Cognitive Systems Research 78 (2023) 87–95

Contents lists available at ScienceDirect

Cognitive Systems Research

journal homepage: www.elsevier.com/locate/cogsys

Machine learning approaches to mild cognitive impairment detection based on structural MRI data and morphometric features

M.O. Zubrikhina [a,*], O.V. Abramova [c,f], V.E. Yarkin [c], V.L. Ushakov [b,c,d], A.G. Ochneva [c,f], A.V. Bernstein [a], E.V. Burnaev [a], D.S. Andreyuk [c], V.B. Savilov [c], M.V. Kurmishev [c], T.C. Syunyakov [c], O.A. Karpenko [c], A.V. Andryushchenko [c], G.P. Kostyuk [c,e], M.G. Sharaev [a,c]

[a] Skolkovo Institute of Science and Technology, Bol'shoy Bul'var, 30 bld. 1, Moscow 121205, Russia
[b] Institute for Advanced Study of the Brain, Lomonosov Moscow State University, Ulitsa Kolmogorova, 1, Moscow 119991, Russia
[c] Mental-health clinic No 1 named after N.A. Alexeev, Zagorodnoye Shosse, 2, Moscow 115191, Russia
[d] National Research Nuclear University MEPhI, Kashira Hwy, 31, Moscow 115409, Russia
[e] Federal State Budgetary Educational Institution of Higher Education "Moscow State University of Food Production", Volokolamskoye Shosse, 11, Moscow 125080, Russia
[f] Department of Basic and Applied Neurobiology, V. Serbsky Federal Medical Research Centre of Psychiatry and Narcology, Kropotkinsky per. 23, 119034 Moscow, Russia

ARTICLE INFO

Keywords:
Machine learning
Mild cognitive impairment
MRI
Morphometry
FreeSurfer

ABSTRACT

Mild cognitive impairment (MCI) is an important public health problem that has enormous consequences for patients, their families, the health care system, and the economy. MCI is highly likely to progress to dementia. Early premedical intervention for MCI could prevent or delay dementia's onset, which in turn would improve patients' quality of life. Thus, the goal of our research is to develop an MCI detection system using MRI-based measures. The dataset for the study consists of 62 healthy volunteers and 145 patients with MCI obtained from the «Memory Clinic» (Moscow, Russia). The morphometric data were extracted from T1-weighted structural MR images. We propose a machine learning method, which achieved an accuracy of 73%, and extract key features in brain structures that can distinguish the healthy control (HC) group from the MCI group. This study confirms the previously identified changes in the structural regions of the brain in MCI patients. Compared to the healthy control group, in the MCI group, volume increases in the Lateral Ventricle, Inferior Lateral Ventricle, and Cerebrospinal Fluid as well as volume reductions in the Amygdala, Hippocampus, and Accumbens have been found.

1. Introduction

Neurocognitive disorders, especially severe dementias, have enormous consequences for patients, their families, the health care system and the economy. According to the WHO, 55 million people (8.1 % of women and 5.4 % of men over the age of 65) suffer from dementia. This number is expected to rise to 78 million by 2030 and to 139 million by 2050 ("https://www.who.int/news-room/fact-sheets/detail/dementia," 2002). The development of dementia primarily affects memory and other cognitive functions, as well as a person's ability to perform everyday tasks ("https://www.who.int/news-room/fact-sheets/detail/dementia," 2002). Neurodegenerative disorders are known to begin many years before symptoms appear, but may be preceded by mild cognitive impairment (MCI). MCI is a borderline condition between cognitive normality and dementia and is considered the prodromal phase of dementia (Breton et al., 2019). The criteria for MCI are cognitive decline, including memory, which, however, does not meet the criteria for dementia, but retains near-normal functional activity and daily independence (Breton et al., 2019; Roberts and Knopman, 2013). The prevalence of MCI in adults over 60 is approximately 6.7 % to 25.2 %. It increases with age, is more common in individuals with lower educational attainment, and is more common in men (Jongsiriyanyong and Limpawattana, 2018).

MCI is an important public health problem because of the increased risk of progression to dementia (Roberts and Knopman, 2013). Approximately half of patients diagnosed with MCI develop dementia

* Corresponding author.
E-mail address: Maria.Zubrikhina@skoltech.ru (M.O. Zubrikhina).

https://doi.org/10.1016/j.cogsys.2022.12.005
Received 15 August 2022; Received in revised form 5 December 2022; Accepted 8 December 2022
Available online 13 December 2022

within 3 years, and between 6 % and 15 % of patients progress to dementia per year from the time of MCI diagnosis (Breton et al., 2019; Roberts and Knopman, 2013). On average, about 20 % of patients with MCI will improve over time. Despite this, subjects who return to normal cognitive functioning may not be cognitively normal. These subjects are more probable to progress to MCI or dementia later in life than subjects who never developed MCI (Jongsiriyanyong and Limpawattana, 2018; Roberts and Knopman, 2013).

Early premedical intervention for MCI could prevent dementia or delay its onset, which in turn would improve patients' quality of life (Ball et al., 2020). The most effective interventions that could slow or stop the progression of neurocognitive decline would be those initiated during the mild cognitive decline phase or before the onset of any symptoms at all. Early diagnosis of cognitive decline has its advantages: the patient's future problems can be foreseen to some extent, and measures can be taken in advance. Moreover, there are some modifiable predictors of transition to dementia (e.g., untreated diabetes) that can be influenced (Breton et al., 2019; Morozova et al., 2022). Despite this, there are no ways to proactively diagnose the onset of cognitive decline. Potentially, one way to make an early diagnosis of cognitive decline could be to identify various biological markers that could identify the onset of disease with a high degree of probability. Keep in mind that cognitive impairment also occurs in some mental disorders, so early diagnosis is relevant for this.

