# Predicting diagnosis 4 years prior to Alzheimer's disease incident

Anqi Qiu [a,b,c,d,e,f,*], Liyuan Xu [e], Chaoqiang Liu [a], for the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Department of Biomedical Engineering, National University of Singapore, Singapore
[b] The N.1 Institute for Health, National University of Singapore, Singapore
[c] Institute of Data Science, National University of Singapore, Singapore
[d] NUS (Suzhou) Research Institute, Suzhou, China
[e] School of Computer Engineering and Science, Shanghai University, China
[f] Department of Biomedical Engineering, the Johns Hopkins University, USA

## ARTICLE INFO

## ABSTRACT

This study employed a deep learning longitudinal model, graph convolutional and recurrent neural network (graph-CNN-RNN), on a series of brain structural MRI scans for AD prognosis. It characterized whole-brain morphology via incorporating longitudinal cortical and subcortical morphology and defined a probabilistic risk for the prediction of AD as a function of age prior to clinical diagnosis. The graph-CNN-RNN model was trained on half of the Alzheimer's Disease Neuroimaging Initiative dataset (ADNI, n = 1559) and validated on the other half of the ADNI dataset and the Open Access Series of Imaging Studies-3 (OASIS-3, n = 930). Our findings demonstrated that the graph-CNN-RNN can reliably and robustly diagnose AD at the accuracy rate of 85% and above across all the time points for both datasets. The graph-CNN-RNN predicted the AD conversion from 0 to 4 years before the AD onset at ~80% of accuracy. The AD probabilistic risk was associated with clinical traits, cognition, and amyloid burden assessed using [18F]-Florbetapir (AV45) positron emission tomography (PET) across all the time points. The graph-CNN-RNN provided the quantitative trajectory of brain morphology from prognosis to overt stages of AD. Such a deep learning tool and the AD probabilistic risk have great potential in clinical applications for AD prognosis.

## 1. Introduction

Alzheimer's Disease (AD) is a type of brain neurodegenerative diseases that become worse with time. It has an astounding impact at individual and societal levels. AD may start to develop 20 years ago before clinical symptoms arise. Changes in the brain, such as amyloid plaques, neurofibrillary tangles, and tissue loss, are not noticeable to the person who is affected until clinical symptoms appear. Prodromal stages of AD are windows of opportunity in reducing the incidence and symptoms of AD.

Structural brain imaging changes are suggested to lie between AD neuropathology and clinical and cognitive decline (Frisoni et al., 2010). Structural brain MRI characterizes tissue damage or loss in the medial temporal lobe (i.e., entorhinal cortex and hippocampus), followed by progressive cortical damage that occurs years before clinical symptoms appear (Chan et al., 2001; Thompson et al., 2003). Hippocampal atrophy is the best established and validated among all structural MRI markers of AD, such as voxel-by-voxel morphology and structural volumes (Modrego, 2006). Meta-analysis also shows that medial temporal atrophy has a sensitivity of ~ 70% and a specificity of ~ 80% for the AD prediction (Yuan et al., 2009). Moreover, structural imaging markers are more sensitive to change from mild cognitive impairment (MCI) to AD than markers of amyloid deposition assessed using cerebrospinal fluid (CSF) (Jack et al., 2009; Sluimer et al., 2010). Hence, structural brain imaging has been recommended to be included as a part of clinical assessment for earlier diagnosis of AD ((Dubois et al., 2007). Frisoni et al. (2010)

([Frisoni et al., 2010](#)) proposed a theoretical trajectory of whole-brain and hippocampal atrophy over the course of AD but did not give detailed quantification. *Also, it is unclear whether a single scalar derived from structural brain imaging can be a precise indicator of AD progression. Moreover, it remains unclear whether such a neuroanatomical scalar can be clinically adopted and clinicians can use it in the same way as cognitive scales.*

Deep learning has received great attention in the field of medical image analysis. In particular, convolutional neural networks (CNNs) have extensively been explored as computer-aided tools for the diagnosis of AD ([Basheera and Sai Ram, 2019; Bashyam et al., 2020; Liu et al., 2020; Liu et al., 2019; Qiu et al., 2020; Wee et al., 2019](#)). Nevertheless, classification performance varies significantly from one study to another due to subject selection criteria, sample size, and image processing, which may raise a question on the generalization of deep learning approaches in clinical practice. Several recent studies performed extensive experiments using structural imaging data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and demonstrated that CNN approaches on cross-sectional 3D brain structural images (pre-processed or minimally processed) achieved 75% to 90% of the AD classification accuracy ([Ansart et al., 2021; Jin et al., 2020; Wen et al., 2020](#)). [Jin et al. (2020)](#) showed the generalizability and reliability of deep learning approaches in the application of early diagnosis of AD based on multiple independent datasets. Moreover, a number of existing studies employed deep learning networks and focused on diagnostic conversion within a fixed time-window to identify subjects converted to AD in near future (e.g.,[Basaia et al., 2019; Lian et al., 2020; Liu et al., 2018](#)). Traditional machine learning models employed longitudinal data and has shown the improvement of diagnostic conversion ([Ye et al., 2012; Zhang et al., 2012; Davatzikos et al., 2011; Risacher et al., 2009; Eskildsen et al., 2015; Moradi et al., 2015](#)). In aid of a series of brain MRI scans, deep learning has potential to further improve the accuracy of AD diagnosis and prognosis. Nevertheless, it is challenging to handle missing brain scans at random time points.

This study aimed to develop a deep learning approach, graph convolutional and recurrent neural network (graph-CNN-RNN) on a series of brain structural MRI data, that can be used to 1) define a diagnosis-guided probabilistic risk for characterizing whole-brain morphology at each time point; 2) quantify the longitudinal trajectory of whole-brain morphology over the course of AD; 3) predict the AD conversion as a function of age prior to clinical diagnosis. This graph-CNN-RNN approach was designed as a longitudinal model that incorporated the information of subcortical volumes and cortical geometry at previous and current time points and computed the AD probabilistic risk at each time point. If there was a missing brain scan at a particular time point, the graph-CNN-RNN approach estimated it based on its previous time points. We trained this deep learning approach using a series of structural brain MRI scans from half of the Alzheimer's Disease Neuroimaging Initiative (ADNI) samples and evaluated it on structural brain MRI data of the other half of the ADNI sample. Moreover, we also employed an independent sample, the Open Access Series of Imaging Studies-3 sample (OASIS-3, n = 930), to evaluate the generalizability and robustness of the graph-CNN-RNN trained by the ADNI dataset. We further demonstrated the usefulness of the AD probabilistic risk derived from our graph-CNN-RNN model via its associations with clinical traits, cognition, amyloid markers obtained from [18F]-Florbetapir (AV45) positron emission tomography (PET). Our results suggested its potential use in clinic and shed light on the quantification of the longitudinal trajectory of brain morphology over the course of AD and the AD prediction 4 years prior to clinical diagnosis.

## 2. Methods

### 2.1. Participants

Data used in this study were obtained from the Alzheimer's Disease

Neuroimaging Initiative (ADNI) database ([http://adni.loni.usc.edu](http://adni.loni.usc.edu)) and the Open Access Series of Imaging Studies-3 (OASIS-3; [http://oasis-brains.org](http://oasis-brains.org)). Institutional review boards approved study procedures across participating institutions.

The ADNI dataset included ADNI-1 (n = 811), ADNI-GO (n = 188) and ADNI-2 (n = 1019). At each visit, subjects were diagnosed as one of three clinical statuses (cognitive normal (CN), mild cognitive impairment (MCI), Alzheimer's disease (AD)) based on the criteria described in the ADNI protocol ([http://adni.loni.usc.edu](http://adni.loni.usc.edu)). The number of visits per subject varied from 1 to 11. Table 1 lists the number of subjects based on the number of visits in each diagnostic group. There were 434 CN subjects, 548 stable MCI (s-MCI) subjects who did not convert to AD across all existing visits, 326 CE subjects, and 251 subjects converted from CN or MCI to AD. In addition, there were 115 cases, including 51 subjects converted from CN to MCI, 49 subjects converted to CN from MCI, and 15 subjects converted back to CN or MCI from AD across the time, which were not used in this study.

