Geriatrics Gerontology International

DOI: 10.1111/ggi.14670

ORIGINAL ARTICLE

EPIDEMIOLOGY CLINICAL PRACTICE AND HEALTH

# Predicting mild cognitive impairment in older adults: A machine learning analysis of the Alzheimer's Disease Neuroimaging Initiative

Yafei Wu,[1,2] Xing Wang[1,2] and Ya Fang[1,2,3]

[1]School of Public Health, Xiamen University, Xiamen, China
[2]Key Laboratory of Health Technology Assessment of Fujian Province, Xiamen, China
[3]National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen, China

**Correspondence**
Ya Fang PhD, School of Public Health, Xiamen University, Xiang'an South Road, Xiang'an District, Xiamen, Fujian, China.
Email: fangya@xmu.edu.cn

*Received: 13 July 2023*
*Revised: 9 August 2023*
*Accepted: 31 August 2023*

**Aim:** Mild cognitive impairment (MCI) in older adults is potentially devastating, but an accurate prediction model is still lacking. We hypothesized that neuropsychological tests and MRI-related markers could predict the onset of MCI early.

**Methods:** We analyzed data from 306 older adults who were cognitive normal (CN) attending the Alzheimer's Disease Neuroimaging Initiative sequentially (474 pairs of visits) within 3 years. There were 231 pairs of MCI conversion (CN to MCI), and 242 pairs of CN maintenance (CN to CN). Variables on demographic, neuropsychological tests, genetic, and MRI-related markers were collected. Machine learning was used to construct MCI prediction models, comparing the area under the receiver operating characteristic curve (AUC) as the primary metric of performance. Important predictors were ranked for the optimal model.

**Results:** The baseline age of the study sample was 74.8 years old. The best-performing model (gradient boosting decision tree) with 13 variables predicted MCI with an AUC of 0.819, and the rank of variable importance showed that intracranial volume, hippocampal volume, and score from task 4 (word recognition) of the Alzheimer's Disease Assessment Scale were important predictors of MCI.

**Conclusions:** With the help of machine learning, fewer neuropsychological tests and MRI-related markers are required to accurately predict MCI within 3 years, thereby facilitating targeted intervention. **Geriatr Gerontol Int 2023; ••: ••–••.**

**Keywords:** Alzheimer's disease, machine learning, mild cognitive impairment, predictors.

## Introduction

Alzheimer's disease (AD) is the most common neurodegenerative disease, with insidious onset and slow progression in the elderly population. AD accounts for about 60%–80% of all dementia etiologies.[1] The degree of cognitive decline in AD patients usually starts slowly, and this change will gradually accelerate as time progresses.[2] Physicians and other caregivers monitor the progression of AD by assessing the extent of patients' cognitive decline, which typically falls into three categories: cognitive normal (CN), mild cognitive impairment (MCI) and dementia.[3] Both MCI and dementia patients suffer from cognitive decline.

Unfortunately, at present, there is no cure for AD, and although the drugs approved by the US Food and Drug Administration (FDA) can improve the symptoms of patients, they cannot block the progression of AD.[4] Patients are often diagnosed with AD when there is significant cognitive decline, in which case the diagnosis becomes too late to implement intervention programs to alleviate or reverse the patient's condition. Studies have shown that MCI may be the prodromal stage of AD, and some MCI patients will progress to AD.[5] Pharmacological and other interventions have been shown to reduce or reverse symptoms of MCI patients.[6–8] The latest edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) includes a special category for such patients called mild neurocognitive disorder, similar to MCI.[9] In view of this, many studies have focused on detecting MCI that has not yet reached AD, in the hope of early intervention to prevent patients from further progressing to AD or experiencing adverse effects in their daily lives. At a later stage, when some cognitive impairment has occurred, MCI can be diagnosed by neuropsychological tests, and by demographical and imaging markers.[10,11] However, early diagnosis or prediction of MCI in the normal population remains difficult.

