Received: 29 September 2023 | Accepted: 25 January 2024

DOI: 10.1111/ene.16235

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

Clinical variables contributing to the identification of biologically defined subgroups within cognitively unimpaired and mild cognitive impairment individuals

Sofia Marcolini[1] | Jaime D. Mondragón[1] | Zeus T. Dominguez-Vega[1] | Peter P. De Deyn[1,2] | Natasha M. Maurits[1] | for the Alzheimer's Disease Neuroimaging Initiative†

[1]University Medical Center Groningen, Department of Neurology, University of Groningen, Groningen, The Netherlands

[2]Laboratory of Neurochemistry and Behavior, Experimental Neurobiology Unit, University of Antwerp, Antwerp, Belgium

Correspondence
Zeus T. Dominguez-Vega, University Medical Center Groningen, Department of Neurology, University of Groningen, PO Box 30001, 9700 RB Groningen, The Netherlands.
Email: z.t.dominguez.vega@umcg.nl; zeus.tlaltecutli@gmail.com

Funding information
Joint Programming Initiative Neurodegenerative Diseases (JPND) HEROES, Grant/Award Number: 733051072; Institutional grant PUSH: Partnership of UMCG – Siemens Healthineers for building the future of Health

Abstract

Background: A lack of consensus exists in linking demographic, behavioral, and cognitive characteristics to biological stages of dementia, defined by the ATN (amyloid, tau, neurodegeneration) classification incorporating amyloid, tau, and neuronal injury biomarkers.

Methods: Using a random forest classifier we investigated whether 27 demographic, behavioral, and cognitive characteristics allowed distinction between ATN-defined groups with the same cognitive profile. This was done separately for three cognitively unimpaired (CU) (112 A-T-N-; 46 A+T+N+/−; 65 A-T+/-N+/−) and three mild cognitive impairment (MCI) (128 A-T-N-; 223 A+T+N+/−; 94 A-T+/-N+/−) subgroups.

Results: Classification-balanced accuracy reached 39% for the CU and 52% for the MCI subgroups. Logical Delayed Recall (explaining 16% of the variance), followed by the Alzheimer's Disease Assessment Scale 13 (14%) and Everyday Cognition Informant (10%), were the most relevant characteristics for classification of the MCI subgroups. Race and ethnicity, marital status, and Everyday Cognition Patient were not relevant (0%).

Conclusions: The demographic, behavioral, and cognitive measures used in our model were not informative in differentiating ATN-defined CU profiles. Measures of delayed memory, general cognition, and activities of daily living were the most informative in differentiating ATN-defined MCI profiles; however, these measures alone were not sufficient to reach high classification performance.

KEYWORDS
Alzheimer's disease, biomarkers, diagnosis, psychological tests, random forest

INTRODUCTION

In clinical settings, dementia diagnosis and treatment monitoring often hinge on cognitive performance assessed through neuropsychological tests [1]. Additionally, behavioral and demographic information is used to support these processes [2]. Despite neuropsychological and behavioral tests being frequently employed, no worldwide consensus is available on their

†Data used in the 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 article. 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.

usage in relation to distinguishing different biological stages of the disease [3].

The current ATN (amyloid, tau, neurodegeneration) research framework proposes to classify Alzheimer's disease (AD) based on a biological biomarker-based system describing three core groups: (1) Aβ burden (amyloid positron emission tomography [PET] or cerebrospinal fluid [CSF] β-amyloid [Aβ42 or Aβ42/Aβ40 ratio]); (2) tau pathology (CSF phosphorylated tau [p-tau] or tau PET); and (3) neuronal injury (CSF total tau [t-tau], anatomic magnetic resonance imaging (MRI), fluorodeoxyglucose (FDG)-PET) [4]. While the recommendation is to use the ATN classification in research contexts, efforts are needed to translate its purpose to clinical practice [5]. Diagnosing AD purely based on biological biomarkers leads to practical issues including generalization, metric performance, accessibility, and thresholds determination [6]. A prior study determined that the ATN classification lacks sufficient consistency for clinical use, reporting insufficient agreement among biomarkers within different pathophysiological categories and noted variations in correlation at different stages of the AD continuum [7]. Another study also showed the ATN system's lack of accuracy for certain non-AD dementias [8]. The International Working Group for the clinical diagnosis of AD has recently reported that the diagnosis of AD should be based on biological biomarker evidence, but also on clinical phenotype [6].

Investigating which demographical, behavioral, and cognitive domains are most relevant in distinguishing biologically defined profiles might be crucial, even to rule out other non-degenerative causes when biomarkers are not available [9]. Previous studies have found that the assessment of cognitive domains, such as attention, processing speed, executive function, memory, and language, can help identify biological disease stages [10, 11]. For instance, two studies have identified delayed recall as the best predictor in distinguishing between amyloid-positive and amyloid-negative mild cognitive impairment individuals [12, 13]. In cognitively normal individuals, subtle impairments in everyday functioning were associated with higher amyloid burden and worse cognition [14]. A major challenge remains to individuate which specific clinical variables are better at distinguishing ATN classified profiles within preclinical and prodromal stages of AD.

In the current study, we aimed to classify ATN-defined biomarker groups with the same cognitive profile (first: cognitively unimpaired [CU] with A-T-N-, A+T+N+/− or A-T+/-N+/− profiles and second: mild cognitive impairment [MCI] with A-T-N-, A+T+N+/− or A-T+/-N+/− profiles) using demographic, behavioral, and cognitive characteristics and a data-driven approach. Additionally, we aimed to identify which demographic, behavioral, and cognitive characteristics are most relevant to identify these biologically defined subgroups within these two cognitive profiles. Our findings will be particularly relevant in contexts where biological biomarker assessment is not available. They will also inform on whether, in different stages of the disease, further biological biomarker assessment has clinical relevance.

METHODS

Data were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu/data-samples/access-data/). The datasets used in this study were downloaded from the ADNI-1, ADNI-2, and ADNI-Grand Opportunity (ADNI-GO) databases between 28 August 2018 and 15 September 2020; further information on the files used to extract cognitive, biological biomarker, and behavioral data used in this study can be found in the Supplementary Material. Participant eligibility criteria for the three ADNI phases are identical and can be found in the ADNI general procedures manual (Alzheimer's Disease Neuroimaging Initiative-I; http://adni.loni.usc.edu/methods/documents/). ADNI was approved by the institutional review boards of all participating centers. Written informed consent was obtained from all patients. For more information see www.adni-info.org.