This study also included the OASIS-3 dataset to evaluate the generalization and robustness of our graph-CNN-RNN model. This study included 650 CN subjects, 47 s-MCI subjects, 233 CE subjects, and 17 subjects converted from CN to AD from the OASIS dataset ([Marcus et al., 2010](#)). The number of visits per subject varied from 1 to 7. Table 2 lists the number of subjects based on the number of visits in each diagnostic group of the OASIS-3 dataset. The diagnosis information of 58 subjects was missing or converted back to CN from MCI and AD, which were not used in this study. Due to the small sample size at each time point, the converted subjects were also excluded in this study (Table 2).

### 2.2. MRI processing

Both the ADNI and OASIS-3 studies acquired structural T1-weighted MRI scans using either 1.5 T or 3 T scanners. FreeSurfer (version 5.3.0) longitudinal analysis pipeline was then used to label each voxel in the usable $T_1$-weighted image as gray matter, white matter, cerebrospinal fluid (CSF), or subcortical structures ([Fischl et al., 2002](#)). Post-processing quality check was conducted following the instruction on [https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData](https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/TroubleshootingData). We employed large deformation diffeomorphic metric mapping (LDDMM) ([Du et al., 2011; Tan and Qiu, 2016; Zhong and Qiu, 2010](#)) to align individual cortical surfaces to the atlas and transferred cortical thickness of each subject to a common space. This study also extracted bilateral subcortical and ventricular volumes from FreeSurfer analysis and included the volumes of the hippocampus, amygdala, thalamus, caudate, putamen, and globus pallidus in the following deep learning.

**Table 1**
The number of subjects who have n number of visits in each diagnostic group of **the ADNI dataset**.

| the number of visits | CN | s-MCI | AD | Conversion |
|---|---|---|---|---|
| 1 | 54 | 45 | 49 | 0 |
| 2 | 62 | 57 | 47 | 7 |
| 3 | 52 | 87 | 75 | 13 |
| 4 | 79 | 127 | 146 | 37 |
| 5 | 94 | 114 | 7 | 68 |
| 6 | 31 | 64 | 2 | 59 |
| 7 | 21 | 28 | 0 | 30 |
| 8 | 17 | 8 | 0 | 18 |
| 9 | 18 | 12 | 0 | 12 |
| 10 | 6 | 6 | 0 | 3 |
| 11 | 0 | 0 | 0 | 4 |
| total | 434 | 548 | 326 | 251 |

Abbreviations: CN, controls; s-MCI, stable mild cognitive impairment; AD, Alzheimer's disease.

**Table 2**
The number of subjects who have n number of visits in each diagnostic group of **the OASIS-3 dataset.**

| the number of visits | CN | s-MCI | AD | Conversion |
|---|---|---|---|---|
| 1 | 242 | 42 | 187 | 0 |
| 2 | 191 | 3 | 44 | 8 |
| 3 | 114 | 2 | 1 | 8 |
| 4 | 64 | 0 | 1 | 1 |
| 5 | 26 | 0 | 0 | 0 |
| 6 | 13 | 0 | 0 | 0 |
| total | 650 | 47 | 233 | 17 |

Abbreviations: CN, controls; s-MCI, stable mild cognitive impairment; AD, Alzheimer's disease.

### 2.3. Diagnostic-Guided probabilistic risk via graph convolution and recurrent neural networks

#### 2.3.1. Graph-CNN-RNN model

This study aimed to derive a probabilistic risk for AD diagnosis from structural MRI that is most relevant to neurodegeneration. We developed a deep learning approach (graph-CNN-RNN) that incorporated vertex-based graph convolutional neural network (vertex-based graph CNN, code available at https://bieqa.github.io/deeplearning.html), (Liu et al., 2021) and minimal recurrent neural network (RNN) (Chen, 2017). This graph-CNN-RNN aimed to generate the probabilistic risk and make the diagnosis and prediction of AD based on longitudinal structural MRI data. We chose vertex-based graph-CNN because of its fast computation and better performance than spectral graph-CNN as shown in (Liu et al., 2021). We employed the minimal RNN instead of long-short term memory (LSTM) because of its efficient model parameterization (fewer parameters) to be estimated (Chen, 2017).

Fig. 1 illustrates the architecture of the vertex-based graph-CNN-RNN that incorporated a vertex-based graph CNN with 4 graph convolutional layers and 1 fully connected layer as well as a minimal RNN. The vertex-based graph-CNN took cortical thickness as input data and the cortical surface represented by a triangulated mesh as a graph. It

conducted filtering of cortical thickness on the cortical surface. Each convolutional layer had 1-ring neighbour filters, rectified linear unit (ReLU), and pooling with a down-sample rate of 4. The four convolutional layers respectively had 8, 16, 32, and 64 filters as shown in Fig. 1. The fully connected layer concatenated all filtered data and reduced the data dimensionality to 128. The mathematical foundation of the vertex-based graph-CNN was described in (Liu et al., 2021). We optimized this network by minimizing a three-class classifier.

The minimal RNN then incorporated the 128 cortical features obtained from the vertex-based graph-CNN with the volumetric data, including the bilateral hippocampus, amygdala, thalamus, basal ganglia, and ventricle volumes. We denoted these morphological measures as $x_t$ at the $t^{th}$ visit. $\tilde{x}_{t+1}$ represents predicted morphological measures (thickness features and volumes) and the diagnostic probability of CN, MCI, and AD at $t + 1$. If actual image data, $x_{t+1}$, were missing, $\tilde{x}_{t+1}$ was used in the model for the estimation of the next time point. Otherwise, the real image measures were used. We minimized the square errors of the image features and cross entropy of diagnostic labels.

#### 2.3.2. Graph-CNN-RNN training and evaluation

We trained our network model using half of the CN, s-MCI, and AD samples at each time point in the ADNI dataset (see Table 1) with a learning rate of 1e-3. The rest of the CN, s-MCI, AD samples in the ADNI dataset were used as a testing dataset to evaluate the entire network (see eTable 1, Supplementary Material). The diagnostic accuracy of each class at each time point was computed as the percentage of correctly classified subjects over the overall subjects of the same class at each time point in the ADNI testing dataset.

We also applied the graph-CNN-RNN model trained using the ADNI dataset to the OASIS-3 dataset to examine its robustness and reliability on the AD classification. We did not use the OASIS-3 as a training dataset due to its limited time points and relatively small sample sizes of s-MCI and AD (see Table 2). The diagnostic accuracy of each class at each time point was computed as the percentage of correctly classified subjects over the overall subjects of the same class at each time point in the
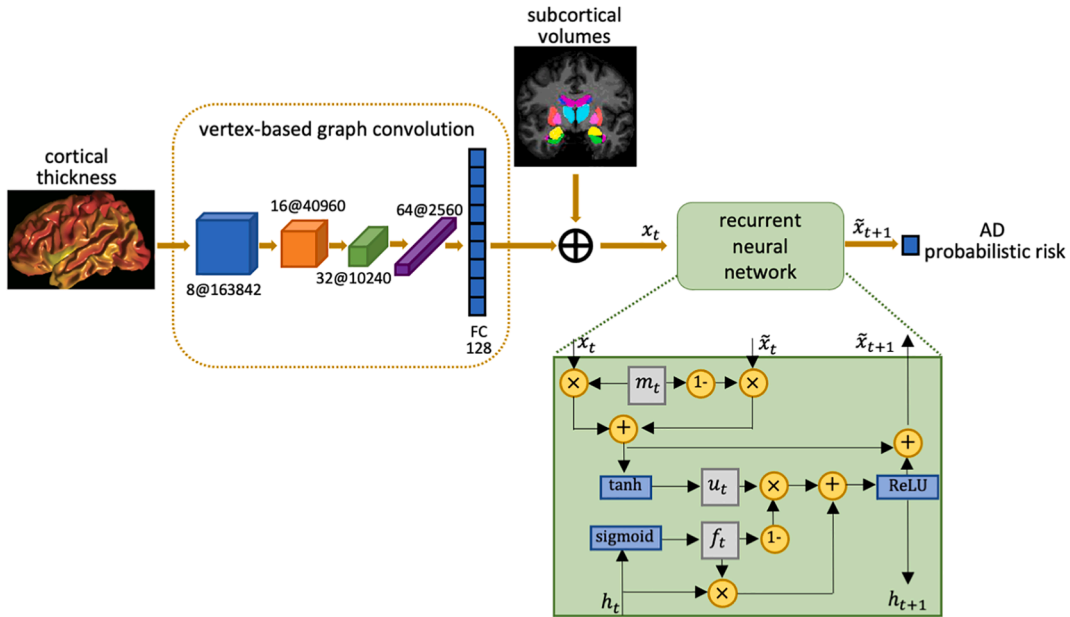