To overcome these challenges, machine learning (ML) has been widely used to classify or predict MCI and AD.[12,13] ML, a multifield interdisciplinary subject, can learn the intrinsic patterns of data automatically. Moreover, ML can be used to solve complex nonlinearity and collinearity, and has the characteristic of high prediction accuracy. For example, in the task of distinguishing between MCI and CN, it is difficult for people to identify the small differences between them in the early stages. However, ML algorithms can extract relevant features of cognitive impairment, optimize the parameters of algorithms by training part of the randomly divided data, and then apply the trained algorithms to the new data to validate models' prediction accuracy, so as to help decision-making.[14,15]

Most recent studies have focused on the differentiation of MCI patients who will progress to AD, with few studies on the progression from healthy people to MCI.[16–18] Therefore, the purpose of this study is to construct ML models to predict the progression of MCI in the CN population within 3 years based on the ADNI database, and further explore key predictors, in order to help

clinical or community workers screen or intervene the high-risk MCI population early.

## Methods

### *Study sample*

Data used in this study were obtained from the ADNI database, which is public and has been ethically approved, no additional ethical approval was required for the current study. We obtained samples from four waves of ADNI (ADNI1, ADNIG0, ADNI2, and ADNI3). This study used sequential visits as the unit of analysis to take advantage of repeated cognitive assessments in the cohort. We named the first visit in a pair as the "index visit", whose data were used as the input to make a prediction of the subsequent cognitive status. For MCI participants with more than two visits, multiple visits were included in the prediction models. We included all participants aged 60 years and above who had sequential visits within 3 years and were cognitively normal in the index visits from all four waves. Finally, 306 participants with a total of 474 pairs of visits (231 pairs of MCI conversion and 242 pairs of CN maintenance) were included. The outcome was defined as incident MCI in the subsequent visits. The detailed sample selection process is shown in Fig. 1.

### *Measurements of predictors and preprocessing*

We included 117 variables from the ADNI database, one of which was the outcome to be predicted. We first excluded variables that were not helpful for model prediction, such as date, time and id number, and then excluded variables with more than 20% of values missing (Tables S1,S2). Finally, 28 variables from demographic, neuropsychological test, gene, and MRI-related information were retained, as shown in Table 1. The description of neuropsychological tests and MRI markers is shown in Text S1.

Data were divided into training (60%), validation (20%) and test (20%) parts, which were used for model training, hyperparameter tuning, and internal validation, respectively. For variables with a missing rate of 20% or less, the MissForest algorithm was applied for imputation. MissForest outperforms other methods such as multiple imputation, especially in data settings where complicated interactions and nonlinear relations are suspected.[19] We conducted one-hot coding for categorical variables and normalized the continuous variables. In the training set, three representative feature selection methods were applied to select fewer but more important variables, namely the variance filtering method, recursive feature elimination (RFE) and lasso cross validation (LassoCV), respectively, and the detailed results are shown in Figs S1–S3.