Participants

Participants included in the ADNI study were between the ages of 55 and 90 years, had completed at least 6 years of education, and were free of any significant neurological disease other than MCI or AD. Exclusion criteria were defined by the ADNI study protocol [15]. Additional inclusion criteria for this study were: availability of biological biomarkers to define the ATN profile (see Section 2.3); no missing demographic, behavioral, or cognitive data; Hachinski score <4; absence of depression as defined by the Geriatric Depression Scale 15 (GDS 15 <9); Clinical Dementia Rating scale (CDR) score <1. Based on the criteria of Petersen (2004) [16] and using the CDR score, participants were considered either CU (CDR = 0) or as having MCI (CDR = 0.5). This resulted in 668 cases with sufficient data for the analyses (see inclusion flow in Figure 1).

Demographic, behavioral, and cognitive variables

Sample demographical characteristics, including age, education, sex, ethnicity, and marital status, were assessed. Behavioral and cognitive measures were assessed by trained staff according to standardized procedures described in the ADNI manual. Details on these variables are available in the ADNI manual (ADNI_GeneralProceduresManual.pdf (usc.edu)); the 27 variables are also described in Table S1 and the reference to each test can be found in Supplementary Material. The functional and behavioral variables included: the Functional Assessment Questionnaire (FAQ); the Everyday Cognition – Participant Self-Report; Everyday Cognition – Study Partner Report; the Neuropsychiatric Inventory; and the Geriatric Depression Scale 15. The cognitive variables included were the following: the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-COG) 13; Montreal Cognitive

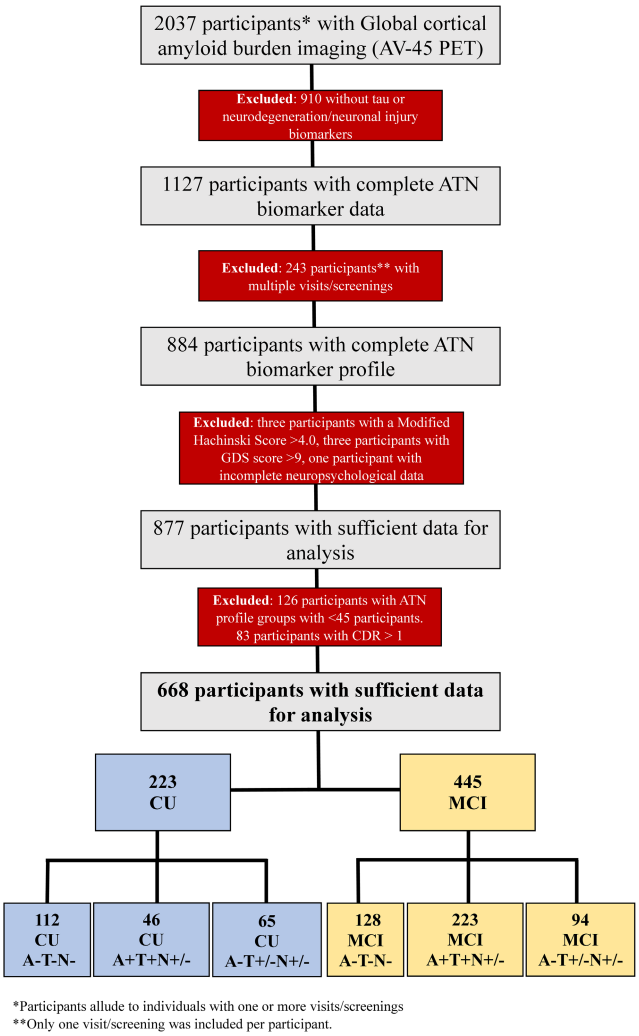

FIGURE 1 Inclusion flow diagram. ATN, amyloid, tau, neurodegeneration classification; CDR, Clinical Dementia Rating scale; CU, cognitively unimpaired; GDS, Geriatric Depression Scale; MCI, mild cognitive impairment; PET, positron emission tomography.

Assessment (MoCA) (adapted from the official MoCA Instruction Guide, Version November 12, 2004 © Z Nasreddine); ADAS-COG 4; Clock Drawing Test, Trail Making Test A and Trail Making Test B; Logical Memory Test Immediate Total; Rey Auditory Verbal Learning Test (RAVLT); Auditory Verbal Learning Test Delayed 30 min (AVDEL30); RAVLT Recognition; RAVLT Forgetting; RAVLT Immediate; Category Fluency Test adapted from the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) administration and scoring procedures for Verbal Fluency and the Boston Naming Test (30 items).

Biological biomarker ATN thresholds and groups classification

For biological biomarker classification, ATN biomarker thresholds have been used as previously validated for the ADNI cohort and summarized in Table S2. Briefly, the florbetapir-PET standardized uptake value ratio (SUVR) ≥1.11 was used to determine Aβ disease burden [17], CSF p-tau levels ≥23 pg/mL were used to determine tau disease burden [18], and FDG-PET SUVR ≤1.21 was used to determine neurodegeneration or neuronal injury [19].

In our study, we classified CU and MCI individuals from three biomarker categories: A-T-N- (normal AD biomarkers); A+T+N+/− (with AD biomarkers); A-T+/-N+/− (non-AD pathological change). In doing this, we aggregated A+T+ individuals (A+T+N- and A+T+N+, both considered part of the Alzheimer's continuum biomarker category) and patients with A- (A-T+N-, A-T+N+, A-T-N+, considered in the non-AD pathological change biomarker category), as shown in Table 2 of Jack et al. [4]. In the CU A+T+N+/− group, 32 were A+T+N- and 14 were A+T+N-. In the MCI A+T+N+/− group, 109 were A+T+N- and 114 were A+T+N-. In the CU A-T+/-N+/− group, 41 were A-T+N-, 17 were A-T-N+, and 7 were A-T+N+. In the MCI A-T+/-N+/− group, 30 were A-T+N-, 45 were A-T-N+, and 13 were A-T+N+.

We then classified individuals also based on their syndromal cognitive stage (CU, MCI, and dementia). From the biomarker categories we decided to only keep participants belonging to groups that represented at least 20% of the sample within the syndromal cognitive stage. Therefore, the A+T-N- (Alzheimer's pathological change), representing only 11.6% of CU and 8% of MCI, and the A+T-N+ (Alzheimer's and concomitant suspected non-Alzheimer's pathological change), representing 2% of CU and 5.6% of MCI, were excluded.

Therefore, after assessing each biomarker and designating a profile based on the cut-off values previously mentioned, each participant was classified into one of six groups derived from the National Institute on Aging-Alzheimer's Association (NIA-AA) research framework [4]: (1) CU with normal biomarkers (CU A-T-N-); (2) CU with AD biomarkers (A+T+N+/−); (3) CU with non-AD pathological change (A-T+/-N+/−); (4) MCI with normal biomarkers (MCI A-T-N-); (5) MCI with AD biomarkers (A+T+N+/−); and (6) MCI with non-AD pathological change (MCI A-T+/-N+/−).