One of the most promising biomarkers that can be applied for early diagnosis are neuroimaging biomarkers. Data from neuroimaging techniques, including magnetic resonance imaging (MRI), are important tools both for understanding the pathology associated with cognitive and clinical symptoms and for differential diagnosis. Neurodegenerative diseases affect the brain and result in changes in brain structure, function, and molecular composition that can be captured by neuroimaging techniques. In addition, cerebrospinal fluid tests and neuroimaging are the key laboratory tests used to assess these changes in the brain in living patients. That said, neuroimaging is the most commonly used medical test for diagnosis in conjunction with a thorough clinical examination and neurocognitive assessment (Risacher and Saykin, 2019).

Neuroimaging indices are used, among other things, to predict progression from cognitively normal to MCI. For example, structural MRI has been shown to observe increases in ventricular volume and reductions in global brain volume and medial temporal lobe volume 10 years before cognitive decline, providing prognostic information (Dickerson et al., 2012; Tondelli et al., 2012). The hippocampal CA1 and subiculum volume are associated with progression to MCI (Apostolova et al., 2010). Finally, atrophy in the prefrontal cortex and rate of atrophy of the basal forebrain have been identified as potential predictors of progression from normal to MCI (Burgmans et al., 2009; Grothe et al., 2013). The data presented point to the potential applicability of neuroimaging biomarkers for predicting cognitive decline. However, approaches to the analysis of neuroimaging data require additional attention. For this purpose, the use of modern approaches to neuroimaging data analysis, one of which is machine learning, appears promising.

Machine learning (ML) is a pattern recognition technique that can be applied, among other things, to medical images. A system of machine learning algorithms calculates image features that are thought to be important for prediction or diagnosis, then determines the best combination of these image features to classify the image or calculate some metric for that area of the image. Automated detection and diagnosis performed using ML algorithms can help clinicians interpret imaging results and reduce interpretation time (Erickson et al., 2017). Including ML in personalized medicine is a growing area of interest. The power of ML has the potential to personalize the approach on many levels: diagnosis, prognosis, and treatment. This applies to many areas of medicine where medical data imaging is used, including in various brain pathologies. Previously, we showed promising results of ML application to neurological (Sharaev et al., 2018b) and psychiatric (Sharaev et al., 2018a) disease diagnostics.

In our study, we collected morphometric MRI data and cognitive testing data, in patients with MCI and in healthy volunteers of the same age. In the next step, we developed a model that accounts for clinically important parameters in the obtained morphometric neuroimaging data and can potentially differentiate a healthy individual from a patient with MCI. Unlike others works, where relevant features are selected manually at the very beginning of the analysis, we have done it automatically using feature selection methods and machine learning methods. Despite the growth in the number of works using deep learning, we haven't used deep learning approaches due to data limitations.

In addition, we wanted to build a more interpretable model for doctors.

1.1. Related work

Currently, there is a lot of works dedicated to building algorithms capable of predicting Alzheimer's disease, or MCI. In the majority of its authors use the ADNI open dataset ("https://adni.loni.usc.edu/," 2022). For instance, in (Ledig et al., 2018; Schmitter et al., 2015), the authors used an approach based on morphometric features extracted from structural T1w MRI data. In the article (Schmitter et al., 2015), the features of interest consist of a few brain structure volumes (hippocampi, lobes, ventricles) were manually selected. The balanced accuracy obtained by the SVM method is 73 %. In (Ledig et al., 2018), 86 features were selected for analysis, which included volumes of cortical and non-cortical structures. The best accuracy of 74 % was also archived by SVM. Unlike these two works, where relevant features were selected manually at the very beginning of the analysis.

In addition to articles where predictions are based on morphometric features, there are a number of studies that use deep learning methods. For example, in one paper (Basaia et al., 2019) 3D Convolutional Neural Network (3D CNN) was used, and in another paper (Parisot et al., 2018) Graph Convolutional Neural Network (GCN) with Spectral graph convolutions is used. In (Basaia et al., 2019), the quality of a convolutional neural network is separately evaluated on HC and stable MCI (s-MCI, 76.1 % accuracy) and HC and MCI, that will convert to AD (p-MCI, 87.1 accuracy). When GCN was applied to predict s-MCI and p-MCI, the accuracy on 10-fold cross-validation was reached at 78.8 % (Parisot et al., 2018).

There are several works that used multimodal approaches to classify MCI and HC (Suk and Shen, 2013; Venugopalan et al., 2021). A multi-kernel SVM was trained on the concatenated latent representations obtained from stacked autoencoders for each of the modalities. The classification accuracy was improved from 0.740 ± 0.021 for only MRI data to 0.850 ± 0.012 for a multimodal approach (MRI, PET, CSF) (Suk and Shen, 2013). In (Venugopalan et al., 2021), authors tried various data integration techniques (feature level combination, intermediate feature level combination, decision level combination). They used 3D CNN for MRI data and Deep Stacked Denoising Autoencoders for electronic health record and single-nucleotide polymorphism. The multimodal approach also showed the best accuracy (0.8 ± 0.002).

2. Materials and methods

2.1. Dataset

The dataset was collected at the Mental-health clinic No. 1 named after N.A. Alexeev, which consists of 207 participants. Each participant has T1w MRI data, Mini-Mental State Examination (MMSE) and Montreal Cognitive Assessment (MoCA) scores. T2-weighted images from Toshiba 1.5 T MR scanner were acquired with the following parameters: TR = 12 ms, TE = 5 ms, 200 sagittal slices, FOV 256 mm, FA 180, TI = 300 ms, voxel size 1x1x1 mm3.

Participants from the control group (HC), a total of 62 (20 males), aged 68.9 ± 5.3, were selected from a pool of healthy volunteers. The

MCI patients, a total of 145 (16 males), aged 73.9 ± 7.3 were selected from «Memory Clinic» (Moscow, Russia). Healthy controls matched with the patients in age were recruited from the volunteers who came for periodic health examinations in the outpatient clinic N° 121 (Moscow, Russia). Informed consent was obtained from all participants. The referral center ethics committee (Mental-health Clinic No. 1 named after N.A. Alekseev of Moscow Healthcare Department) approved (Protocol N° 2 28.10.2020) the patient recruitment and collection protocols were in accordance with ethical standards according to the World Medical Association Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects.