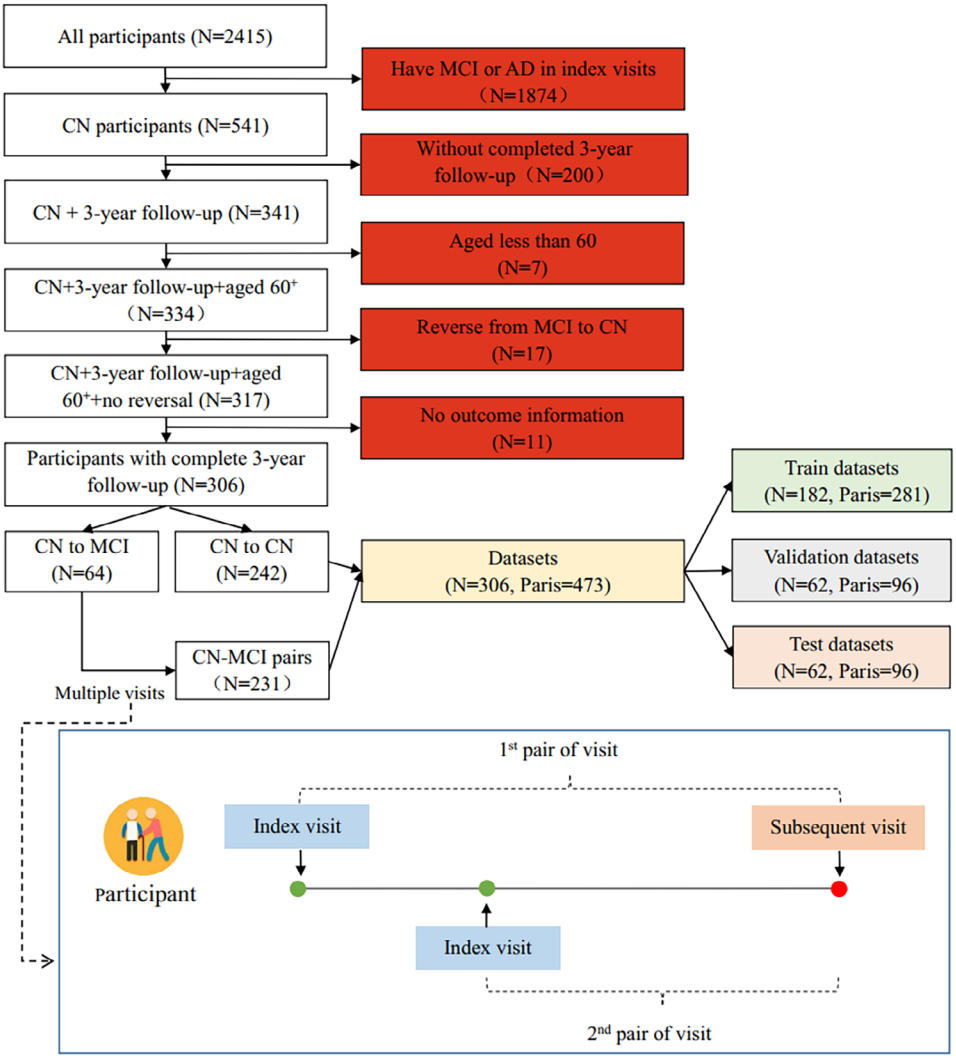


**Figure 1** Description of the sampling process. All samples were selected within 3 years.

**Table 1** Reserved variables for model construction

| Categories | Number | Variables |
|---|---|---|
| Demographics | 6 | AGE, PTGENDER, PTEDUCAT, PTETHCAT, PTRACCAT, PTMARRY |
| Neuropsychological tests | 14 | CDRSB, ADAS11, ADAS13, ADASQ4, MMSE, RAVLT_immediate, RAVLT_learning, RAVLT_forgetting, RAVLT_perc_forgetting, LDELTOTAL, TRABSCOR, FAQ, mPACCdigit, mPACCtrailsB |
| Genetic factors | 1 | APOE-4 |
| MRI-related markers | 7 | Ventricles, Hippocampus, WholeBrain, Entorhinal, Fusiform, MidTemp, ICV |

PTGENDER, Gender; PTEDUCAT, Education Level; PTETHCAT, Ethnicity; PTRACCAT, Race; PTMARRY, marital status; ADAS11, Unweighted sum of 11 items from the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog); ADAS13, Unweighted sum of 13 items from ADAS-Cog; ADASQ4, Score from Task 4 (word recognition) of the Alzheimer's Disease Assessment Scale (ADAS); CDRSB, Clinical Dementia Rating-Sum of Boxes Score; FAQ, Functional Activities Questionnaire; LDELTOTAL, Delayed Total Recall; MMSE, Mini-Mental State Examination; RAVLT_forgetting, Rey's Auditory Verbal Learning Test-Forgetting score; RAVLT_immediate, Rey's Auditory Verbal Learning Test-Immediate Recall score; RAVLT_percentage_forgetting, Rey's Auditory Verbal Learning Test-Percent Forgetting; mPACCdigit, Modified Preclinical Alzheimer Cognitive Composite with Digit Test; mPACCtrailsB, Modified Preclinical Alzheimer Cognitive Composite with Trails Test; Ventricles, volume of ventricles; WholeBrain, volume of whole brain; ICV, intracranial volume; Hippocampus, hippocampal volume; Entorhinal, volume of entorhinal cortex; Fusiform, volume of fusiform gyrus; MidTemp, volume of the middle temporal lobe.

### *Statistical analysis*

SPSS 26 was used for statistical analysis. Continuous variables were expressed as mean ± standard deviation ($\overline{x} \pm s$) or median and interquartile range (IQR), and categorical variables were expressed as number and ratio (%). Frequencies and percentages of categorical variables were compared using the chi-square test or Fisher exact test. For continuous variables, the *t*-test or nonparametric Kruskal–Wallis test was considered. *P*-values less than 0.05 were considered statistically significant.