Machine learning approach: data exploration, classification, and ADASYN

The analysis was done separately for the three CU groups and the three MCI groups. For this step, 27 features based on demographic ($n=5$), behavioral ($n=5$), and cognitive ($n=17$) characteristics were used for both analyses (see Table S1) and a script was made in python version 3.7 using scikit-learn version 0.32.

For data preparation, first, a table was created with the information of the participants who fulfilled the inclusion criteria. This table consisted of 27 columns, where each column represents a variable, and 668 rows, where each row represents a participant (sample). Second, 20-fold stratified cross-validation was used to split the data into training and test sets. Here, data from one fold was used as test set and the remaining information was used for training the model. Third, the data were transformed to follow a

TABLE 1 Sample characteristics.

Characteristic	CU A-T-N- (*n* = 112)	CU A+T+N+/− (*n* = 46)	CU A-T+/-N+/− (*n* = 65)	Total (*n* = 223)	*P*-value	MCI A-T-N- (*n* = 128)	MCI A+T+N+/− (*n* = 223)	MCI A-T+/-N+/−(*n* = 94)	Total (*n* = 445)	*P*-value
Age, years (mean ± SD)	72.4 ± 5.7	77 ± 6.1	74.7 ± 8.0	74.0 ± 6.7	<0.001*	69.1 ± 7.2	73.8 ± 6.9	74.2 ± 7.7	72.6 ± 7.5	<0.001*
Education, years (mean ± SD)	16.6 ± 2.5	16.3 ± 2.8	17.3 ± 2.1	16.7 ± 2.5	0.06	16.7 ± 2.5	16.0 ± 2.7	16.0 ± 2.6	16.2 ± 2.6	0.03*
Sex (%)										
Female	51.8	67.4	43.1	52.5	0.04*	46.1	45.7	39.4	44.5	0.53
Ethnicity (%)										
White	82.1	95.7	89.2	87.0	0.01*	86.7	92.8	93.6	91.2	0.11
Black	8.9	2.2	0	4.9		4.7	2.2	0	2.5	
Hispanic	8	2.2	1.5	4.9		4.7	2.2	1.1	2.7	
Asian	0.9	0	4.6	1.8		1.6	0.9	4.3	1.8	
Indian/Alaskan	0	0	1.5	0.4		0	0	1.1	0.2	
Hawaiian/Other	0	0	0	0		0.8	0.4	0	0.4	
More than one	0	0	3.1	0.9		1.6	1.3	0	1.1	
Marital status (%)										
Never married	3.6	4.3	7.7	4.9	0.46	2.3	2.7	2.1	2.5	0.01*
Married	74.1	63	63.1	68.6		71.9	79.8	79.8	77.5	
Divorced	11.6	10.9	10.8	11.2		16.4	6.3	7.4%	9.4	
Widowed	10.7	21.7	18.5	15.2		6.3	11.2	9.6	9.4	
Unknown	0	0	0	0		3.1	0	1.1	1.1	

Abbreviations: ATN, amyloid, tau, neurodegeneration classification; CU, cognitively unimpaired; MCI, mild cognitive impairment; SD, standard deviation.

*P-value < 0.05

normal distribution, using the 'Quantile Transformer' from scikit-learn. Finally, the data were normalized, using the 'Normalizer Transformer' from scikit-learn, which normalizes samples individually to have unit norm.

For classification, a random forest classifier was used. Random forest is an ensemble technique which uses decision trees and averaging to improve predictive performance [20]. Random forest was chosen as a classifier because it provides insight into the most relevant features used for classification, thereby contributing to 'explainable artificial intelligence'. This characteristic of random forest could be used to identify whether demographic, behavioral, or cognitive data are more relevant for the classification task. Three hundred decision trees were employed. The Gini index was used as the separation criterion on each node to decrease node impurity [21]. One of the drawbacks of the current dataset is the imbalance of the classes (see Table 1). To overcome this problem an adaptive synthetic sampling technique (ADASYN) [22] was used, oversampling the minority classes, thereby obtaining a balanced dataset.

Additionally, to investigate the contribution of features to the classification of each of the MCI groups, Shapley values were calculated by using the SHAP (Shapely Addictive exPlanations) method [23]. SHAP values aim to explain the prediction of a class (or a group) by computing the contribution of each feature to the prediction; more information on their usage in the context of AD can be found elsewhere [24]. Positive SHAP values indicate that for that class the feature is improving accuracy, while negative values instead indicate that the feature impacts the accuracy in distinguishing that class negatively.

Finally, to evaluate the classifier performance and taking into account the nature of the imbalance dataset, a group of evaluation metrics are reported: balance accuracy (arithmetic mean of sensitivity and specificity); sensitivity (true positive rate, i.e., proportion of participants with positive test who were correctly identified); specificity (true negative rate, i.e., proportion of a negative test result predicted by the model); precision (positive predictive value, i.e., proportion of participants with a positive test of all positive predicted cases); and the f1 score (harmonic mean of precision and recall) [25, 26]. There is no single metric to evaluate performance with an imbalanced dataset, therefore these several evaluation metrics are reported.

This process was repeated 100 times to obtain the overall performance of the classifier, and tables and figures were created using descriptive statistics of this information.

Statistical analysis

Statistical analysis was performed using SPSS 28 (SPSS Inc., Chicago, IL, USA). Sample characteristics (age, education, sex, ethnicity, marital status) were compared between the three CU groups and then between the three MCI groups using one-way analyses of variance (ANOVA) for continuous normally distributed variables (means and standard deviations displayed) and chi-square tests for categorical variables (percentages displayed). All analyses considered a significance level of $\alpha = 0.05$.

RESULTS

A total of 668 participants were categorized into six ATN groups: 112 CU with A-T-N-; 46 CU with A+T+N+/−; 65 CU with A-T+/-N+/−; 128 MCI with A-T-N-; 223 MCI with A+T+N+/−; 94 MCI with A-T+/-N+/−. Sample characteristics are displayed in Table 1. CU individuals with A-T-N- profile were significantly younger than CU individuals with A+T+N+/− ($p < 0.001$) or A-T+/-N+/− profile ($p = 0.02$) and included more Black ($p = 0.02$) and Hispanic participants ($p = 0.02$). The CU individuals with A+T+N+/− profile comprised fewer females ($p = 0.01$) and White participants ($p = 0.02$) compared to the other two CU groups. MCI individuals with A-T-N- profile were significantly younger than MCI with A+T+N+/− or A-T+/-N+/− profile (both $p < 0.001$) and were more educated than MCI with A+T+N+/− ($p = 0.01$) or A-T+/-N+/− profile ($p = 0.04$); they also included more divorced participants and more participants with unknown marital status compared to the other two groups ($p < 0.001$).

TABLE 2 Classification performance metrics for the cognitively unimpaired and mild cognitive impairment groups.