**Fig. 1.** The architecture of the vertex-based graph convolution and recurrent neural network (graph-CNN-RNN). The graph-CNN reduces the dimensionality of cortical thickness data and the minimal RNN updates anatomical measures and predicts the diagnosis of controls (CN), mild cognitive impairment (MCI), or Alzheimer's disease (AD) at each time point. $x_t$ denotes an observed data at time $t$, comprising 128 cortical features, subcortical volumes, and diagnosis. $\tilde{x}_{t+1}$ represents the predicted value of $x_t$ at time $t + 1$. The hidden state $h_t$ is a combination of the previous hidden state and the transformed input and hence encodes the longitudinal information of $x_t$ and $h_t$. The gate $m_t$ is on if $x_t$ exists. The forget gate $f_t$ weights the contribution of the previous hidden state $h_t$ and current transformed input $u_t$ toward the hidden state $h_{t+1}$ at time $t + 1$.

OASIS-3 sample.

We further applied the graph-CNN-RNN model to individual brain scans of conversion samples and computed its ability to predict future diagnosis (CN, MCI, or AD) of the ADNI sample. We defined the AD prediction accuracy as a function of years prior to AD onset. This study did not perform this experiment to the OASIS-3 sample due to its small sample size of conversion subjects at each time point (see Table 2), which cannot provide the precision of the prediction accuracy (1/the number conversion subjects at each time point).

### 2.3.3. AD probabilistic risk

We further defined an AD probabilistic risk as a ratio of the diagnostic probability of AD to the sum of the diagnostic probability of AD and CN from the graph-CNN-RNN model. It was a continuous scalar ranging from 0 to 1. The smaller value, the healthier the subject was. We used this AD probabilistic risk to represent brain morphological degeneration.

### 2.4. Diagnosis, cognition, and amyloid burden

This study included all available data of diagnosis, cognition, and biomarkers available in the ADNI and OASIS databases. For diagnostic measures, we included mini-mental state exam (MMSE), clinical dementia rating scale sum of boxes (CDR-SOB), the 13 item version of the Alzheimer's disease assessment scale-cognitive subscale (ADAS-Cog) at every time point. This study also included cognitive measures at all available time points. The ADNI study included Rey Auditory Verbal Learning Test immediate and delayed scores (RAVLT IR, RAVLT DR) for memory, category fluency and Boston naming tests for semantic memory and language, Trail Making Test A and B scores (TMT-A, TMT-B) for attention and executive function, and clock test for visuospatial function. The OASIS-3 study also included Wechsler Memory Scale (WMS)-Revised measures, LOGIMEM and MEMUNITS, for episodic memory.

Both ADNI and OASIS-3 assessed amyloid burden based on standard uptake value ratios (SUVRs) derived from [18F]-Florbetapir (AV45) positron emission tomography (Villars et al., 2010). AV45 SUVRs were computed as the value averaged over the cingulate, frontal, parietal, and temporal regions (Su et al., 2013; Su et al., 2019).

### 2.5. Statistical analysis

Partial correlation was used to examine the relationship of the AD probabilistic risk with diagnostic measures, cognition, and amyloid burden at each time point. Age and gender were considered as covariates. This partial correlation analysis was applied to both the ADNI testing dataset and OASIS-3 dataset to quantify the use of the AD probabilistic risk in associations with clinical and pathological measures.

### 2.6. Data availability

The ADNI and OASIS data are available at https://adni.loni.usc.edu and https://www.oasis-brains.org, respectively.

## 3. Results

### 3.1. Subject characteristics

This study included both the ADNI and OASIS-3 datasets. Tables 3 and 4 list the demographic, clinical, and cognitive information, as well as amyloid burden assessed using AV45 SUVRs at the baseline time point of the ADNI and OASIS-3 datasets, respectively. In the ADNI dataset, there were significant group differences in age, gender, and years of education among the four diagnostic groups at the baseline (all $p < 0.01$). The AD patients showed the worst memory, language, attention, and visuospatial functions, and the highest amyloid burden in the

**Table 3**

Demographic, clinical, cognition, and amyloid burden characteristics of the ADNI dataset at the baseline visit.

| | Control (n = 434) | s-MCI (n = 548) | AD (n = 326) | Conversion (n = 251) |
|---|---|---|---|---|
| Age (range, years) | 74.0 (59.7–90.1) | 73.1 (54.4–91.4) | 75.1 (55.1–92.3) | 73.7 (55–88.3) |
| Gender (females, %) | 52.5 | 42.2 | 44.8 | 41.0 |
| Education (SD, yrs) | 16.4 (2.70) | 15.8 (2.96) | 15.1 (2.94) | 15.9 (2.72) |
| **Clinical** | | | | |
| MMSE (SD) | 29.0 (1.14) | 27.9 (1.73) | 23.2 (2.22) | 26.9 (1.79) |
| CDR-SOB (SD) | 0.04 (0.14) | 1.34 (0.79) | 4.45 (1.79) | 1.84 (1.00) |
| ADAS-Cog (SD) | 8.85 (4.28) | 14.8 (6.03) | 30.1 (8.58) | 20.7 (6.04) |
| **Cognition** | | | | |
| **Memory** | | | | |
| RAVLT IR (SD) | 45.4 (9.77) | 36.6 (10.7) | 22.6 (7.60) | 28.6 (7.80) |
| RAVLT DR (SD) | 12.9 (2.59) | 11.2 (3.22) | 7.06 (3.90) | 9.15 (3.63) |
| **Language** | | | | |
| Animal fluency (SD) | 20.6 (5.45) | 17.7 (5.08) | 12.3 (5.08) | 15.5 (4.88) |
| Boston naming (SD) | 28.0 (2.60) | 26.5 (3.64) | 22.2 (6.14) | 25.4 (4.15) |
| **Visuospatial** | | | | |
| Clock (SD) | 4.68 (0.64) | 4.45 (0.79) | 3.34 (1.35) | 4.06 (1.10) |
| **Attention/ Executive** | | | | |
| TMT-A (SD) | 34.4 (12.0) | 39.7 (16.9) | 65.5 (36.0) | 46.9 (23.8) |
| TMT-B (SD) | 84.2 (42.9) | 108.7 (58.5) | 199.4 (86.4) | 141.3 (76.3) |
| **AV45 SUVR** | | | | |
| Cingulate (SD) | 1.41 (0.28) | 1.48 (0.30) | 1.70 (0.29) | 1.67 (0.28) |
| Frontal (SD) | 1.30 (0.25) | 1.37 (0.29) | 1.60 (0.28) | 1.57 (0.28) |
| Parietal (SD) | 1.32 (0.27) | 1.38 (0.29) | 1.60 (0.28) | 1.55 (0.26) |
| Temporal (SD) | 1.22 (0.23) | 1.28 (0.26) | 1.48 (0.27) | 1.46 (0.26) |

**Abbreviations:** MMSE, Mini-Mental State Examination; CDR-SOB, clinical dementia rating scale sum of boxes; ADAS-Cog, the Alzheimer's disease assessment scale-cognitive subscale;
RAVLT IR, Rey Auditory Verbal Learning Test (RAVLT) total immediate recall; RAVLT DR, RAVLT total delayed recognition; Clock, clock drawing; TMT-A, trail making test part A; TMT-B, trail making test part B; SD, standard deviation.

cingulate, frontal, parietal, and temporal brain regions compared to the s-MCI, conversion patients, and controls (all $p < 0.001$).

