>0.50</td>
<td>0.85</td>
<td>Global Cognition, Processing speed, Episodic memory, SCD, APOE Genotype, Sex¹²</td>
<td>0.53</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>4-8 Years</td>
<td>Hippocampus, Caudate, Temporal Lobe WMH, Processing speed, SCD, APOE Genotype, Age at Visit</td>
<td>0.43</td>
<td>0.91</td>
<td>Global Cognition, Processing speed, Episodic memory, SCD, APOE Genotype, Sex¹²</td>
<td>0.55</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>8+ Years</td>
<td>Amygdala, Putamen, Thalamus, Entorhinal, Episodic memory, SCD, APOE Genotype</td>
<td>0.45</td>
<td>0.87</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex¹²</td>
<td>0.62</td>
<td>0.75</td>
</tr>
<tr>
<td>Neurofibrillary Tangles</td>
<td>0-4 Years</td>
<td>Amygdala, Putamen, Thalamus, Episodic memory, SCD, APOE Genotype, Sex¹³</td>
<td>0.74</td>
<td>0.90</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex¹³</td>
<td>0.82</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>4-8 Years</td>
<td>Hippocampus, Amygdala, Caudate, Frontal Lobe WMH, Episodic memory, SCD, APOE Genotype²</td>
<td>0.59</td>
<td>0.93</td>
<td>Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex¹³</td>
<td>0.67</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>8+ Years</td>
<td>Amygdala, Parahippocampal, Processing speed, Episodic memory, SCD, APOE Genotype, Age at Visit</td>
<td>0.92</td>
<td>0.82</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Sex¹³</td>
<td>0.92</td>
<td>0.82</td>
</tr>
<tr>
<td>Amyloid Plaques</td>
<td>0-4 Years</td>
<td>Total WMH, Parahippocampal, Frontal Lobe WMH, Episodic memory, SCD, APOE Genotype, Sex¹</td>
<td>0.68</td>
<td>0.90</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, APOE Genotype, Sex, Age at Visit¹²</td>
<td>0.82</td>
<td>0.85</td>
</tr>
<tr>
<td>Age Group</td>
<td>Regions</td>
<td>Global Cognition, Processing speed, Working memory, Episodic memory, SCD, APOE Genotype, Age at Visit</td>
<td>4-8 Years Total WMH, Ventricles, Parahippocampal, Working memory, Episodic memory, SCD, APOE Genotype²</td>
<td>8+ Years Ventricles, Parahippocampal, Frontal Lobe WMH, Occipital Lobe WMH, Episodic memory, SCD, Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-4 Years</td>
<td>Ventricle, Entorhinal, Parahippocampal, Temporal Lobe WMH, Occipital Lobe WMH, Processing speed, SCD¹³</td>
<td>0.44 0.86</td>
<td>0.62 0.93</td>
<td>0.68 0.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-8 Years</td>
<td>Parahippocampal, Frontal Lobe WMH, Occipital Lobe WMH, Processing speed, Episodic memory, SCD, APOE Genotype²</td>
<td>0.49 0.88</td>
<td>0.62 0.93</td>
<td>0.68 0.90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8+ Years</td>
<td>Amygdala, Ventricle, Thalamus, Episodic memory, SCD, APOE Genotype, Age at Visit</td>
<td>0.62 0.83</td>
<td>0.69 0.91</td>
<td>0.80 0.86</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: WMH = White matter hyperintensity, SCD = Subjective cognitive decline

¹ indicates significant differences between the 0-4 and 4-8 years comparison.
² indicates significant differences between the 4-8 and 8+ years comparison.
³ indicates significant differences between the 0-4 and 8+ years comparison.
Figure 1: Scatter plots showing the performance of the best models across post-mortem neuropathologies. The models compare actual versus predicted values for neuropathological outcomes across MRI and Clinical (A-E), and Clinical only (F-J). The regression line represents the best linear fit through the data points. The shaded area around the regression line represents the 95% confidence interval.
Declarations

Ethics approval and consent to participate

Not applicable.

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.

Competing interests

The authors declare no competing interests.

Funding

This research was supported by a grant from the Canadian Institutes of Health Research (CIHR). Dr. Kamal is supported by a scholarship from Fonds de Recherche du Québec - Santé (FRQS). Dr. Dadar reports receiving research funding from the Quebec Bio-Imaging Network and Fonds de Recherche du Québec - Santé (FRQS), Natural Sciences and Engineering Research Council of Canada (NSERC), Healthy Brains for Healthy Lives (HBHL), Alzheimer Society Research Program (ASRP), CIHR, and Douglas Research Centre (DRC). Dr. Morrison is supported by CIHR.

Consent Statement

Written informed consent was obtained from participants or their study partner

Availability of data and materials

Data used in preparation of this article were also obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu).
Disclosures

The authors report no disclosures relevant to the manuscript.

Contributions

F.K, M.O, M.D, and C.M, were involved with the conceptualization and design of the work. F.K. and M.D. completed analysis and C.M, M.D, M.O, and F.K were involved with data interpretation. F.K. organized figures. F.K. wrote the manuscript. and C.M, M.O, M.D, and F.K revised and approved the submitted version.

Consent for publication

Not applicable.