Actual	Balanced accuracy	Sensitivity	Specificity	Precision	f1-score
Model 1					
CU A-T-N-	0.54 (0.01)	0.46 (0.02)	0.63 (0.02)	0.55 (0.02)	0.50 (0.02)
CU A+T+N+/−	0.59 (0.01)	0.47 (0.03)	0.72 (0.02)	0.30 (0.02)	0.37 (0.02)
CU A-T+/-N+/−	0.50 (0.02)	0.26 (0.03)	0.74 (0.02)	0.29 (0.02)	0.27 (0.02)
Model 2					
MCI A-T-N-	0.71 (0.01)	0.69 (0.01)	0.72 (0.01)	0.50 (0.01)	0.58 (0.01)
MCIA+T+N+/−	0.69 (0.01)	0.62 (0.01)	0.76 (0.01)	0.72 (0.01)	0.67 (0.01)
MCI A-T+/-N+/−	0.54 (0.01)	0.25 (0.02)	0.84 (0.01)	0.29 (0.02)	0.27 (0.02)

Abbreviations: ATN, amyloid, tau, neurodegeneration classification; CU, cognitively unimpaired; MCI, mild cognitive impairment.

TABLE 3 Confusion matrix of the averaged 100 iterations for the two models.

	Prediction		
	CU A-T-N-	CU A+T+N+/−	CU A-T+N+/−
Actual			
CU A-T-N-	45.95 (2.4)	24.11 (2.0)	29.95 (2.4)
CU A+T+N+/−	35.30 (2.6)	46.87 (3.5)	17.82 (2.8)
CU A-T+/-N+/−	38.86 (2.1)	35.55 (2.1)	25.58 (2.6)
	Prediction		
	MCI A-T-N-	MCI A+T+N+/−	MCI A-T+N+/−
Actual			
MCI A-T-N-	68.66 (1.2)	14.98 (0.9)	16.36 (1.2)
MCI A+T+N+/−	22.39 (0.8)	61.70 (0.8)	15.92 (1.0)
MCI A-T+/-N+/−	40.00 (1.9)	35.36 (1.6)	24.64 (1.9)

Abbreviations: ATN, amyloid, tau, neurodegeneration classification; CU, cognitively unimpaired; MCI, mild cognitive impairment.

Classification of CU groups

The classification of the three CU groups, over 100 iterations, led to a balanced accuracy of 39%. Table 2 displays the classification performance metrics and Table 3 the confusion matrix. Classification-balanced accuracy on group level was 54.0% (1.0%) (mean and standard deviation) for the CU A-T-N- group, 59.0% (1.0%) for the CU A+T+N+/− group, and 50.0% (2.0%) for the CU A-T+/-N+/− group. Considering the low accuracy achieved, feature importance was not evaluated for these three CU groups.

Classification of MCI groups

The classification of the three MCI groups, using leave one patient out over 100 iterations, led to a balanced accuracy of 52%. Table 2 displays the classification performance metrics and Table 3 the confusion matrix. Classification-balanced accuracy on group level was 71.0% (1.0%) for the MCI A-T-N- group, 69.0% (1.0%) for the MCI A+T+N+/− group, and 54.0% (1.0%) for the MCI A-T+/-N+/− group.

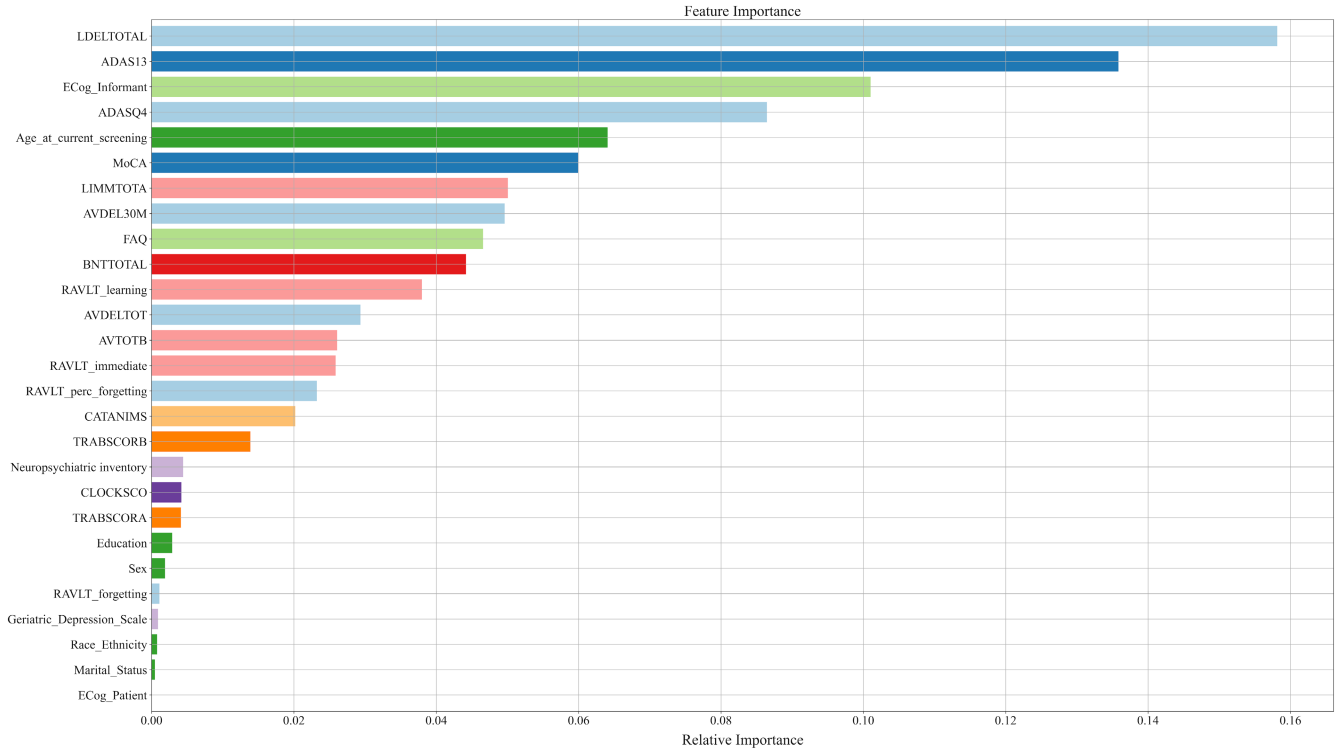

FIGURE 2 Feature importance for the mild cognitive impairment (MCI) groups. ADAS, Alzheimer's Disease Assessment Scale (Cognitive Subscale); AVTOTB, Auditory Verbal Learning Test Total B; AVDEL30, Auditory Verbal Learning Test Delayed 30 min; AVDELTOT, Auditory Verbal Learning Test Delayed Total (Recognition); BNTTOTAL, Boston Naming Test Total; CATANIMS, Category Fluency Test Animals; CLOCKSCO, Clock Drawing Test; ECog, Everyday Cognition; FAQ, Functional Assessment Questionnaire; LDELTOTAL, Logical Memory Delayed Total; LIMMTOTA, Logical Memory Test Immediate Total A; MoCA, Montreal Cognitive Assessment; NPI, Neuropsychiatric Inventory; RAVLT, Rey Auditory Verbal Learning Test; TRABSCORA, Trail Making Test A; TRABSCORB, Trail Making Test B. Lines with the same color represent variables belonging to the same cognitive domain.