All diagnoses were assigned at a daily consensus conference among neurologists, neuropsychologists and psychiatrists. The International Classification of Diseases (ICD-10) was used to determine the diagnosis status. Participants received standardized neurobehavioral exams, including neurological examination, neuropsychological testing. All subjects were checked, including subjective memory complaints, normal general cognition, preserved performance of daily living activities, and a measurable impairment in one or more cognitive functions. Patients over the age of 65 who had been diagnosed with mild cognitive impairment (F06.7) were eligible. Psychiatric illness, positive family histories (first-degree relatives) of psychiatric illness, substance abuse, and a heavy comorbid severe somatic or neurological disorder were all exclusion criteria. Information about vascular risk factors (including hypertension, hypercholesterolemia, diabetes mellitus, a history of stroke, heart disease) and a family history of dementia was provided by the patients or their caregivers.

The MoCA is widely used to assess the cognitive functions of patients with various diseases. Results are scored on a scale between 0 and 30, with a score of 26 or higher considered normal. The MMSE is a short 30-item questionnaire widely used to assess and screen for cognitive impairment, including dementia, also used to assess the dynamics of cognitive function against the background of ongoing therapy.

2.2. Preprocessing

We applied FreeSurfer 6.0.1 software package ("https://surfer.nmr.mgh.harvard.edu/," 2022)with the thickness of the cortical surface limit of 10 mm for extracting morphometric features from T1w MRI data. We used recon-all with all steps, such as motion correction, intensity normalization, registration, skull striping, white matter segmentation, spherical mapping, cortical parcellation, etc. We obtained 906 non-null features after all steps, including the voxel count, number of vertices, surface area of various brain structures, and thicknesses and curvatures of cortical areas. Since many features are highly correlated, for example the number of vertices, surface area and volume of the brain structure, we removed features describing the volume of the same brain structure in different variations. For example, delete volume in mm3 and surface area but keep the number of voxels.

2.3. ML pipeline

Due to the small dataset, we decided to apply a classical Machine Learning (ML) approach rather than deep learning. We conducted a common ML grid search procedure to examine various ML models with all types of extracted features. After the grid search, we built a final model using the top-performing ML algorithm. The pipeline included cross-validation and feature selection procedures. Due to an imbalanced dataset, for estimation of model performance, we used 10-fold cross-validation with the F1-metric.

We examined various ML models: Logistic Regression, K-Nearest Neighbor, Gradient Boosting Classifier, Random Forest, C-Support Vector (scikit-learn Python library, version 0.20.1) and XGBClassifier (xgboost Python library, version). Before fitting data in the model, all data was standardized by removing the mean and scaling to unit variance. The models with maximized F1-macro were chosen from Randomized CV Search and standard Grid Search (scikit-learn Python library, version 0.20.1) due to imbalanced data. For the feature selection procedure, we used SelectKBest (scikit-learn Python library, version 0.20.1) with a chi-squared test.

Table 1
Comparison of the ML model performance to classify MCI vs Control on 10-fold cross-validation.

ML model	Best params	ROC-AUC	Accuracy	F1-macro
XGBoost	'select_k': 150, 'subsample': 0.3, 'scale_pos_weight': 0.6, 'n_estimators': 276, 'max_depth': 2, 'lambda': 0.1,'gamma': 0.01, 'eta': 0.2	0.644 +- 0.149	0.7 +- 0.114	0.644 +- 0.12
KNeighbors Classifier	'n_neighbors': 10, 'weights': 'uniform', 'select_k': 500	0.593 +- 0.09	0.637 +- 0.081	0.557+- 0.070
GradientBoosting Classifier	'select_k': 150, 'n_estimators': 215, 'min_samples_split': 2, 'min_samples_leaf': 1, 'max_features': 'sqrt', 'max_depth': 20, 'learning_rate': 0.3	0.687 +- 0.095	0.67 +- 0.097	0.638 +- 0.09
SVC	'C': 1, 'class_weight': 'balanced', 'kernel': 'rbf', 'select_k': 500	**0.735 +- 0.076**	**0.73 +- 0.059**	**0.683 +- 0.063**
RandomForest Classifier	'select_k': 550, 'n_estimators': 338, 'min_samples_split': 3, 'min_samples_leaf': 1, 'max_features': 'auto', 'model_max_depth': 2, 'model_class_weight': 'balanced_subsample', 'model_bootstrap': True	0.666 +- 0.118	0.71 +- 0.105	0.653 +- 0.11
Logistic Regression	'model_C': 0.01, 'model_class_weight': None, 'model_penalty': 'l2', 'model_solver': 'liblinear', 'select_k': 600	0.711 +- 0.097	0.681 +- 0.075	0.648 +- 0.072

2.4. Feature research

In order to understand which features of the brain systems are related to MCI, we explored the predictive power of the most important features from the best models and also clustered highly correlated features. The clusters show that changes in some structures are strongly related to each other, and the model looks not only at one feature of the brain structure but at the set of features related to the corresponding system of structures. In order to understand that the models don't overfit and that features are stable on different folds, we have run the best models 50 times with 10-fold cross-validation and added the top 10 features obtained by the model at each step to the list of important features. As a result, we got the list of the important features with the number of times this feature has been selected (i.e., the selection frequency). After that, we sorted the features according to this score and left the top 15 of them. Next, we used an agglomerative clustering method on the selected features from the two best models described in 2.3 paragraph, using Pearson correlation as a distance metric with an average linkage function (e.g., distance between clusters was calculated as an average of distances between pairs of cluster elements). The number of clusters was selected subjectively based on the dendrogram and inter-cluster distance.