According to the No Free Lunch theorem,[20] 10 machine learning models (Text S2) were chosen to predict MCI. We evaluated the model by using discrimination and calibration. For discrimination evaluation, area under the receiver operating characteristic curve (AUC) was used as the primary metric. In addition, the accuracy, precision, recall, and F1-score were also evaluated; the calculation formulas of these indexes are shown in Table S3. To avoid the randomness of model predictions, we computed the mean and standard deviation of all the above metrics after 100-time iterations. The calibration curve was plotted to evaluate the consistency of predicted and observed probabilities. In addition, we also analyzed the important predictors of the optimal model with two commonly used methods including feature importance and SHapley Additive exPlanations (SHAP), which can uncover the black-box of ML predictions. All the packages used in analysis are shown in Table S4.

## Results

### *Sample characteristics*

Table S5 shows the comparison of included variables between CN ($n = 242$) and MCI ($n = 231$). Age and race were statistically significant between CN and MCI, but sex, marriage, education, and ethnicity were not. All neuropsychological tests were statistically significant between CN and MCI. There was no significant difference in apolipoprotein (APOE) genotype between groups. Except for intracranial volume, other MRI-related features showed statistical significance between groups.

### *Feature selection and model performance*

Tables 2 and Table S6 show the feature selection results. Finally, we included 13 predictors by LassoCV because of the relatively fewer variables and better accuracy compared to the feature-free selection and other two feature selection methods (VT and RFE).

As shown in Table 3. We found that gradient boosting decision tree (GBDT) showed the best performance in predicting MCI, with its accuracy, precision, recall, F1-score, and AUC achieving 0.74, 0.74, 0.74, 0.74, and 0.82, respectively. We evaluated the calibration degree of the GBDT (Fig. S4), and the result showed that there was relatively small deviation from the perfect calibrated line. The AUC curve of the GBDT model is shown in Fig. S5.

### *Feature importance of optimal model*

Figure 2 shows the variable importance of the GBDT model. For the default importance of GBDT (feature importance), the top five predictors were the hippocampus, score from task 4 of the Alzheimer's Disease Assessment Scale (ADASQ4), intracranial volume (ICV), ventricles and Unweighted sum of 13 items from

**Table 2** Comparison of the accuracy in validation sets for different feature selection methods

| Model | Feature selection methods | | | |
|---|---|---|---|---|
| | Feature-free selection | VT | RFE | LassoCV |
| KNN | 0.632 | 0.658 | 0.697 | 0.789 |
| SVC | 0.671 | 0.671 | 0.737 | 0.737 |
| DT | 0.658 | 0.645 | 0.658 | 0.658 |
| RF | 0.724 | 0.803 | 0.737 | 0.710 |
| LR | 0.750 | 0.737 | 0.750 | 0.710 |
| MLP | 0.750 | 0.776 | 0.750 | 0.737 |
| GBDT | 0.763 | 0.763 | 0.789 | 0.737 |
| Adaboost | 0.829 | 0.750 | 0.829 | 0.763 |
| LightGBM | 0.789 | 0.750 | 0.829 | 0.750 |
| ExtraTrees | 0.789 | 0.763 | 0.750 | 0.803 |

*Note*: For the variance method, the threshold was set as 0.02; for RFE, LinearSVC (Linear SVM) was used as the base model, and the number of screening features was 28; for LassoCV, the threshold was "1.1 × Mean".

KNN, K nearest neighbors; SVC, support vector classifier; DT, decision tree; RF, random forest; LR, logistic regression; MLP, multi-layer perceptron; GBDT, gradient boosting decision tree; VT, variance selection; RFE, recursive feature elimination; LassoCV, Lasso cross-validation.

**Table 3** Performance comparison of different models (mean ± std*)

| Model | Accuracy (mean ± std) | Precision (mean ± std) | Recall (mean ± std) | F1-score (mean ± std) | AUC (mean ± std) |
|---|---|---|---|---|---|
| KNN | 0.737 | 0.737 | 0.736 | 0.736 | 0.790 |
| SVC | 0.705 | 0.705 | 0.706 | 0.705 | 0.786 |
| DT | 0.653 | 0.652 | 0.651 | 0.651 | 0.649 |
| RF | 0.728 | 0.729 | 0.728 | 0.728 | 0.804 |
| LR | 0.726 | 0.726 | 0.727 | 0.726 | 0.782 |
| MLP | 0.726 | 0.726 | 0.726 | 0.726 | 0.801 |
| GBDT | 0.742 | 0.743 | 0.743 | 0.742 | 0.819 |
| Adaboost | 0.752 | 0.753 | 0.753 | 0.752 | 0.753 |
| LightGBM | 0.747 | 0.748 | 0.748 | 0.747 | 0.813 |
| ExtraTrees | 0.730 | 0.730 | 0.729 | 0.729 | 0.812 |

*Note*: The metrics of all models were obtained by 100-time iterations.