Feature importance MCI groups

Random forest allows us to obtain the most relevant features used in the classification of patients in one of the three MCI groups. In Figure 2 the features are presented from the most to the least relevant, displaying the median of relative feature importance over 100 iterations (relative feature importance is also reported in Table S3). Logical Delayed Recall was the most relevant feature (explaining 16% of the variance), followed by the Alzheimer's Disease Assessment Scale-Cognitive Subscale 13 (14%), Everyday Cognition Informant (10%), and Alzheimer's Disease Assessment Scale 4 (9%). Race and ethnicity, marital status, and Everyday Cognition Patient were not relevant (0%).

SHAP values, showing the contribution of each feature in distinguishing each MCI group, are displayed in Figure 3 and are reported in Table S4. Logical Delayed Recall, MoCA, and Auditory Verbal Learning Test Delayed 30 min contribute most to the prediction of the MCI A-T-N- group, while Alzheimer's Disease Assessment Scale-Cognitive Subscale 13, Everyday Cognition Informant, and Alzheimer's Disease Assessment Scale 4 contribute most to the prediction of the MCI A+T+N+/− group.

DISCUSSION

This study aimed to classify CU and MCI individuals with A-T-N-, A+T+N+/−, or A−T+/-N+/− profiles, based on demographic, behavioral, and cognitive characteristics. Several previous studies have predicted cognitive profiles based on biological information [27–31]. Our approach instead is to use clinical information to classify ATN-defined cognitive profiles to bridge the biological and clinical classification of cognitive impairment and to inform clinicians in case of biomarker assessment inaccessibility. With our set of variables, the best classification was achieved for the three MCI groups and especially for the MCI A+T+N+/−, for which 71% balanced accuracy was reached. Lower balanced accuracy was achieved for the other two MCI groups and all CU groups. Two measures of delayed memory, one of general cognition, and one of activities of daily living, when answers were filled in by the informant, were the most important variables to classify the three MCI groups, although balanced accuracy was still not high enough for clinical utility. Marital status, sex, and activities of daily living, when filled in by the patient, were the least relevant.

From our study we find that behavioral and cognitive assessments from a single domain are not enough to classify cognitive

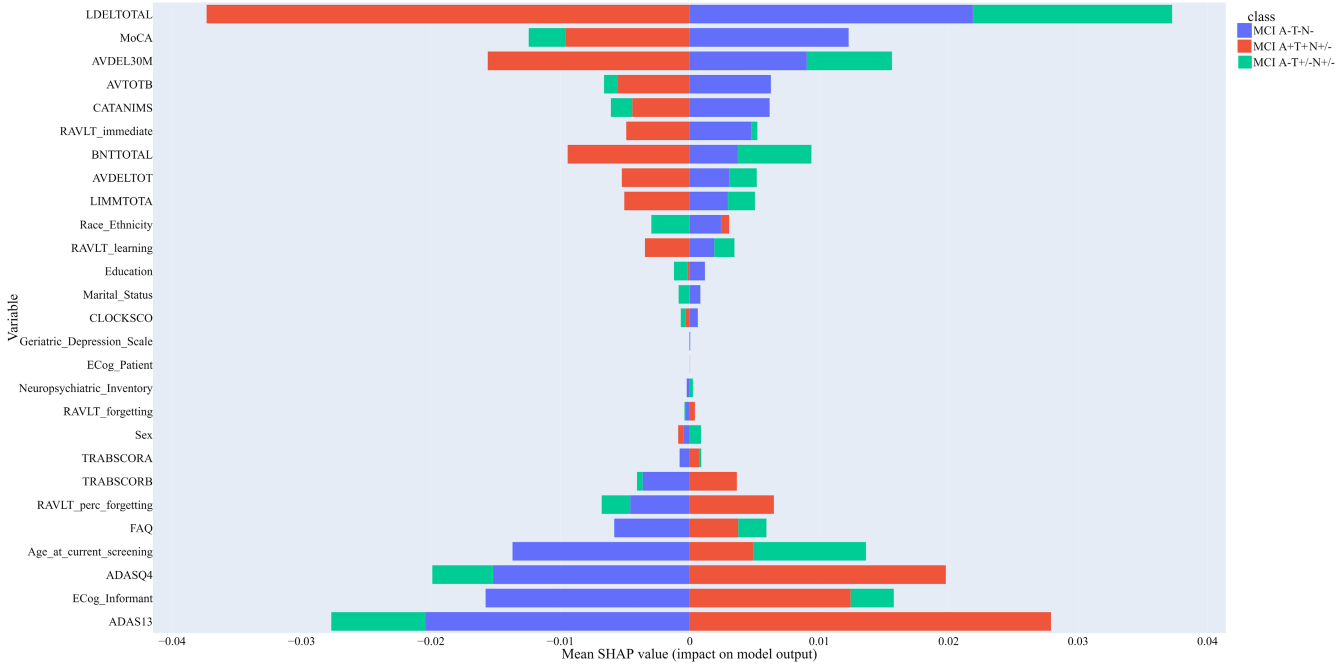

FIGURE 3 Mean SHAP values for the classification of mild cognitive impairment (MCI) groups. Mean SHapley Additive exPlanations (SHAP) values are used to investigate the contribution of all features to classifying each MCI group. Negative values indicate that the feature is contributing negatively to the prediction of classifying that group, while positive values indicate that the variable is improving the prediction for classifying that group. It should be noted that – in line with our classification results – only few features improve the prediction for classifying the MCI A-T+N+/− group (green). ADAS, Alzheimer's Disease Assessment Scale (Cognitive Subscale); AVDEL30, Auditory Verbal Learning Test Delayed 30 min; AVDELTOT, Auditory Verbal Learning Test Delayed Total (Recognition); AVTOTB, Auditory Verbal Learning Test Total B; BNTTOTAL, Boston Naming Test Total; CATANIMS, Category Fluency Test Animals; CLOCKSCO, Clock Drawing Test; ECog, Everyday Cognition; FAQ, Functional Assessment Questionnaire; LDELTOTAL, Logical Memory Delayed Total; LIMMTOTA, Logical Memory Test Immediate Total A; MoCA, Montreal Cognitive Assessment; NPI, Neuropsychiatric Inventory; RAVLT, Rey Auditory Verbal Learning Test; TRABSCORA, Trail Making Test A; TRABSCORB, Trail Making Test B.