The OASIS-3 dataset showed significant group differences in baseline age ($p < 0.001$) and gender ($p = 0.006$) but not in years of education ($p = 0.812$). Similar to the ADNI dataset, the OASIS-3 dataset also showed the worst memory, language, and attention function as well as the highest amyloid burden in the cingulate, frontal, parietal, and temporal brain regions in the AD patients when compared to the other groups (all $p < 0.001$). Unlike the ADNI dataset, the OASIS-3 had a few conversion subjects at each time point that were not included in this study (see Table 2).

### 3.2. Diagnostic prediction of stable controls and AD patients

This study included 2491 subjects from the ADNI and OASIS-3 datasets (Tables 1-2,). For the ADNI dataset, each subject had multiple scans that were taken over the course of 9 ∼ 10 years. But, subjects with MRI scans more than 4 years were few (Table 1). For the OASIS-3 dataset, each subject had MRI scans over the course of 4 years, however, a few subjects had MRI scans over the course of more than 2 years (Table 2).

For the stable controls and AD classification, this study employed half of the ADNI sample to train the graph-CNN-RNN and the other half

**Table 4**
Demographic, clinical, cognition, and amyloid burden characteristics of the OASIS-3 dataset at the baseline visit.

| | Control (n = 650) | s-MCI (n = 47) | AD (n = 233) |
|---|---|---|---|
| Age (range, years) | 67.6 (42.5–97) | 75.1 (61.5–89.5) | 76.0 (49.5–95.5) |
| Gender (females, %) | 40.9 | 44.7 | 51.5 |
| Education (SD, yrs) | 15.6 (2.55) | 15.9 (2.27) | 16.3 (10.8) |
| **Clinical** | | | |
| **MMSE (SD)** | 29.0 (1.44) | 27.3 (2.66) | 22.9 (5.13) |
| **CDR-SOB (SD)** | 0.14 (0.91) | 1.46 (1.25) | 4.67 (3.24) |
| **Cognition** | | | |
| **Memory** | | | |
| WMS LOGIMEM (SD) | 13.8 (3.92) | 9.57 (4.72) | 5.68 (4.92) |
| WMS MEMUNITS (SD) | 12.9 (4.29) | 7.23 (4.62) | 3.58 (4.77) |
| **Language** | | | |
| Animal fluency (SD) | 20.4 (5.72) | 15.9 (5.22) | 12.2 (5.64) |
| Boston naming (SD) | 27.3 (3.09) | 24.8 (4.99) | 21.7 (6.28) |
| **Attention/Executive** | | | |
| TMT-A (SD) | 35.1 (17.3) | 43.0 (19.1) | 71.3 (46.3) |
| TMT-B (SD) | 91.1 (45.4) | 131.8 (83.4) | 188.0 (92.9) |
| **AV45 SUVR** | | | |
| Cingulate (SD) | 1.44 (0.51) | 2.83 (2.05) | 2.54 (1.01) |
| Frontal (SD) | 1.05 (0.52) | 2.30 (1.92) | 2.14 (0.88) |
| Parietal (SD) | 0.90 (0.45) | 1.84 (1.19) | 1.80 (0.75) |
| Temporal (SD) | 1.15 (0.49) | 2.75 (2.13) | 2.28 (0.71) |

**Abbreviations:** MMSE, Mini-Mental State Examination; CDR-SOB, clinical dementia rating scale sum of boxes; ADAS-Cog, the Alzheimer's disease assessment scale-cognitive subscale;
WMS, Wechsler Memory Scale; TMT-A, trail making test part A; TMT-B, trail making test part B; SD, standard deviation.

of the ADNI sample to validate the classification accuracy (eTable 1, Supplementary Material). Since the graph-CNN-RNN was a longitudinal model, the prediction of diagnosis at the current time was based on the MRI data and diagnosis information at the previous time points. Fig. 2**A** shows the classification accuracy between stable controls and AD patients of the ADNI dataset at each time point. The accuracy rate ranged from 85.4% to 92.3% and had relatively a small fluctuation over the 6 years from the baseline. Nevertheless, the 95% confidence interval increased over time due to the drop in the number of subjects available in the ADNI dataset (Fig. 2**B**).

Moreover, the robustness and reliability of the graph-CNN-RNN model were tested using the OASIS-3 dataset (Table 2). We directly applied the trained graph-CNN-RNN model to the OASIS-3 sample and achieved the classification accuracy ranging from 82.6% to 95.2% over the 2 years from the baseline (Fig. 2**C**). There was an increasing trend in the classification accuracy over time mainly because of the incorporation of longitudinal MRI scans despite the small sample size at 1.5-year and 2-year visits (Fig. 2**D** and Table 2).

*3.3. The discrimination map of stable controls and AD patients*

We evaluated the discriminative power of various brain regions for AD diagnosis over time. For this, we evaluated the contribution of each brain region to the AD classification as follows. The measure within a brain region was first set to the difference between the CN and AD subjects and the measure of the rest of brain regions was set to be the averaged value of the CN subjects. The graph-CNN-RNN model then computed the AD probabilistic risk of this region. We repeated this for all brain regions. Fig. 2**E** shows that the hippocampal volume was ranked with the highest discriminative power of AD diagnosis, followed by the cortical thickness, lateral ventricular, thalamus, and amygdala volumes (range: 0.020 ~ 0.475). Among cortical regions, the bilateral medial temporal lobes (MTL) had the highest discriminative power of AD diagnosis (Fig. 2**F**). Our findings were highly consistent to the

existing findings on the hippocampal and MTL atrophy in AD (Leung et al., 2013; Liedes et al., 2019; Qiu et al., 2009).

*3.4. The AD probabilistic risk predicts AD incident 4 years in advance*

We computed the AD probabilistic risk for the ADNI testing sample (eTable 1, Supplementary Material) and the OASIS-3 samples (Table 2) based on the graph CNN-RNN model. Fig. 3**A, B** illustrate the distribution of the AD probabilistic risk in each diagnostic group of the ADNI and OASIS-3 datasets, respectively. Both datasets showed the same trend, that is, AD patients had the largest AD probabilistic risk score, while normal controls had the lowest AD probabilistic risk score (all p < 0.001). Nevertheless, the AD probabilistic risk in each diagnostic group of the OASIS-3 dataset was smaller than the respective value in the ADNI dataset partly because of the differences in age and the sample size of s-MCI and AD patients in the two datasets. The standard deviations of clinical measures in the ADNI dataset were smaller than those in the OASIS-3 dataset (Tables 3 and 4).

We then applied the graph-CNN-RNN model to earlier MRI scans to compute the longitudinal trajectory of the probabilistic risk for the prediction of AD conversion at a later time point. Fig. 3**C** shows the prediction accuracy of the AD conversion from 0 to 4 years prior to the onset of AD. For instance, our model can predict the AD conversion at the accuracy of 79.7% while using the MRI scan acquired 4 years before the onset of AD. The prediction accuracy of the AD conversion was relatively consistent across 0 to 4 years prior to the onset of AD (range: 78.3% ~ 81.5%). When the sample size (Fig. 3**D**) was larger, the prediction accuracy was relatively higher. The 95% confidence interval of the prediction accuracy increased as the sample size reduced.