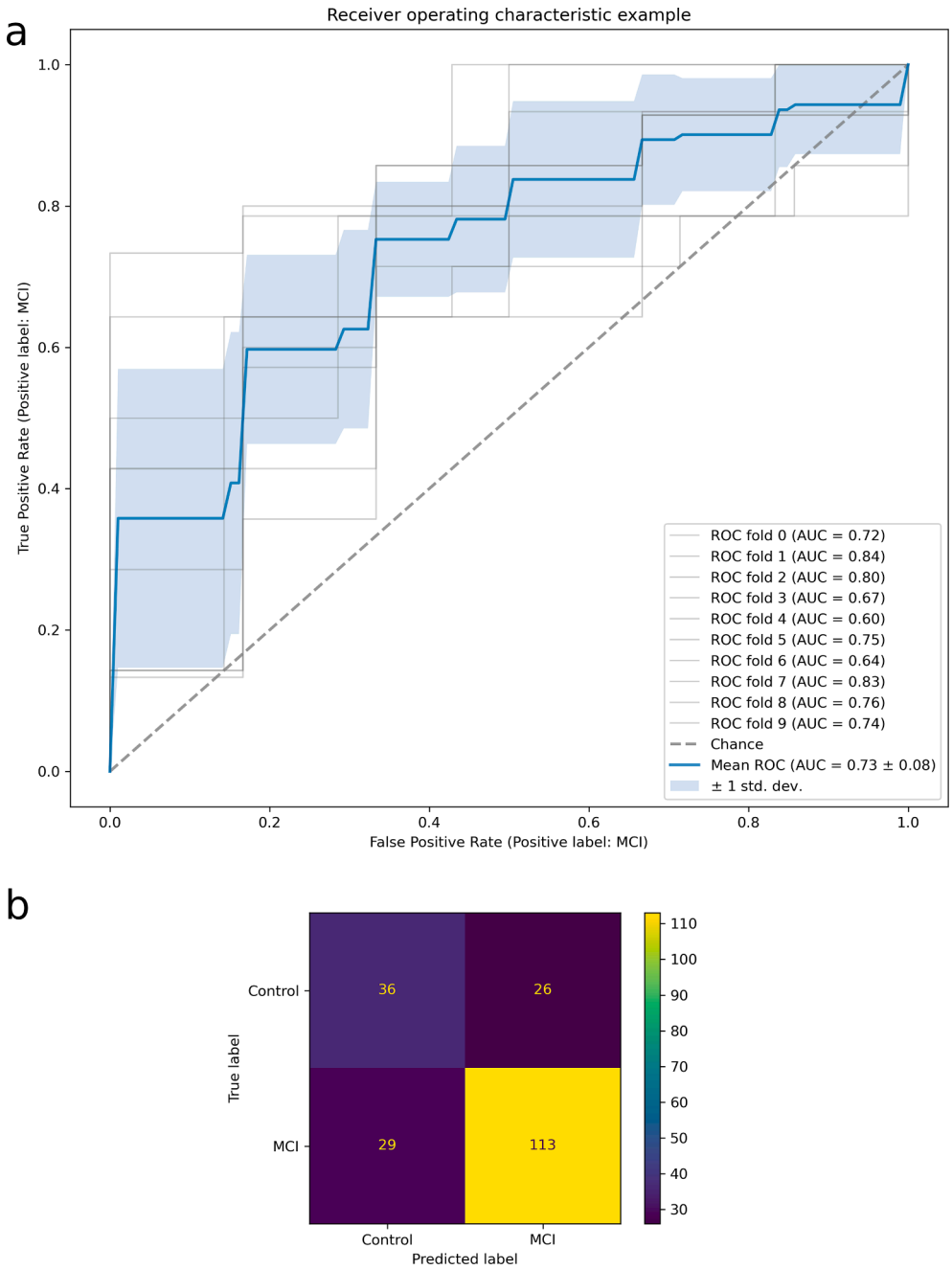

Fig. 1. (a) The best svc model's roc curves after 10-fold cross-validation for mci/control classification. (b) The confusion matrix of the svc model for mci/control classification is calculated as the sum of fn, tp, fp, tn on a test set for all 10 folds.

3. Results and discussion

3.1. Model performance

As we can see from Table 1, the best and most stable results were obtained by SVC (0.68 F1, 0.735 ROC -AUC), Random Forest Classifier (0.653 F1, 0.666 ROC-AUC) showed moderate results, and KNN was the worst one (0.593 ROC-AUC). The ROC curves and the confusion matrix of the SVC model are shown in Fig. 1. The model shows better recall (0.796) and precision (0.816) for the MCI group compared to recall (0.583) and precision (0.583) for the Control group. For a better ratio of recall and precision between classes, weight balancing was used, which allows adjusting weights inversely proportional to class frequencies in the input data. In this task, for evaluating models, the F1-macro metric turned out to be more effective, which allows to just evaluate the quality ratio for two classes: MCI and Control.

3.2. Feature importance

The importance of the features was evaluated for several models using the approach described in paragraph 1.4. Since building the best SVC model with a non-linear kernel feature extraction is not a trivial task, we used Random Forest Classifier and Logistic Regression models for it. The most important features from the morphometric data for both models are presented in Table 2. In Fig. 2, we show the correlation matrix with the combined features of both models, where it is clearly seen that some features are cross-correlated and form clusters. For example, the number of voxels of left inferior lateral ventricles, left lateral ventricle and CSF form the cluster, as well as the left and right Amygdala and right hippocampus. This means that selected features

Table 2
The most informative and stable features for MCI/ Control classification. The significant features with p-value < 0.05 are highlighted in bold.