profiles with different biological profiles; instead, the classifier requires information from several domains to achieve higher balanced accuracy. Our results are in line with a substantial amount of literature pointing at delayed memory and daily functioning as important domains for the diagnosis of MCI due to AD [12, 32]. In a previous study, a naming task together with wordlist forgetting and recognition, phonemic fluency, and cognitive estimations were the best measures to distinguish between individuals from the AD continuum and those with non-AD change [33]. Changes in everyday functioning are also seen throughout the AD continuum [34]. In a study comparing MCI groups with different pathologies, those with non-AD pathology were shown to have worse performance in almost all cognitive domains compared to MCI with negative biomarkers, but better performance than MCI with positive amyloid and neurodegeneration biomarkers [35]. Also, delayed recall was identified as the best predictor in distinguishing between amyloid-positive and amyloid-negative MCI individuals [12, 13]. A recent study has found that a measure of semantic verbal fluency, and not semantic-phonemic discrepancy, was predictive of abnormal CSF phosphorylated tau 181 levels [36]. Importantly, the measures used in the study of Aiello et al. [36] are specific measures of semantic fluency which are not normally incorporated in routine neuropsychological batteries. Therefore, to reach higher balanced accuracy in predicting an individual's biological status, more domain-specific cognitive measures might be needed.

Notably, among the three MCI groups, MCI individuals with positive AD biomarkers are best distinguished. This might suggest that while cognitively equally classified as MCI, in MCI with positive AD biomarkers a more advanced disease status is also reflected in clinical changes compared to the other two biomarker groups. These results also suggest that biomarker assessment in individuals with MCI might add clinical relevance. In case of inaccessibility to biomarker evaluation, the cognitive tests identified in our study as most relevant for MCI biomarker profile classification can provide further information on whether AD biomarkers are present. This is important since for individuals with positive AD biomarkers the risk of progression to dementia increases [37, 38]. It should be kept in mind, however, that the clinical variables used are not sufficient to reach high balanced accuracy in distinguishing MCI individuals with different biomarker profiles and therefore the distinction between groups solely with these measures is not yet feasible.

The MCI group with amyloid negativity, but with tau and or neurodegenerative markers positivity, is the group with the worst classification performance. Looking at the SHAP values, we see that a measure of general cognition (ADAS13) is the one contributing negatively to the classification of that group. We could speculate that non-AD pathological changes might lead to a more complex cognitive profile or that the variables used were originally developed to detect impairment in individuals presenting with Alzheimer's pathological changes. This aligns with findings from a previous study, which highlighted a lack of accuracy in classifying non-AD dementias [8]. The complexity also arises from the diverse clinical and pathological entities underpinning these profiles, including primary age-related tauopathy, argyrophilic grain disease, and Lewy body disease [39].

Whether CU individuals with positive biomarkers should be described as preclinical AD [40] or merely as at risk for cognitive impairment [6] is still currently debated. A previous study on the ADNI dataset found that AD biomarkers were found in 36% of a cognitively normal group [41], and another study established that concentrations of total and phosphorylated tau and amyloid-beta 42 in CSF were highly related to future development of AD in individuals with MCI [42]. Our results show that based on demographic, behavioral, and cognitive information, it is not possible to distinguish between CU individuals with different biomarker profiles. This suggests that demographic, behavioral, and cognitive measures are not sufficient to distinguish changes in biological status when cognitive impairment is not present yet and, therefore, that the definition of biomarker profiles for CU individuals might have limited added value in clinical contexts. It should also be noted that these tests were developed several decades ago and were aimed at distinguishing CU individuals from individuals with dementia, while the characterization of biomarker profiles is recent. Reduced thresholds in cognitive testing made it possible to increase sensitivity for subtle cognitive changes, but reduced specificity as several other disorders could be the cause of such changes such as psychiatric or metabolic disorders [6]. In CU individuals, subtle impairments in everyday functioning have been related to higher amyloid burden and worse cognition [14]; additional studies involving CU individuals with different biomarker profiles are thus still required.

The precision in detecting individuals with MCI increases when a full neuropsychological examination is added, and for instance when adding the ADAS-COG to the Mini-Mental State Examination (MMSE) [43]. The need to use tests assessing specific cognitive domains rather than global cognitive composites has been extensively discussed [44, 45]. Yet, for treatment monitoring, most studies only use general tests of cognition [46, 47]. Future efforts are needed to establish which tests are more sensitive in measuring treatment efficacy in participants with different biological statuses. In the case of positive effects of a specific treatment aimed at ameliorating cognitive complaints, knowing which cognitive domains benefited from treatment is also of paramount importance [45].

A strength of our study is that our included variables are almost all among the tests recently recommended to diagnose mild neurocognitive impairment in Europe [48]. The work done on the harmonization of neuropsychological tests usage could be extended by providing recommendations also based on the biological stage of disease. In future work we want to incorporate biological information into our models (i.e., APOE status, structural and perfusion MRI) currently not incorporated in the ATN framework, to see if that added information would improve classification-balanced accuracy, especially for the CU groups. Vascular and inflammatory factors are currently not included in the ATN framework, but several expert panels have recommended their inclusion. Additionally, a recent study has shown that measures of hippocampal volume

and white matter hyperintensities, neither considered in the ATN classification, were important predictors in different stages of the AD continuum [49].

A limitation of the current study is that the current machine learning model first needs to be externally validated. Additionally, biomarker threshold determination is still under debate. In our study we used thresholds determined on the ADNI population; future studies will therefore need to use thresholds specific to their population. In our study, some of the ATN profiles were aggregated in their corresponding biomarker categories described in the 2018 NIA-AA Research Framework [4]. This was necessitated by the limited number of ATN profiles within the three non-AD pathological change categories and in two groups of the Alzheimer's continuum. While this aggregation may have introduced greater heterogeneity to the groups, potentially influencing our findings, the groups that have been aggregated still belonged to the same biomarker category group described in the NIA-AA Research Framework. Thus, we believe that this choice has not impacted our results. Also, there was insufficient data to include the A+T-N- and the A+T-N+ profiles. It is important to acknowledge that ATN profiles are not static, and their longitudinal changes remain incompletely understood. Consequently, future studies exploring the dynamic nature of ATN profiles are important for a comprehensive understanding. Moreover, we do not specifically detect cognitive and behavioral test score thresholds to distinguish one subgroup from the other, which could help to give more specific recommendations on tests usage, once balanced accuracy increases. Which specific scores of the individuated tests should be considered informative about biomarker profile remains a question.