*All standard deviations are less than 0.01.

KNN, K nearest neighbors; SVC, support vector classifier; DT, decision tree; RF, random forest; LR, logistic regression; MLP, multi-layer perceptron; GBDT, gradient boosting decision tree.

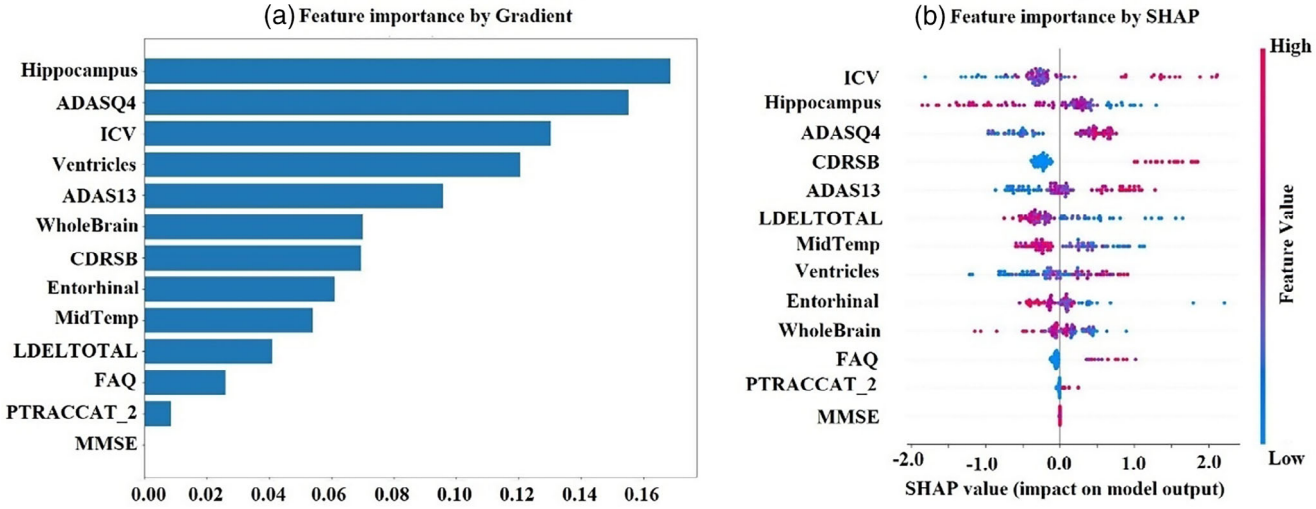


**Figure 2** Feature importance of the gradient boosting decision tree.

ADAS-Cog (ADAS13) (Fig. 2a). For SHAP analysis, the top five predictors were ICV, hippocampus, ADASQ4, Clinical Dementia Rating-Sum of Boxes Score (CDRSB) and ADAS13 (Fig. 2b). No matter what kind of method is used, ICV, hippocampus, ADASQ4, and ADAS13 all showed great importance for the prediction of MCI.

## Discussion

At present, there are few studies with high enough accuracy on the early prediction of MCI. By utilizing a variety of feature selection methods, this study constructed an optimal ML model to predict MCI in the CN population within 3 years, and the AUC obtained by the best model was 0.82. It should not be ignored that the detection of dementia is still challenging in clinical practice.[21,22] A retrospective study on the prediction of MCI progression to AD found that the average accuracy was 75.4%,[23] which means that about 24.6% of AD individuals may be missed by models. If our model can be used before AD, which is equivalent to the multi-stage screening of the same individual, the risk of missing an individual will be reduced, which is undoubtedly significant for healthy people.