Model 1 (Logistic Regression)					**Model 2 (Random Forest)**				
Feature name	count	p-value	test	Changes in brain structures (MCI)	Feature name	count	p-value	test	Changes in brain structures (MCI)
Caudal anterior cingulate (ThickStd_lh)	500	**<0,001**	*U* test	**increase**	**Left Lateral Ventricle (NVoxels)**	494	**<0,001**	*U* test	**increase**
Left Cerebellum (CortexnormMax)	476	**<0,001**	*U* test	**decrease**	**Left Amygdala (NVoxels)**	490	**<0,001**	*U* test	**decrease**
Frontal pole (SurfArea_rh)	445	**<0,01**	T-test	**decrease**	**CSF (NVoxels)**	471	**<0,001**	*U* test	**increase**
Brain Stem (normMax)	363	0,08	*U* test	decrease	**Left-Inf-Lat-Vent (NVoxels)**	461	**<0,001**	*U* test	**increase**
5th-Ventricle (NVoxels)	358	0,26	*U* test	decrease	**isthmus cingulate (GrayVol_lh)**	454	**<0,01**	*U* test	**decrease**
Temporal pole (NumVert_lh)	236	**<0,01**	T-test	**decrease**	**Precentral (GrayVol_lh)**	446	**<0,01**	*U* test	**decrease**
Brain stem (normStdDev)	215	**<0,01**	*U* test	**decrease**	**Postcentral (GrayVol_rh)**	358	**<0,01**	*U* test	**decrease**
Caudal anterior cingulate (NumVert_lh)	196	0,36	*U* test	increase	**Caudal middle frontal (GrayVol_lh)**	354	**<0,01**	*U* test	**decrease**
Left Pallidum (normStdDev)	193	0,13	*U* test	increase	**Right Hippocampus (NVoxels)**	327	**<0,001**	*U* test	**decrease**
Left Lateral Ventricle (normStdDev)	192	**<0,001**	T-test	**decrease**	**Right Accumbens area (NVoxels)**	242	**<0,01**	T-test	**decrease**
CC_Anterior (normRange)	179	0,15	T-test	decrease	**Cuneus (GrayVol_lh)**	125	**0,02**	*U* test	**decrease**
CC Mid Posterior (normMin)	172	**<0,01**	*U* test	**increase**	**Right Amygdala (NVoxels)**	117	**<0,01**	*U* test	**decrease**
entorhinal (GrayVol_rh)	163	0,45	*U* test	increase	Precentral (GrayVol_rh)	70	0,11	*U* test	decrease
precuneus (MeanCurv_rh)	137	0,09	*U* test	decrease	**Frontal pole (GrayVol_rh)**	70	**0,03**	*U* test	**decrease**
medial orbito frontal (MeanCurv_lh)	128	0,36	*U* test	decrease	isthmus cingulate (NumVert_lh)	50	0,18	*U* test	decrease

from different folds describe the same system (or a set of structures) of the brain that could be affected by disease (thus the features are correlated) and confirms that the model is reliable.

Both models select the Left Lateral Ventricle and Frontal Pole as important brain structures for MCI prediction, which is also supported by a number of other studies (Ledig et al., 2018; Nestor et al., 2008). We can conclude that machine learning models select reliable and relevant features for MCI vs Control classification because Amygdala, Right Hippocampus, Left-Inf-Lat-Vent, Brain Stem, Left Cerebellum, and Caudal Anterior Cingulate influence prediction to a large extent, as shown in the article (Bidelman et al., 2017; Nemoto et al., 2006; Song et al., 2021).

3.3. Post-hoc T-test

After selecting the important features, post-hoc analysis was carried out to show the differences in brain regions between the MCI group and the HC group. The results of this analysis are shown in Table 2. The Mann-Whitney *U* test and *T*-test were used to test the hypotheses about the means of the selected features from the brain regions. To test the applicability of the *T*-test, the Shapiro test was used to test distributions for normality, and the Levene's test was used to test the equality of variances. If these assumptions were not met, the Mann-Whitney *U* test was used.

For all important features, their distributions are shown on Fig. 3 to highlight the differences between the control group and MCI. As we can see, Lateral Ventricle, Inferior Lateral Ventricle and Cerebrospinal fluid volumes are increased in the MCI group, while the number of voxels of Amygdala, Right Hippocampus and Right Accumbens are reduced. A gray volume in the left hemisphere of the precentral, caudal middle frontal, isthmus of the cingulate gyrus, and a gray volume in the right hemisphere of the postcentral gyrus are also reduced. The changes in these regions are also confirmed in some studies (Karas et al., 2008; Machulda et al., 2020; Reinvang et al., 2012).

Since in the top-15 features the model selects several structures from only one hemisphere (i.e., Right Hippocampus, Right Accumbens, Left Lateral Ventricle, Left-Inf-Lat-Vent) we checked whether this structure changes in the other hemisphere in the same way. We conducted post-hoc analysis for the symmetric structures and concluded that in the MCI group, selected structures in both hemispheres change symmetrically (i.e., increase or decrease). (See, Table 3.).

4. Conclusion

The application of ML in clinical practice holds tremendous promise. ML reveals complex patterns in multivariate data, which can then be used to make automated clinical predictions in new datasets (Dimitriadis et al., 2018; for the Alzheimer's Disease Neuroimaging Initiative et al., 2020). ML is a promising method for specific prediction, including MCI and dementia (for the Alzheimer's Disease Neuroimaging Initiative et al., 2020). The task of predicting the transition from cognitively normal to MCI is inherently difficult to accomplish because it is very difficult to distinguish cognitive decline associated with MCI symptoms from cognitive decline with stable cognitive performance at baseline (Yue et al., 2021). From a clinical perspective, it is important to build a computational model to predict the possible progression to MCI in a healthy individual. Developing improved models may contribute to the understanding of early manifestations of cognitive decline, which may improve clinical care (for the Alzheimer's Disease Neuroimaging Initiative et al., 2020; Suk et al., 2016). Therefore, our study aimed to create a model that could further identify the risk of MCI in cognitively healthy older adults before the manifestation of symptoms.