Overall, the demographic, behavioral, and cognitive variables inserted in our model were not able to distinguish three CU ATN subgroups. Measures of delayed memory, general cognition, and activities of daily living were informative in classifying ATN-defined MCI profiles, although not yet sufficient to reach high accuracy. Future efforts are therefore needed to obtain new behavioral and cognitive variables which are more sensitive in distinguishing profiles with different biological biomarker statuses.

AUTHOR CONTRIBUTIONS

Sofia Marcolini: Conceptualization; writing – original draft; project administration; data curation; investigation; visualization. **Jaime D. Mondragón:** Conceptualization; writing – review and editing; project administration; investigation; data curation. **Zeus T. Dominguez-Vega:** Formal analysis; conceptualization; writing – review and editing; visualization; methodology; software; data curation; validation. **Peter P. De Deyn:** Funding acquisition; writing – review and editing; resources; supervision; conceptualization. **Natasha M. Maurits:** Resources; supervision; conceptualization; writing – review and editing; methodology; validation.

FUNDING INFORMATION

S.M. is supported by an institutional grant PUSH: Partnership of UMCG - Siemens Healthineers for building the future of Health and by the Joint Programming Initiative Neurodegenerative Diseases (JPND) HEROES project (No.: 733051072).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from ADNI. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from https://adni.loni.usc.edu/ with the permission of ADNI.

ORCID

Sofia Marcolini https://orcid.org/0000-0002-1856-9348

REFERENCES

1. Pasternak E, Smith G. *Cognitive and Neuropsychological Examination of the Elderly*. Vol 167. 1st ed Elsevier B.V.; 2019. doi:10.1016/B978-0-12-804766-8.00006-6
2. Marshall GA, Amariglio RE, Sperling RA, Rentz DM. Activities of daily living: where do they fit in the diagnosis of Alzheimer's disease? *Neurodegener Dis Manag*. 2012;2(5):483-491. doi:10.2217/nmt.12.55
3. Maruta C, Guerreiro M, de Mendonça A, Hort J, Scheltens P. The use of neuropsychological tests across Europe: the need for a consensus in the use of assessment tools for dementia. *Eur J Neurol*. 2011;18(2):279-285. doi:10.1111/j.1468-1331.2010.03134.x
4. Jack CRJ, Bennett DA, Blennow K, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562. doi:10.1016/j.jalz.2018.02.018
5. Tromp K, Smedinga M, Richard E, Perry M, Schermer MHN. Views on early diagnosis of Alzheimer's disease among Dutch physicians: a qualitative interview study. *J Alzheimers Dis*. 2021;79(2):917-927. doi:10.3233/JAD-200884
6. Dubois B, Villain N, Frisoni GB, et al. Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol*. 2021;20(6):484-496. doi:10.1016/S1474-4422(21)00066-1
7. Illán-Gala I, Pegueroles J, Montal V, et al. Challenges associated with biomarker-based classification systems for Alzheimer's disease. *Alzheimers Dement*. 2018;10:346-357. doi:10.1016/j.dadm.2018.03.004
8. Baldeiras I, Silva-Spínola A, Lima M, et al. Alzheimer's disease diagnosis based on the amyloid, tau, and neurodegeneration scheme (ATN) in a real-life multicenter cohort of general neurological centers. *J Alzheimers Dis*. 2022;90(1):419-432. doi:10.3233/JAD-220587
9. Villain N, Dubois B. Alzheimer's disease including focal presentations. *Semin Neurol*. 2019;39(2):213-226. doi:10.1055/s-0039-1681041
10. Weintraub S, Salmon D, Mercaldo N, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord*. 2009;23(2):91-101. doi:10.1097/WAD.0b013e318191c7dd
11. di Pucchio A, Vanacore N, Marzolini F, et al. Use of neuropsychological tests for the diagnosis of dementia: a survey of Italian memory clinics. *BMJ Open*. 2018;8(3):1-8. doi:10.1136/bmjopen-2017-017847
12. Clarens MF, Crivelli L, Calandri I, et al. Neuropsychological profile of Alzheimer's disease based on amyloid biomarker findings results from a South American cohort. *Appl Neuropsychol Adult*. 2022;29(3):345-350. doi:10.1080/23279095.2020.1756816
13. Alves L, Cardoso S, Silva D, et al. Neuropsychological profile of amyloid-positive versus amyloid-negative amnestic mild cognitive impairment. *J Neuropsychol*. 2021;15(S1):41-52. doi:10.1111/jnp.12218