Fig. 4 illustrates the longitudinal trajectory of the AD probabilistic risk that was averaged over all conversion subjects in the ADNI dataset. This trajectory quantified brain morphology over the course of AD. This trajectory quantified the theoretical model of whole-brain morphology proposed by Frisoni et al. (2010) (Frisoni et al., 2010).

*3.5. The AD probabilistic risk in relation with diagnostic measures, cognition, and biological markers*

For the ADNI dataset, the higher AD probabilistic risk score was associated with lower MMSE, and higher CDR-SOB and ADAS-Cog scores at all the time points (Fig. 5**A** and eTable 2, Supplementary Material) after adjusting for age and gender. Fig. 5**B** shows that the higher AD probabilistic risk score was correlated with lower memory, language, attention, and visuospatial cognitive functions over the 5 years from the baseline visit (see statistical values and sample sizes in eTable 3, Supplementary Material). Similarly, the higher AD probabilistic risk score was associated with higher amyloid burdens in the cingulate, frontal, parietal, and temporal regions at most of the time points (Fig. 5**C**), except at 0.5, 1, and 5 years from the baseline due to the limited samples (eTable 5, Supplementary Material).

The above correlation patterns shown in the ADNI dataset were replicated in the OASIS-3 dataset (Fig. 5**D-F**). **ETables 2, 4, 5** (Supplementary Material) provided the detailed statistical values and sample sizes for the OASIS-3 dataset. Overall, the AD probabilistic risk was highly correlated with diagnostic status, cognitive ability, and AD neuropathology, suggesting its potential use in the integration with the clinical and pathological measures of AD.

**4. Discussion**

This study developed the graph-CNN-RNN model and defined a scalar measure, AD probabilistic risk, that well characterizes whole-brain morphology. The graph-CNN-RNN was a longitudinal model that incorporated longitudinal cortical and subcortical morphology for the diagnosis and prediction of AD. This study employed the two largest AD datasets with the most time points and demonstrated that the graph-
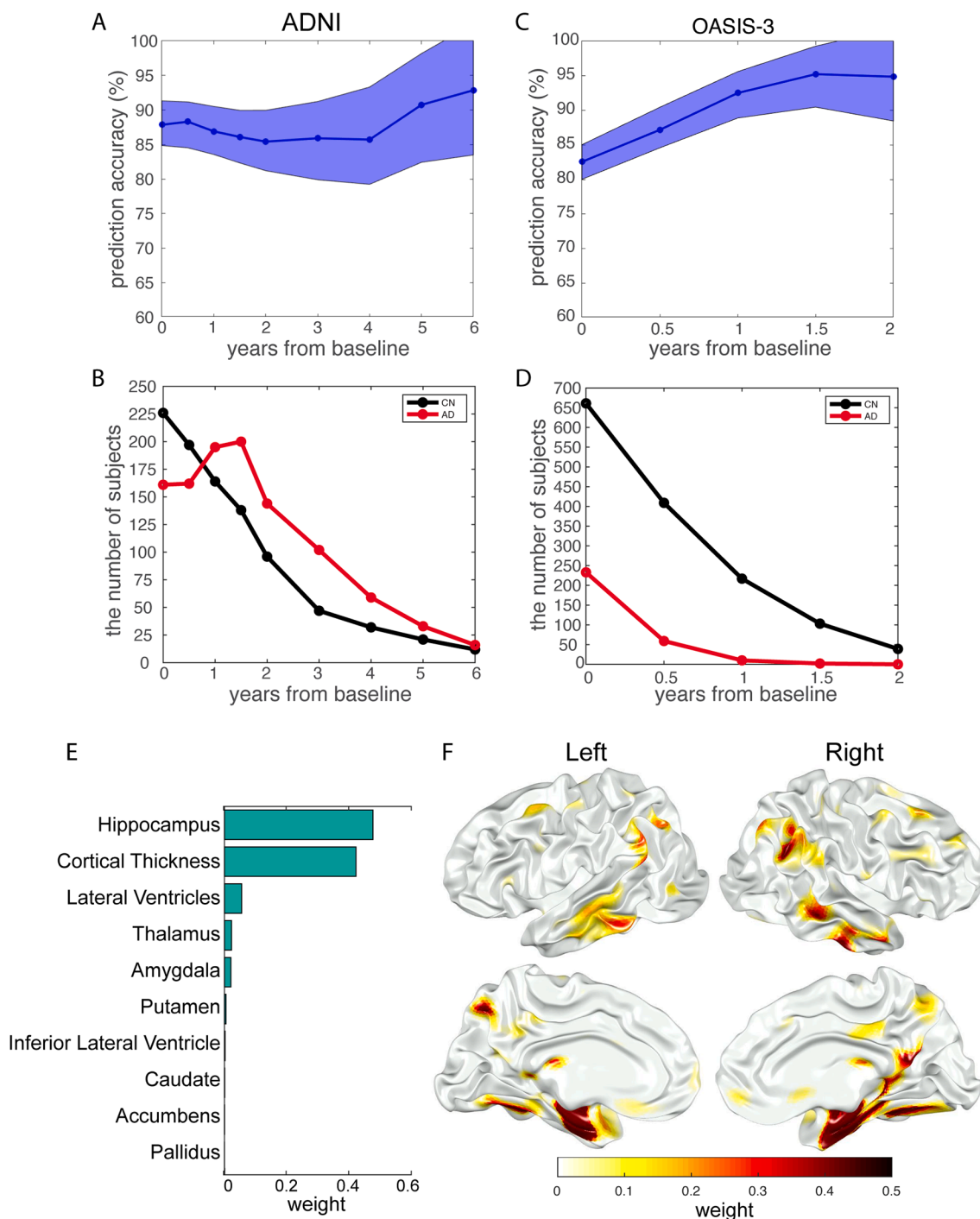
**Fig. 2.** Prediction accuracy of stable controls and AD patients and attention maps. A, B) The classification accuracy of stable controls and AD patients over time for the ADNI and OASIS-3 samples, respectively. The shading area represents the 95% confidence interval for the classification accuracy. C, D) show the sample sizes of stable controls and AD patients from the ADNI and OASIS-3 datasets used to test the robustness and generalizability of the graph-CNN-RNN model, respectively. E) The discriminative map indicates the contribution of various brain regions for the AD diagnosis over time. F) The discriminative map indicates the contribution of cortical thickness for the AD diagnosis over time.

CNN-RNN can reliably and robustly diagnose AD at the accuracy rate of 85% and above at all the time points. The graph-CNN-RNN predicted the AD conversion 4 years prior to the AD onset at ~ 80% of accuracy. The AD probabilistic risk was associated with clinical traits, cognition, and amyloid burden across all the time points. Our findings suggested the possibility of constructing a scalar measure to quantify brain morphology and successfully predict the onset of AD 4 years prior to clinical diagnosis.