We used the ML method to analyze morphometric neuroimaging data obtained with MRI to determine the most important neuroimaging morphometric parameters and to differentiate between cognitively healthy individuals and patients with MCI of the same age. We identified

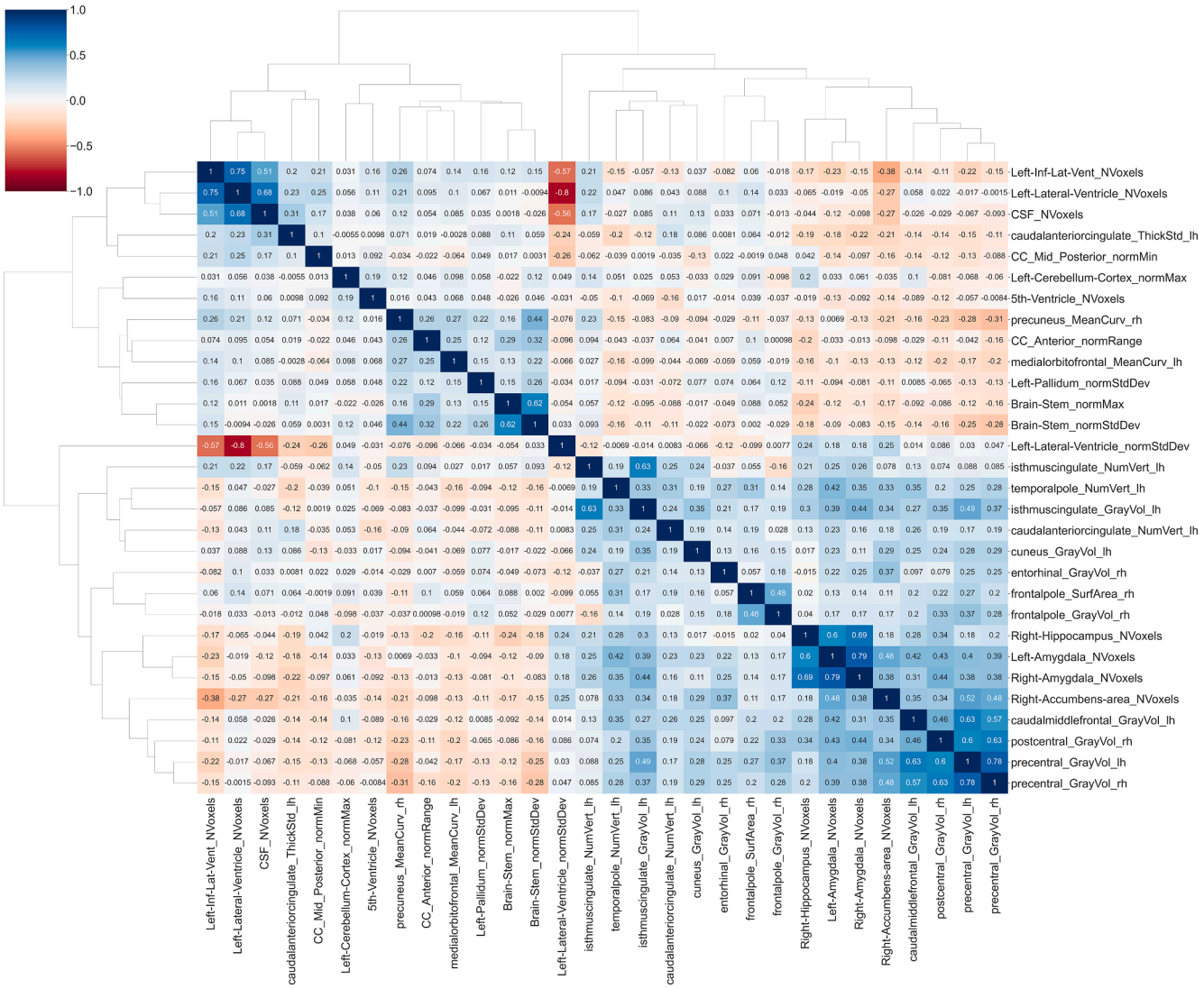

Fig. 2. The correlation matrix with dendrogram between important morphometric features. The color bar indicates the value of the correlation coefficient r (ranging from − 1 to 1).

the most important morphometric parameters and constructed the model to determine cognitive status that could potentially be applied to diagnose cognitive decline in the elderly.

Our model identified several of the most important morphometric indicators for the transition from normal to MCI. These indicators are related to a number of parameters of brain structures such as the lateral ventricles, amygdala, grey matter, hippocampus, and nucleus accumbens. The findings are consistent with many neuroimaging studies that have shown similar potential predictors of progression from cognitive normal to MCI, such as increases in ventricular volume and reductions in global brain volume (Dickerson et al., 2012; Tondelli et al., 2012), decreased hippocampal and subiculum volume, (Apostolova et al., 2010), atrophy in the prefrontal cortex and rate of atrophy of the basal forebrain (Burgmans et al., 2009; Grothe et al., 2013) and alterations in the parietal lobe, including white matter hyperintensity scores (Brickman et al., 2012). This demonstrates on the one hand the validity of our model construction, and on the other hand, provides further evidence that morphometric indicators that can be captured by structural MRI can indeed be predictors of cognitive decline if the data are analyzed correctly. Neuroimaging has indeed made it possible to quantify pathological alterations in the brain. However, integrating neuroimaging with ML methods has the potential to improve our understanding of structural and functional modifications in the pathological brain, so building improved models is an important task (Dimitriadis et al., 2018).