We selected 3 years as the prediction span because our aim was to identify the population at high risk of MCI as early as possible. However, previous studies have shown that CN people have a low probability of progression to MCI in the short term.[24] Therefore, multiple observations were used in our study, which has been validated in previous studies;[25–27] for instance, Tavares et al. used two sequential checkup visits (pair of visits) as the unit of analysis to take advantage of repeated metabolic syndrome assessments in the cohort, and successfully predicted the metabolic syndrome with sensitivity, specificity, and AUC reaching 87.8%, 70.2%, and 86.0%, respectively.[26] It should be noted that our enhancement using data from multiple visits of the same individual requires dealing with two issues: data leakage and independence. Data leakage refers to the fact that the training set and the test set contain follow-up information from the same individual. In this study, we ensure that subjects do not overlap by dividing the data at the individual level rather than at the outcome of disease progression; that is, subjects with multiple observations would only be divided into the training set or test set. Sample independence is another concern; that is, the data may be consistent between multiple follow-up visits, which may lead to some individuals not being independent. However, because we only

conducted data enhancement on subjects who progressed to MCI, the process from CN to MCI is constantly changing, and different follow-up visits are in different progression processes, which is not invariable. This practice is similar to the idea of Synthetic Minority Oversampling Technique (SMOTE).[28] In addition, the reason we only performed data enhancement on subjects who progressed to MCI is that we hope to balance positive and negative samples, if data is imbalanced, that is, the proportion of MCI patients far exceeds that of cognitive normal people, then the prediction results of ML model will be biased.[29]

Regardless of which method is used to evaluate feature importance, ICV, hippocampus, and ADASQ4 are the top three important predictors in MCI progression, indicating that these three features are of great importance in predicting MCI. Some studies have confirmed that hippocampal atrophy often occurs in the early stage of AD.[30,31] The hippocampus is the region of interest (ROI) commonly used by MCI researchers,[32] which is consistent with the suggestion in this study that the hippocampus is one of the most important predictors in MCI prediction. ICV is also an important early marker of MCI. Different from hippocampus, different intracranial volumes in healthy people will affect the capacity for cognitive reserve: individuals with larger ICVs may have a greater capacity for "brain damage" than those with smaller ICVs. We hypothesized that ICV can predict MCI, which may reflect how much the volume of a single or multiple intracranial regions can change, such as through changes in the average cortical thickness, hippocampal atrophy, and medial temporal lobe atrophy. These regions are often associated with cognitive decline,[33–35] which indicates that cognitive impairment or decline first begins in the brain, suggesting that further analysis of changes in some intracranial regions and even changes in cerebral blood flow on cognitive impairment is needed in the future to help people determine the cause of MCI.[36] The ADASQ4 rating scale, reflecting the ability to recognize words, shows importance in MCI early prediction, this suggests that the damage of word recognition ability may occurr before MCI. The brain regions corresponding to word recognition are the temporal lobe and the hippocampus, suggesting that changes in the temporal lobe and hippocampus led to the change in word recognition ability. It is worth noting that Mini-Mental State Examination (MMSE) showed relatively low importance in distinguishing MCI from CN, while Clinical Dementia Rating (CDR) was more predictive than MMSE, which is consistent with some findings.[15] For the rest of the features, it's not sure whether they can better reflect the features of early MCI owing to the large difference in the importance evaluation of the two methods (feature importance and SHAP), which needs further research and verification.

Our findings confirm that ML can be used to predict the risk of progression frmo CN to MCI at an early stage. Early intervention can minimize the risk of adverse events such as AD, and the benefits from reducing risk of cognitive impairment have been discussed previously. Unlike other auxiliary diagnostic studies, our study was longitudinal-based and showed good performance (AUC = 0.82). If the model is applied before predicting MCI progression to AD, the number of MCI individuals who progress to AD may be significantly reduced.

Our study had some limitations. First, we failed to introduce variables such as images and cerebrospinal fluid biomarkers, which could have further improved the accuracy of prediction models[16] and helped us to find new MCI markers.[37] Second, we only predicted the risk within 3 years, so a long-term prediction is warranted in the future. Third, the ML models were validated only on the ADNI database, so their generalizability to other populations cannot be guaranteed. Therefore, these models should be externally validated in large cohorts and other regions before application. Given the "black-box" nature of ML, for the time being, ML models should be used with caution and should not be used as a tool to dominate clinical diagnosis.