Deep learning methods have increasingly been used in the computer-

aided diagnosis of AD due to their flexibility and ability to learn neuroimage features that have the most discriminative power of AD diagnosis (e.g., Ansart et al., 2021; Jin et al., 2020; Wen et al., 2020). In the past 10 years, a substantial body of research mainly employed convolutional neural networks (CNNs) on 2D slices, 3D patches/regions of interest (ROIs), or volumetric features of cross-sectional data for AD diagnosis. However, due to participant selection, image processing, sample size, or validation procedure across studies (e.g., Basaia et al., 2019; Jin et al., 2020; Lian et al., 2020; Liu et al., 2018;
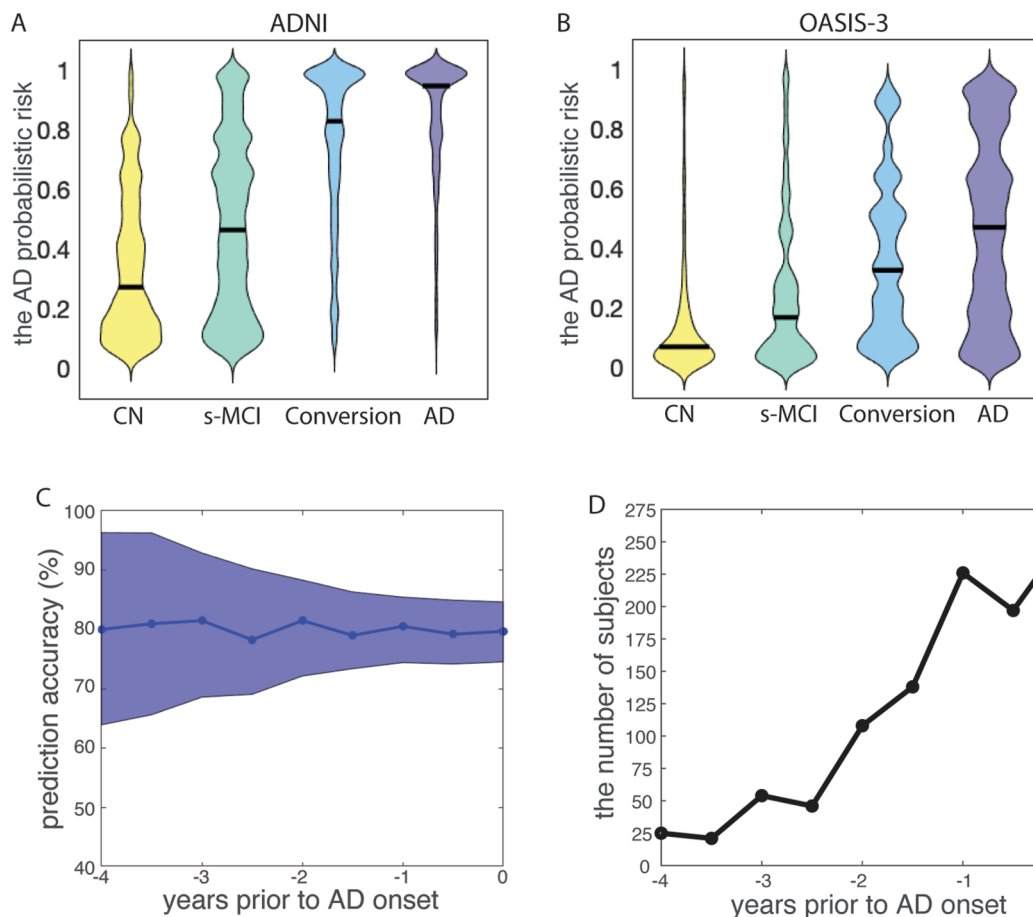
**Fig. 3.** The distributions of the AD probabilistic risk and the prediction accuracy of AD conversion. A, B) The distributions of the AD probabilistic risk in each diagnostic group of the ADNI and OASIS-3 datasets, respectively. C, D) The prediction accuracy of the AD conversion prior to the AD onset and the number of the AD conversion subjects used.
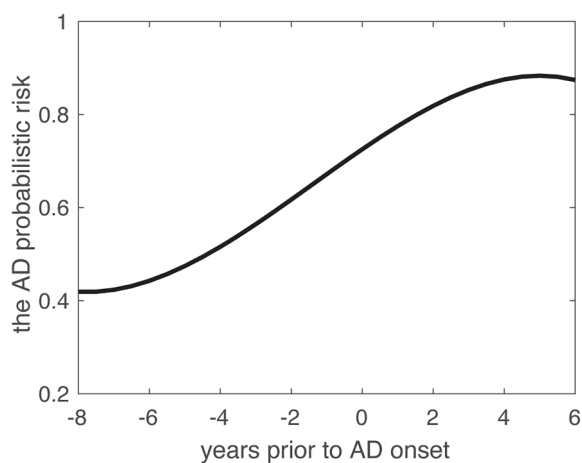


**Fig. 4.** The longitudinal trajectory of the AD probabilistic risk averaged over all conversion subjects in the ADNI dataset.

Wee et al., 2019), their classification results may not be comparable. A recent review (Wen et al., 2020) implemented the existing CNN models, such as CNN on 2D slices, 3D patches/regions of interest (ROIs), or 3D images to overcome the above variations and provided the most comparable classification results across the existing CNN models in literature. It demonstrated that 3D CNN approaches (3D images, 3D-ROI, 3D-patch) achieved the best classification performance between CN and AD

(accuracy: 75% ~ 90%). Instead of 3D volumetric images, our graph-CNN-RNN model took an advantage of cortical geometry and applied the convolution along the cortical surface. Moreover, our model comprised the recurrent neural network and incorporated the longitudinal information of the data, which allowed the interpolation of missing data points based on previous time points and stabilized the classification performance across all the time points. In other words, our model provided a diagnosis of individual subjects at any time point. Even though our model was trained on half of the ADNI sample, both the ADNI and OASIS-3 datasets showed similar performance (accuracy: 85.4% ~ 92.3% for ADNI; 82.6% ~ 95.2% for OASIS-3) across all the time points, suggesting that our graph-CNN-RNN model was reproducible, robust, and generalizable.