Research and modeling efforts are underway to successfully predict pathology. There are several studies that predict the transition from cognitive normality to MCI (Albert et al., 2018; for the Alzheimer's Disease Neuroimaging Initiative et al., 2020; Mofrad et al., 2021; Ruijuan Yun et al., 2013; Yue et al., 2021). (Albert et al., 2018) used imaging biomarkers associated with the hippocampus and entorhinal cortex, obtaining a sensitivity of 64 % in predicting conversion to MCI. (Yue et al., 2021) obtained an accuracy/sensitivity of 63 %/42 % when predicting decline to MCI using MRI-derived features alone, improving their results to 70 % accuracy and 63 % sensitivity when including multidomain features. (Mofrad et al., 2021) derived a new flexible predictive framework based on mixed-effects and ML models for longitudinal MRI that is robust to noise and different examination intervals and numbers of scans. The framework is based on measurements of hippocampal and lateral ventricular volumes in individual subjects over time. The prediction accuracy for the transition from cognitively normal to MCI is 73 % (for the Alzheimer's Disease Neuroimaging Initiative et al., 2020; Mofrad et al., 2021) attempted to examine whether baseline moderate behavioral impairment status, used to quantify neuropsychiatric symptoms, along with features of brain morphology, is predictive of a subsequent diagnosis. Healthy individuals and MCI patients were studied 40 months apart. Only two characteristics out of more than 200

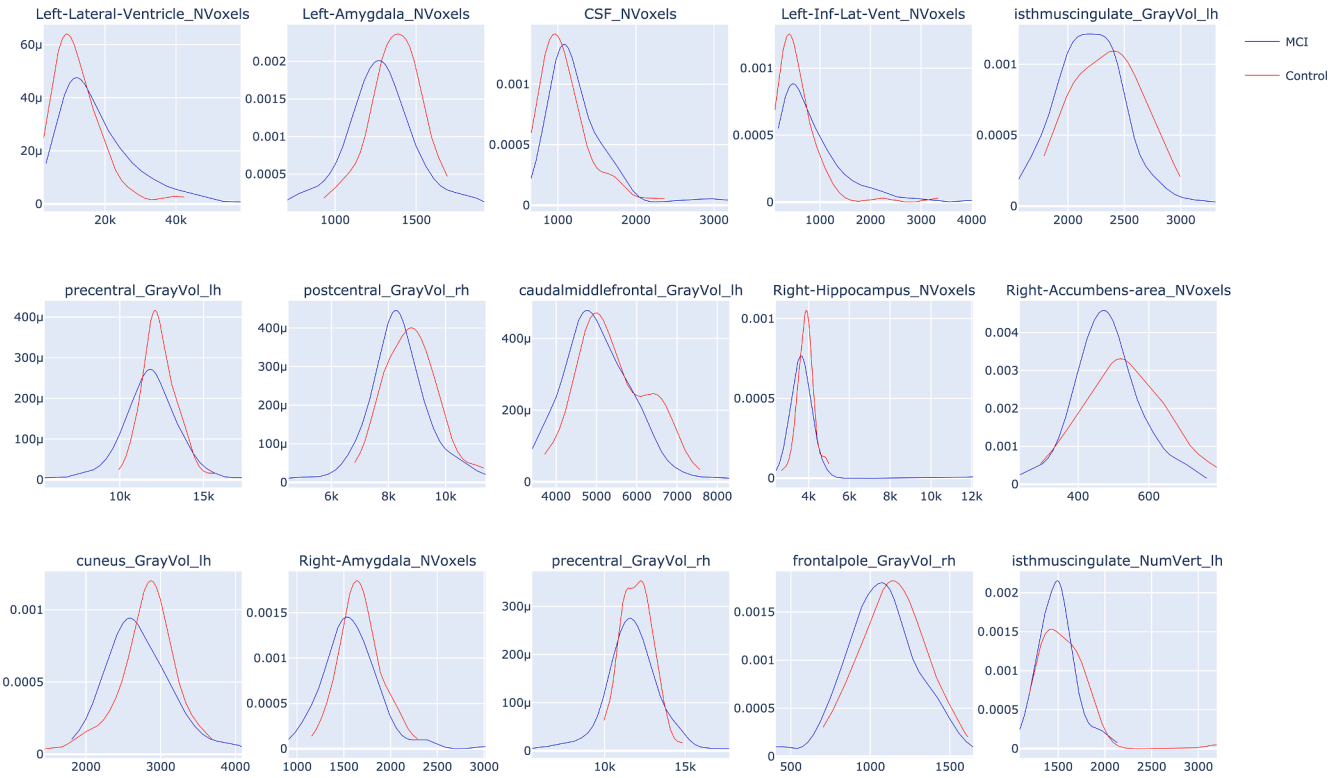

Fig. 3. The distribution plots of the most important predictive features for the best ML model. The X axis presents the feature values, the Y axis indicates the probability density. The red line - HC, the blue line - MCI patients. Overall, gray matter volume decreases in MCI patients in different brain areas, while CSF volume increases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Post-hoc T-test for symmetric hemispheres of the Left Lateral Ventricle, Left-Inf-Lat-Vent, Right Hippocampus, Right Accumbens. The significant features with p-value < 0.05 are highlighted in bold.

Feature name	p-value	test	Changes in brain structures (MCI)
Right-Lateral-Ventricle_NVoxels	**<0,001**	***U* test**	**increase**
Right-Inf-Lat-Vent_NVoxels	0.06	*U* test	increase
Left-Hippocampus_NVoxels	**<0,05**	***U* test**	**reduce**
Left-Accumbens-area_NVoxels	**<0,05**	***U* test**	**reduce**

were required for the optimal ML model: total neuropsychiatric inventory questionnaire score and left hippocampal volume. These characteristics correctly classified participants as remaining normal or developing cognitive impairment with 84.4 % accuracy (for the Alzheimer's Disease Neuroimaging Initiative et al., 2020). The authors of another study used diffusion tensor imaging (DTI) to construct a structural brain network from data from 96 healthy older adults. These characteristics were then integrated to predict cognitive abilities. A linear regression model and a Gaussian process model demonstrated the best ability to predict cognitive function. Moreover, these extracted topological properties of the brain structural network derived from DTI can also be considered as biosignals for further assessment of brain degeneration in healthy elderly and for early diagnosis of MCI (Ruijuan Yun et al., 2013). In our study, we presented another model, which was constructed based on data from a relatively small number of patients and healthy volunteers. However, we demonstrated the effectiveness of this model for differentiating between healthy individuals and patients with MCI. This model may be useful in clinical practice for the early diagnosis of cognitive decline in the elderly.