## Conclusion

In this study, we developed ML algorithms with a small number of neuropsychological tests and MRI-related markers for the accurate prediction of MCI. Our model can be used as a complement to or in combination with models that predict MCI progression to AD. It can facilitate timely intervention for serious diseases such as dementia.

## Acknowledgements

This study was supported by the National Natural Science Foundation of China (grant number: 81973144). We also thank the Alzheimer's Disease Neuroimaging Initiative (ADNI) for providing us with the data.

## Disclosure statement

The authors have no conflicts of interest to declare.

## Data availability statement

The Alzheimer's disease Neuroimaging initiative (ADNI) database was publicly available to researchers, which can be accessed through https://adni.loni.usc.edu/.

## References

1 Reitz C, Mayeux R. Alzheimer disease: epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochem Pharmacol* 2014; **88**: 640–651.
2 Moore PJ, Lyons TJ, Gallacher J. Random forest prediction of Alzheimer's disease using pairwise selection from time series data. *PLoS One* 2019; **14**: e0211558.
3 Marinescu RV, Oxtoby NP, Young AL *et al*. Tadpole challenge: prediction of longitudinal evolution in Alzheimer's disease, 2018.
4 Neugroschl J, Wang S. Alzheimer's disease: diagnosis and treatment across the spectrum of disease severity. *Mt Sinai J Med* 2011; **78**: 596–612.
5 Gauthier S, Reisberg B, Zaudig M *et al*. Mild cognitive impairment. *Lancet* 2006; **367**: 1262–1270.
6 Robinson L, Tang E, Taylor JP. Dementia: timely diagnosis and early intervention. *BMJ* 2015; **350**: h3029.
7 Gonzalez SC, Bamiou D-E, Lewis G *et al*. Feasibility of treating hearing loss to prevent dementia in MCI: results of the face to face and remote interventions in the treating auditory impairment and cognition trial (TACT). *Alzheimers Dement* 2021; **17**: e055148.
8 Lee JH, Kim DY, Kim MG *et al*. Preventing cognitive decline in older adults with mild cognitive impairment using integrated Korean and Western treatments: initial results. *Alzheimers Dement* 2021; **17**: e052545.
9 Battle DE. Diagnostic and statistical manual of mental disorders (DSM). *Codas* 2013; **25**: 191–192.
10 Battista P, Salvatore C, Berlingeri M, Cerasa A, Castiglioni I. Artificial intelligence and neuropsychological measures: the case of Alzheimer's disease. *Neurosci Biobehav Rev* 2020; **114**: 211–228.
11 Molinuevo JL, Ayton S, Batrla R *et al*. Current state of Alzheimer's fluid biomarkers. *Acta Neuropathol* 2018; **136**: 821–853.
12 Pellegrini E, Ballerini L, Hernandez M *et al*. Machine learning of neuroimaging for assisted diagnosis of cognitive impairment and dementia: a systematic review. *Alzheimers Dement (Amst)* 2018; **10**: 519–535.
13 Samper-González J, Burgos N, Fontanella S *et al*. Yet another ADNI machine learning paper? Paving the way towards fully-reproducible