Structural MRI has become a part of the clinical assessment of AD. Frisoni et al. (Frisoni et al., 2010) proposed a theoretical model on the progression of AD biological markers from preclinical to overt stages of AD. Nevertheless, there is still a lack of the quantification of these biomarkers over the course of AD. Brain morphological changes have mainly been restricted to the volumes of whole-brain or ROIs, such as the hippocampus and medial temporal structures. This study employed the deep learning approach and encoded whole brain morphology as a scalar measure, AD probabilistic risk. Our study provided the quantification of the longitudinal trajectory of brain morphology over the course of AD. This trajectory allowed early prediction of AD and achieved the prediction accuracy of ~ 80% from 0 to 4 years before clinical diagnosis. Moreover, the trajectory of the AD probabilistic risk was highly correlated not only with the trajectory of clinical traits and cognition but also with the trajectories of AD pathology, amyloid burden. These findings
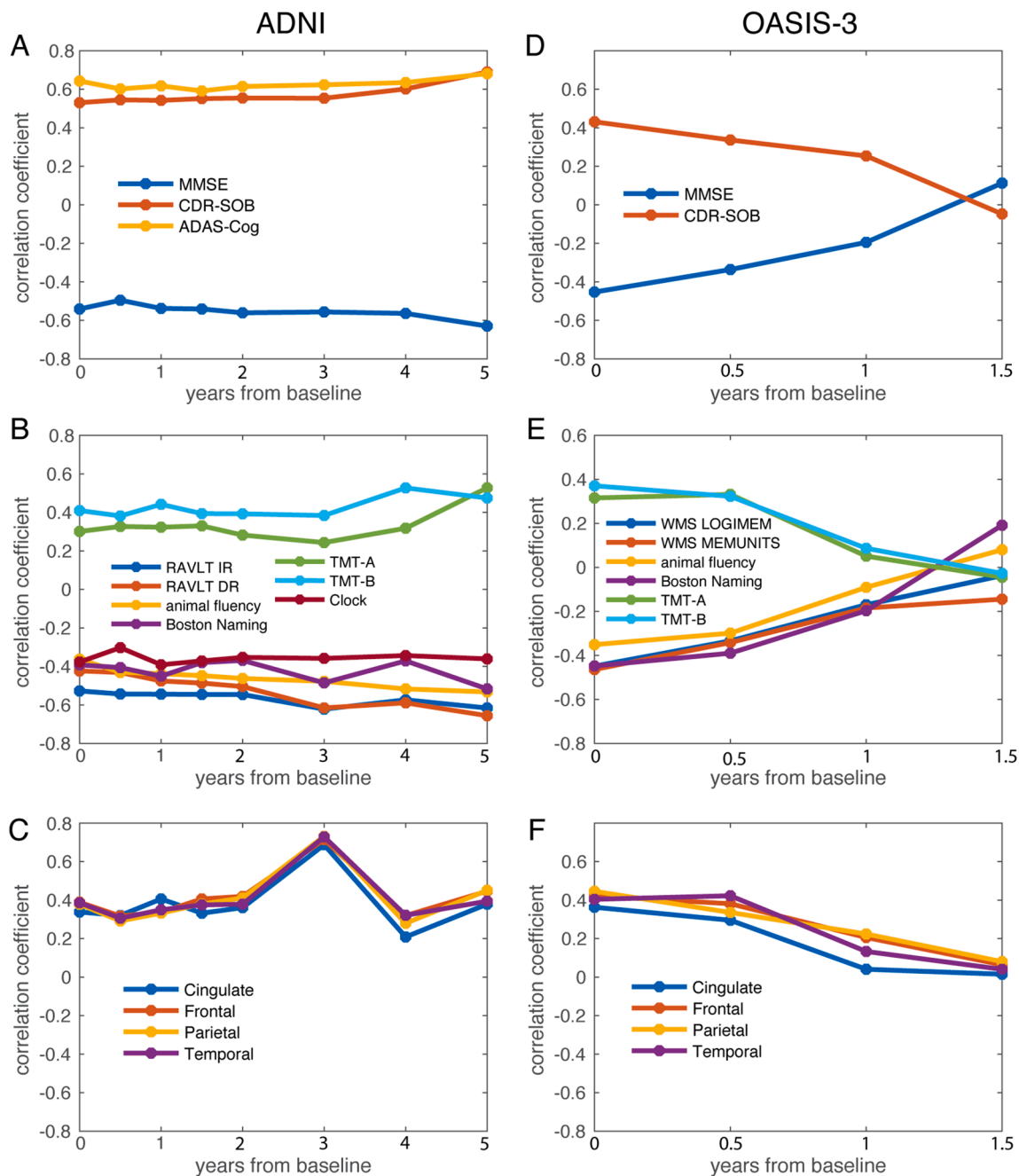
**Fig. 5.** Correlation of the AD probabilistic risk with diagnosis, cognition, and [18F]-Florbetapir (AV45) standard uptake ratio (SUVR) at each time point of the ADNI and OASIS-3 datasets. The left column shows the results for the ADNI dataset, while the right column illustrates the results for the OASIS-3 dataset. The sample sizes of each time points are listed in **eTables 2**–**5** (Supplementary Material). **Abbreviations:** MMSE, mini mental state exam; CDR-SOB, clinical dementia rating scale sum of boxes; ADAS-Cog, the Alzheimer's disease assessment scale-cognitive subscale; RAVLT IR and DR, Rey Auditory Verbal Learning Test Immediate and Delayed scores; TMT-A and TMT-B, Trail Making Test A and B scores; WMS, Wechsler Memory Scale.

implicated the feasibility of using this graph-CNN-RNN as a computer-aided tool for the prognosis of AD in clinical practice.

While the experimental results highlighted a promising potential of clinical diagnosis and prognosis applications of the graph-CNN-RNN for AD classification and prediction, this study has some limitations that warrant consideration. The graph-CNN-RNN model achieved good prediction results based on the ADNI dataset. Nevertheless, there is lack of longitudinal deep learning methods for fair comparison. The OASIS-3 dataset was lack of conversion subjects with multiple time points. The performance and robustness of the proposed model should be further tested on longitudinal datasets with larger samples. Moreover, as multi-modal brain image data become available, our model can be extended to

other image modalities and create a functional, metabolic, and pathological signature for the diagnosis and prognosis of AD, which can facilitate the quantification of the trajectory of image biomarkers over the course of AD.

## 5. Conclusion

This study offered a promising deep learning tool for individualized diagnosis and prognosis over the course of AD. This study employed the two major AD datasets longitudinally, defined the AD probabilistic risk, and quantified the trajectory of whole-brain morphology via a scalar measure from prognosis to overt stages of AD. Such a deep learning tool

and the AD probabilistic risk have great potential in clinical practice.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2022.102993.

## References

Ansart, M., Epelbaum, S., Bassignana, G., Bône, A., Bottani, S., Cattai, T., Couronné, R., Faouzi, J., Koval, I., Louis, M., Thibeau-Sutre, E., Wen, J., Wild, A., Burgos, N., Dormont, D., Colliot, O., Durrleman, S., 2021. Predicting the progression of mild cognitive impairment using machine learning: A systematic, quantitative and critical review. Med. Image Anal. 67, 101848.

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., Filippi, M., Neuroimaging, A.D., I.,, 2019. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. Neuroimage Clin. 21, 101645.

Basheera, S., Sai Ram, M.S., 2019. Convolution neural network-based Alzheimer's disease classification using hybrid enhanced independent component analysis based segmented gray matter of T2 weighted magnetic resonance imaging with clinical valuation. Alzheimers Dement (N Y) 5 (1), 974–986.

Bashyam, V.M., Erus, G., Doshi, J., Habes, M., Nasrallah, I., Truelove-Hill, M., Srinivasan, D., Mamourian, L., Pomponio, R., Fan, Y., Launer, L.J., Masters, C.L., Maruff, P., Zhuo, C., Volzke, H., Johnson, S.C., Fripp, J., Koutsouleris, N., Satterthwaite, T.D., Wolf, D., Gur, R.E., Gur, R.C., Morris, J., Albert, M.S., Grabe, H. J., Resnick, S., Bryan, R.N., Wolk, D.A., Shou, H., Davatzikos, C., 2020. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. Brain 143, 2312–2324.

Chan, D., Fox, N.C., Scahill, R.I., Crum, W.R., Whitwell, J.L., Leschziner, G., Rossor, A.M., Stevens, J.M., Cipolotti, L., Rossor, M.N., 2001. Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. Ann. Neurol. 49 (4), 433–442.

Chen, M., 2017. Minimalrnn: Toward more interpretable and trainable recurrent neural networks. arXiv:1711.06788.

Davatzikos, C., Bhatt, P., Shaw, L.M., Batmanghelich, K.N., Trojanowski, J.Q., 2011. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiol. Aging 32 (2322), e2319–2327.

Du, J., Younes, L., Qiu, A., 2011. Whole brain diffeomorphic metric mapping via integration of sulcal and gyral curves, cortical surfaces, and images. Neuroimage 56 (1), 162–173.

Dubois, B., Feldman, H.H., Jacova, C., DeKosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P.J., Scheltens, P., 2007. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol. 6 (8), 734–746.

Eskildsen, S.F., Coupe, P., Fonov, V.S., Pruessner, J.C., Collins, D.L., Neuroimaging, A.D., I., 2015. Structural imaging biomarkers of Alzheimer's disease: predicting disease progression. Neurobiol. Aging 36 (Suppl 1), S23–31.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33 (3), 341–355.

Frisoni, G.B., Fox, N.C., Jack, C.R., Scheltens, P., Thompson, P.M., 2010. The clinical use of structural MRI in Alzheimer disease. Nat. Rev. Neurol. 6 (2), 67–77.

Jack Jr., C.R., Lowe, V.J., Weigand, S.D., Wiste, H.J., Senjem, M.L., Knopman, D.S., Shiung, M.M., Gunter, J.L., Boeve, B.F., Kemp, B.J., Weiner, M., Petersen, R.C., Alzheimer's Disease Neuroimaging, I., 2009. Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease. Brain 132, 1355–1365.

Jin, D., Zhou, B., Han, Y., Ren, J., Han, T., Liu, B., Lu, J., Song, C., Wang, P., Wang, D., Xu, J., Yang, Z., Yao, H., Yu, C., Zhao, K., Wintermark, M., Zuo, N., Zhang, X., Zhou, Y., Zhang, X., Jiang, T., Wang, Q., Liu, Y., 2020. Generalizable, Reproducible, and Neuroscientifically Interpretable Imaging Biomarkers for Alzheimer's Disease. Adv. Sci. (Weinh) 7 (14), 2000675.