14. Marshall GA, Sikkes SAM, Amariglio RE, et al. Instrumental activities of daily living, amyloid, and cognition in cognitively normal older adults screening for the A4 study. *Alzheimers Dement (Amst)*. 2020;12(1):e12118. doi:10.1002/dad2.12118
15. Mueller SG, Weiner MW, Thal LJ, et al. The Alzheimer's disease neuroimaging initiative. *Neuroimaging Clin N Am*. 2005;15(4):869-77, xi-xii. doi:10.1016/j.nic.2005.09.008
16. Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004;256(3):183-194. doi:10.1111/j.1365-2796.2004.01388.x
17. Landau SM, Breault C, Joshi AD, et al. Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. *J Nucl Med*. 2013;54(1):70-77. doi:10.2967/jnumed.112.109009
18. Shaw LM, Vanderstichele H, Knapik-Czajka M, et al. Qualification of the analytical and clinical performance of CSF biomarker analyses in ADNI. *Acta Neuropathol*. 2011;121(5):597-609. doi:10.1007/s00401-011-0808-0
19. Dowling NM, Johnson SC, Gleason CE, Jagust WJ. The mediational effects of FDG hypometabolism on the association between cerebrospinal fluid biomarkers and neurocognitive function. *NeuroImage*. 2015;105:357-368. doi:10.1016/j.neuroimage.2014.10.050
20. Breiman L. Random forests. *Mach Learn*. 2001;45:1. doi:10.1023/A:1010933404324
21. Dominguez-Vega ZT, Dubber D, Elting JWJ, Sival DA, Maurits NM. Instrumented classification of patients with early onset ataxia or developmental coordination disorder and healthy control children combining information from three upper limb SARA tests. *Eur J Paediatr Neurol*. 2021;34:74-83. doi:10.1016/j.ejpn.2021.07.009
22. He H, Bai Y, Garcia EA, Li S. ADASYN: adaptive synthetic sampling approach for imbalanced learning. 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, 2008, pp. 1322–1328. doi:10.1109/IJCNN.2008.4633969
23. Lundberg S, Lundberg SM, Allen PG, Lee SI. A Unified Approach to Interpreting Model Predictions ChromNet View Project Shapley Additive ExPlanations (SHAP) View Project A Unified Approach to Interpreting Model Predictions. 2017. Accessed May 2023. https://github.com/slundberg/shap.
24. Bogdanovic B, Eftimov T, Simjanoska M. In-depth insights into Alzheimer's disease by using explainable machine learning approach. *Sci Rep*. 2022;12(1):6508. doi:10.1038/s41598-022-10202-2
25. Hicks SA, Strümke I, Thambawita V, et al. On evaluation metrics for medical applications of artificial intelligence. *Sci Rep*. 2022;12(1):1-9. doi:10.1038/s41598-022-09954-8
26. Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. *Lect Notes Comput Sci*. 2005;3408(2014):345-359. doi:10.1007/978-3-540-31865-1_25
27. Jitsuishi T, Yamaguchi A. Searching for optimal machine learning model to classify mild cognitive impairment (MCI) subtypes using multimodal MRI data. *Sci Rep*. 2022;12(1):1-14. doi:10.1038/s41598-022-08231-y
28. Syaifullah AH, Shiino A, Kitahara H, Ito R, Ishida M, Tanigaki K. Machine learning for diagnosis of AD and prediction of MCI progression from brain MRI using brain anatomical analysis using diffeomorphic deformation. *Front Neurol*. 2021;11:1-13. doi:10.3389/fneur.2020.576029
29. Wang Y, Xu C, Park JH, et al. Diagnosis and prognosis of Alzheimer's disease using brain morphometry and white matter connectomes. *Neuroimage Clin*. 2019;23:101859. doi:10.1016/j.nicl.2019.101859
30. Farina MP, Saenz J, Crimmins EM. Does adding MRI and CSF-based biomarkers improve cognitive status classification based on cognitive performance questionnaires? *PLoS One*. 2023;18(5):e0285220. doi:10.1371/journal.pone.0285220
31. 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(1):162. doi:10.1186/s13195-021-00900-w
32. Weissberger GH, Strong J, Stefanidis KB, Summers MJ, Bondi MW, Stricker NH. Diagnostic accuracy of memory measures in Alzheimer's dementia and mild cognitive impairment: a systematic review and meta-analysis. *Neuropsychol Rev*. 2017;27(4):354-388. doi:10.1007/s11065-017-9360-6
33. Florean I, Penolazzi B, Menichelli A, et al. Using the ATN system as a guide for the neuropsychological assessment of Alzheimer's disease. *J Clin Exp Neuropsychol*. 2021;43(9):926-943. doi:10.1080/13803395.2022.2036327
34. Dubbelman MA, Jutten RJ, Tomaszewski Farias SE, et al. Decline in cognitively complex everyday activities accelerates along the Alzheimer's disease continuum. *Alzheimers Res Ther*. 2020;12(1):1-11. doi:10.1186/s13195-020-00706-2
35. Wisse LEM, Butala N, Das SR, et al. Suspected non-AD pathology in mild cognitive impairment. *Neurobiol Aging*. 2015;36(12):3152-3162. doi:10.1016/j.neurobiolaging.2015.08.029
36. Aiello EN, Verde F, Solca F, et al. Lower semantic fluency scores and a phonemic-over-semantic advantage predict abnormal CSF P-tau181 levels in Aβ+ patients within the Alzheimer's disease clinical spectrum. *Neurol Sci*. 2023;44(6):1979-1985. doi:10.1007/s10072-023-06643-w
37. Rostamzadeh A, Bohr L, Wagner M, Baethge C, Jessen F. Progression of subjective cognitive decline to MCI or dementia in relation to biomarkers for Alzheimer disease: a meta-analysis. *Neurology*. 2022;99(17):E1866-E1874. doi:10.1212/WNL.0000000000201072
38. Parnetti L, Chipi E, Salvadori N, D'Andrea K, Eusebi P. Prevalence and risk of progression of preclinical Alzheimer's disease stages: a systematic review and meta-analysis. *Alzheimers Res Ther*. 2019;11(1):1-13. doi:10.1186/s13195-018-0459-7
39. Wisse LEM, de Flores R, Xie L, et al. Pathological drivers of neurodegeneration in suspected non-Alzheimer's disease pathophysiology. *Alzheimers Res Ther*. 2021;13(1):100. doi:10.1186/s13195-021-00835-2
40. Ossenkoppele R, Pichet Binette A, Groot C, et al. Amyloid and tau PET-positive cognitively unimpaired individuals are at high risk for future cognitive decline. *Nat Med*. 2022;28(11):2381-2387. doi:10.1038/s41591-022-02049-x
41. De Meyer G, Shapiro F, Vanderstichele H, et al. Diagnosis-independent Alzheimer disease biomarker signature in cognitively normal elderly people. *Arch Neurol*. 2010;67(8):949-956. doi:10.1001/archneurol.2010.179
42. Hansson O, Zetterberg H, Buchhave P, Londos E, Blennow K, Minthon L. Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *Lancet Neurol*. 2006;5(3):228-234. doi:10.1016/S1474-4422(06)70355-6
43. Wouters H, Appels B, Van Der Flier WM, et al. Improving the accuracy and precision of cognitive testing in mild dementia. *J Int Neuropsychol Soc*. 2012;18(2):314-322. doi:10.1017/S1355617711001755
44. Schneider LS, Goldberg TE. Composite cognitive and functional measures for early stage Alzheimer's disease trials. *Alzheimers Dement (Amst)*. 2020;12(1):1-9. doi:10.1002/dad2.12017
45. Harrison JE. Commentary: composite cognitive and functional measures for early stage Alzheimer's disease trials. *Alzheimers Dement (Amst)*. 2020;12(1):10-12. doi:10.1002/dad2.12009
46. van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med*. 2023;388(1):9-21. doi:10.1056/NEJMoa2212948
47. Rahman A, Hossen MA, Chowdhury MFI, et al. Aducanumab for the treatment of Alzheimer's disease: a systematic review. *Psychogeriatrics*. 2023;23:512-522. doi:10.1111/psyg.12944
48. Boccardi M, Monsch AU, Ferrari C, et al. Harmonizing neuropsychological assessment for mild neurocognitive disorders in Europe. *Alzheimers Dement*. 2022;18(1):29-42. doi:10.1002/alz.12365

49. Prosser L, MacDougall A, Sudre CH, et al. Predicting cognitive decline in nondemented elders using baseline metrics of AD pathologies, cerebrovascular disease, and neurodegeneration. *Neurology*. 2022;100(8):e834-e845. doi:10.1212/wnl.0000000000201572

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Marcolini S, Mondragón JD, Dominguez-Vega ZT, De Deyn PP, Maurits NM, . Clinical variables contributing to the identification of biologically defined subgroups within cognitively unimpaired and mild cognitive impairment individuals. *Eur J Neurol*. 2024;00:e16235. doi:10.1111/ene.16235