research on classification of Alzheimer's disease. Machine Learning in Medical Imaging: 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings; 2017: Springer.
14 Bringas S, Salomón S, Duque R, Lage C, Montaña JL. Alzheimer's disease stage identification using deep learning models. *J Biomed Inform* 2020; **109**: 103514.
15 Nagaraj S, Duong TQ. Deep learning and risk score classification of mild cognitive impairment and Alzheimer's disease. *J Alzheimers Dis* 2021; **80**: 1079–1090.
16 Grueso S, Viejo-Sobera R. Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review. *Alzheimers Res Ther.* 2021; **13**: 162.
17 Jo T, Nho K, Saykin AJ. Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. *Front Aging Neurosci* 2019; **11**: 220.
18 Sarica A, Cerasa A, Quattrone A. Random forest algorithm for the classification of neuroimaging data in Alzheimer's disease: a systematic review. *Front Aging Neurosci* 2017; **9**: 329.
19 Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2011; **28**: 112–118.
20 Wolpert DH, Macready WG. No free lunch theorems for optimization. *IEEE Trans Evol Comput* 1997; **1**: 67–82.
21 Zhang D, Shen D. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *Neuroimage* 2012; **59**: 895–907.
22 Hinrichs C, Singh V, Xu G, Johnson SC. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population. *Neuroimage* 2011; **55**: 574–589.
23 Langa KM, Levine DA. The diagnosis and management of mild cognitive impairment: a clinical review. *JAMA* 2014; **312**: 2551–2561.
24 Liew TM. Trajectories of subjective cognitive decline, and the risk of mild cognitive impairment and dementia. *Alzheimers Res Ther* 2020; **12**: 135.
25 Feng X, Provenzano FA, Small SA. A deep learning MRI approach outperforms other biomarkers of prodromal Alzheimer's disease. *Alzheimers Res Ther* 2022; **14**: 45.
26 Daniel Tavares L, Manoel A, Henrique Rizzi Donato T *et al*. Prediction of metabolic syndrome: a machine learning approach to help primary prevention. *Diabetes Res Clin Pract* 2022; **191**: 110047.
27 Vodencarevic A, Weingärtner M, Caro JJ *et al*. Prediction of recurrent ischemic stroke using registry data and machine learning methods: the Erlangen stroke registry. *Stroke* 2022; **53**: 2299–2306.
28 Han H, Wang WY, Mao BH. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. *Lect Notes Comput Sci* 2005; **3644**: 878–887.
29 Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw* 2018; **106**: 249–259.
30 Devanarayan P, Devanarayan V, Llano DA. Identification of a simple and novel cut-point based cerebrospinal fluid and MRI signature for predicting Alzheimer's disease progression that reinforces the 2018 NIA-AA research framework. *J Alzheimers Dis* 2019; **68**: 537–550.
31 Madusanka N, Choi HK, So JH, Choi BK, Park HG. One-year follow-up study of hippocampal subfield atrophy in Alzheimer's disease and normal aging. *Curr Med Imaging Rev* 2019; **15**: 699–709.
32 Platero C, Lin L, Tobar MC. Longitudinal neuroimaging hippocampal markers for diagnosing Alzheimer's disease. *Neuroinformatics* 2019; **17**: 43–61.
33 Kehoe EG, McNulty JP, Mullins PG, Bokde AL. Advances in MRI biomarkers for the diagnosis of Alzheimer's disease. *Biomark Med* 2014; **8**: 1151–1169.
34 Dubois B, Feldman HH, Jacova C *et al*. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* 2014; **13**: 614–629.
35 Illán-Gala I, Pegueroles J, Montal V *et al*. Challenges associated with biomarker-based classification systems for Alzheimer's disease. *Alzheimers Dement (Amst)* 2018; **10**: 346–357.
36 Ferreira LK, Rondina JM, Kubo R *et al*. Support vector machine-based classification of neuroimages in Alzheimer's disease: direct comparison of FDG-PET, rCBF-SPECT and MRI data acquired from the same individuals. *Braz J Psychiatry* 2018; **40**: 181–191.
37 Li Y, Haber A, Preuss C *et al*. Transfer learning-trained convolutional neural networks identify novel MRI biomarkers of Alzheimer's disease progression. *Alzheimers Dement (Amst)* 2021; **13**: e12140.

## Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

**Text S1.** Description of neuropsychological and MRI-related variables.
**Text S2.** Machine learning models.
**Table S1.** Excluded variables that were not helpful for model prediction.
**Table S2.** Excluded variables with missing rate (>20%).
**Table S3.** Formula for the machine learning model evaluation.
**Table S4.** Model building platform and corresponding version.
**Table S5.** Comparison of variables between adults with cognitive normal and mild cognitive impairment.
**Table S6.** The number of features selected by different feature selection methods.
**Figure S1.** Hyperparameter tuning of the variance method.
**Figure S2.** Hyperparameter tuning of recursive feature elimination.
**Figure S3.** Hyperparameter tuning of LassoCV.
**Figure S4.** Calibration plots of gradient boosting decision tree (reliability curve).
**Figure S5.** Receiver operating characteristic curve of gradient boosting decision tree.

**How to cite this article:** Wu Y, Wang X, Fang Y. Predicting mild cognitive impairment in older adults: A machine learning analysis of the Alzheimer's Disease Neuroimaging Initiative. Geriatr. Gerontol. Int. 2023;1–6. https://doi.org/10.1111/ggi.14670