Leung, K.K., Bartlett, J.W., Barnes, J., Manning, E.N., Ourselin, S., Fox, N.C., Neuroimaging, A.D., I.,, 2013. Cerebral atrophy in mild cognitive impairment and Alzheimer disease: rates and acceleration. Neurology 80, 648–654.

Lian, C., Liu, M., Zhang, J., Shen, D., 2020. Hierarchical Fully Convolutional Network for Joint Atrophy Localization and Alzheimer's Disease Diagnosis Using Structural MRI. IEEE Trans. Pattern Anal. Mach. Intell. 42 (4), 880–893.

Liedes, H., Lötjönen, J., Kortelainen, J.M., Novak, G., van Gils, M., Gordon, M.F., 2019. Multivariate Prediction of Hippocampal Atrophy in Alzheimer's Disease. J. Alzheimers Dis. 68 (4), 1453–1468.

Liu, C., Ji, H., Qiu, A., 2021. Fast vertex-based graph convolutional neural network and its application to brain images. Neurocomputing 434, 1–10.

Liu, M., Li, F., Yan, H., Wang, K., Ma, Y., Alzheimer's Disease Neuroimaging, I., Shen, L., Xu, M., 2020. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage 208, 116459.

Liu, M., Zhang, J., Adeli, E., Shen, D., 2018. Landmark-based deep multi-instance learning for brain disease diagnosis. Med. Image Anal. 43, 157–168.

Liu, M., Zhang, J., Adeli, E., Shen, D., 2019. Joint classification and regression via deep multi-task multi-channel learning for alzheimer's disease diagnosis. IEEE Trans. Biomed. Eng. 66 (5), 1195–1206.

Marcus, D.S., Fotenos, A.F., Csernansky, J.G., Morris, J.C., Buckner, R.L., 2010. Open access series of imaging studies: longitudinal MRI data in nondemented and demented older adults. J. Cogn. Neurosci. 22, 2677–2684.

Modrego, P.J., 2006. Predictors of conversion to dementia of probable Alzheimer type in patients with mild cognitive impairment. Curr. Alzheimer Res. 3, 161–170.

Modari, E., Pepe, A., Gaser, C., Huttunen, H., Tohka, J., Alzheimer's Disease Neuroimaging, I., 2015. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. Neuroimage 104, 398–412.

Qiu, A., Fennema-Notestine, C., Dale, A.M., Miller, M.I., Alzheimer's Disease Neuroimaging, I., 2009. Regional shape abnormalities in mild cognitive impairment and Alzheimer's disease. Neuroimage 45, 656–661.

Qiu, S., Joshi, P.S., Miller, M.I., Xue, C., Zhou, X., Karjadi, C., Chang, G.H., Joshi, A.S., Dwyer, B., Zhu, S., Kaku, M., Zhou, Y., Alderazi, Y.J., Swaminathan, A., Kedar, S., Saint-Hilaire, M.H., Auerbach, S.H., Yuan, J., Sartor, E.A., Au, R., Kolachalama, V.B., 2020. Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. Brain 143, 1920–1933.

Risacher, S.L., Saykin, A.J., West, J.D., Shen, L., Firpi, H.A., McDonald, B.C., Alzheimer's Disease Neuroimaging, I., 2009. Baseline MRI predictors of conversion from MCI to probable AD in the ADNI cohort. Curr. Alzheimer Res. 6, 347–361.

Sluimer, J.D., Bouwman, F.H., Vrenken, H., Blankenstein, M.A., Barkhof, F., van der Flier, W.M., Scheltens, P., 2010. Whole-brain atrophy rate and CSF biomarker levels in MCI and AD: a longitudinal study. Neurobiol. Aging 31 (5), 758–764.

Su, Y., D'Angelo, G.M., Vlassenko, A.G., Zhou, G., Snyder, A.Z., Marcus, D.S., Blazey, T. M., Christensen, J.J., Vora, S., Morris, J.C., Mintun, M.A., Benzinger, T.L.S., Chen, K., 2013. Quantitative analysis of PiB-PET with FreeSurfer ROIs. PLoS ONE 8 (11), e73377.

Su, Y., Flores, S., Wang, G., Hornbeck, R.C., Speidel, B., Joseph-Mathurin, N., Vlassenko, A.G., Gordon, B.A., Koeppe, R.A., Klunk, W.E., Jack, C.R., Farlow, M.R., Salloway, S., Snider, B.J., Berman, S.B., Roberson, E.D., Brosch, J., Jimenez-Velazques, I., Dyck, C.H., Galasko, D., Yuan, S.H., Jayadev, S., Honig, L.S., Gauthier, S., Hsiung, G.R., Masellis, M., Brooks, W.S., Fulham, M., Clarnette, R.,

Masters, C.L., Wallon, D., Hannequin, D., Dubois, B., Pariente, J., Sanchez-Valle, R., Mummery, C., Ringman, J.M., Bottlaender, M., Klein, G., Milosavljevic-Ristic, S., McDade, E., Xiong, C., Morris, J.C., Bateman, R.J., Benzinger, T.L.S., 2019. Comparison of Pittsburgh compound B and florbetapir in cross-sectional and longitudinal studies. Alzheimers Dement (Amst) 11 (1), 180–190.

Tan, M., Qiu, A., 2016. Large deformation multiresolution diffeomorphic metric mapping for multiresolution cortical surfaces: a coarse-to-fine approach. IEEE Trans. Image Process. 25 (9), 4061–4074.

Thompson, P.M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M., Toga, A.W., 2003. Dynamics of gray matter loss in Alzheimer's disease. J. Neurosci. 23 (3), 994–1005.

Villars, H., Oustric, S., Andrieu, S., Baeyens, J.P., Bernabei, R., Brodaty, H., Brummel-Smith, K., Celafu, C., Chappell, N., Fitten, J., Frisoni, G., Froelich, L., Guerin, O., Gold, G., Holmerova, I., Iliffe, S., Lukas, A., Melis, R., Morley, J.E., Nies, H., Nourhashemi, F., Petermans, J., Ribera Casado, J., Rubenstein, L., Salva, A., Sieber, C., Sinclair, A., Schindler, R., Stephan, E., Wong, R.Y., Vellas, B., 2010. The primary care physician and Alzheimer's disease: an international position paper. J. Nutr. Health Aging 14 (2), 110–120.

Wee, C.Y., Liu, C., Lee, A., Poh, J.S., Ji, H., Qiu, A., Neuroimage, A.D., I.., 2019. Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations. Neuroimage Clin 23, 101929.

Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-Gonzalez, J., Routier, A., Bottani, S., Dormont, D., Durrleman, S., Burgos, N., Colliot, O., Alzheimer's Disease Neuroimaging, I., Australian Imaging, B., Lifestyle flagship study of, a., 2020. Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation. Med Image Anal 63, 101694.

Ye, J., Farnum, M., Yang, E., Verbeeck, R., Lobanov, V., Raghavan, N., Novak, G., DiBernardo, A., Narayan, V.A., Alzheimer's Disease Neuroimaging, I., 2012. Sparse learning and stability selection for predicting MCI to AD conversion using baseline ADNI data. BMC Neurol. 12, 46.

Yuan, Y., Gu, Z.-X., Wei, W.-S., 2009. Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis. AJNR Am. J. Neuroradiol. 30 (2), 404–410.

Zhang, D., Shen, D., Alzheimer's Disease Neuroimaging, I., 2012. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS ONE 7, e33182.

Zhong, J., Qiu, A., 2010. Multi-Manifold Diffeomorphic Metric Mapping for Aligning Cortical Hemispheric Surfaces. NeuroImage 49 (1), 355–365.