The development of new and improved models based on neuroimaging data has potential clinical relevance. Predicting the clinical progression of cognitive impairment can improve clinical research and clinical decision-making. A major opportunity and motivation for applying ML to neuroimaging testing in patients is the ability to make predictions for individuals. Such imaging procedures and data analysis will support personalized medicine (for the Alzheimer's Disease Neuroimaging Initiative et al., 2019; Mofrad et al., 2021). As mentioned above, early presymptomatic prediction of cognitive decline can provide an opportunity to prevent and delay the development of dementia, improve patients' quality of life, and reduce the burden on society and the economy. Another application of ML is to choose the next step in managing patient care. Given that some diagnostic tests are invasive or costly, selecting and screening the appropriate group of patients who are more likely to benefit from such tests can reduce unwanted side effects and costs in the entire population (for the Alzheimer's Disease Neuroimaging Initiative et al., 2019). Another major clinical application of these predictive models is to improve the efficiency of clinical trials by reducing the estimated sample size needed to observe the effect of the intervention. Trials may be more effective by including participants who are more likely to have cognitive impairment (for the Alzheimer's Disease Neuroimaging Initiative et al., 2019).

Despite the clinical relevance of our developed model, several factors need to be considered for its further use in clinical practice. The ML model developed should be evaluated using prospectively collected independent data in the next phase for further validation (for the Alzheimer's Disease Neuroimaging Initiative et al., 2020). Long-term observation is crucial because a subject with "compensatory normal cognition" may take almost a decade to transition to MCI (Molinuevo et al., 2017), so this factor must be considered. In addition, any methodological progression of ML should be tested at multiple sites in

different neuroimaging labs with the same or different equipment (Dimitriadis et al., 2018). Finally, the fact that there is no single neuroimaging method that can achieve absolute accuracy for automatic prediction should be considered, but integrating the best features from different methods can have a greater effect (Rathore et al., 2017), this factor should be considered for more accurate modeling.

In summary, our study uses baseline structural MRI data in cognitively healthy older adults and patients diagnosed with MCI to create a model using ML to predict cognitive decline in older adults. Our study confirms that ML models can be used to predict the progression of cognitive decline in patients with preclinical or prodromal disease. Our results may stimulate research in the field of early detection of cognitive decline and dementia. ML methods have enormous potential for use as clinical decision-making tools, but they must be carefully tested and validated against traditional diagnoses in different clinical settings, in different population cohorts, and in long-term studies.

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.

Data availability

The authors do not have permission to share data.

Acknowledgements

This research was supported by Russian Science Foundation grant #21-71-10136 (machine learning models creation and interpretation). This work was in part (MRI data collection and preprocessing) supported by the Ministry of Science and Higher Education of the Russian Federation (state assignment project No. 0723-2020-0036).

References

Albert, M., Zhu, Y., Moghekar, A., Mori, S., Miller, M. I., Soldan, A., … Wang, M.-C. (2018). Predicting progression from normal cognition to mild cognitive impairment for individuals at 5 years. *Brain, 141*, 877–887. https://doi.org/10.1093/brain/awx365

Apostolova, L. G., Mosconi, L., Thompson, P. M., Green, A. E., Hwang, K. S., Ramirez, A., … de Leon, M. J. (2010). Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. *Neurobiology of Aging, 31*, 1077–1088. https://doi.org/10.1016/j.neurobiolaging.2008.08.008

Ball, H. A., McWhirter, L., Ballard, C., Bhome, R., Blackburn, D. J., Edwards, M. J., … Carson, A. J. (2020). Functional cognitive disorder: Dementia's blind spot. *Brain, 143*, 2895–2903. https://doi.org/10.1093/brain/awaa224

Basaia, S., Agosta, F., Wagner, L., Canu, E., Magnani, G., Santangelo, R., & Filippi, M. (2019). Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. *NeuroImage: Clinical, 21*, Article 101645. https://doi.org/10.1016/j.nicl.2018.101645

Bidelman, G. M., Lowther, J. E., Tak, S. H., & Alain, C. (2017). Mild Cognitive Impairment Is Characterized by Deficient Brainstem and Cortical Representations of Speech. *The Journal of Neuroscience, 37*, 3610–3620. https://doi.org/10.1523/JNEUROSCI.3700-16.2017

Breton, A., Casey, D., & Arnaoutoglou, N. A. (2019). Cognitive tests for the detection of mild cognitive impairment (MCI), the prodromal stage of dementia: Meta-analysis of diagnostic accuracy studies. *International Journal of Geriatric Psychiatry, 34*, 233–242. https://doi.org/10.1002/gps.5016

Brickman, A. M., Provenzano, F. A., Muraskin, J., Manly, J. J., Blum, S., Apa, Z., … Mayeux, R. (2012). Regional White Matter Hyperintensity Volume, Not Hippocampal Atrophy, Predicts Incident Alzheimer Disease in the Community. *Archives of Neurology, 69*, 1621. https://doi.org/10.1001/archneurol.2012.1527

Burgmans, S., van Boxtel, M. P. J., Smeets, F., Vuurman, E. F. P. M., Gronenschild, E. H. B. M., Verhey, F. R. J., … Jolles, J. (2009). Prefrontal cortex atrophy predicts dementia over a six-year period. *Neurobiology of Aging, 30*, 1413–1419. https://doi.org/10.1016/j.neurobiolaging.2007.11.028

Dickerson, B.C., Wolk, D.A., Alzheimer's Disease Neuroimaging Initiative, 2012. MRI cortical thickness biomarker predicts AD-like CSF and cognitive decline in normal adults. Neurology 78, 84–90. Doi: 10.1212/WNL.0b013e31823efc6c.

Dimitriadis, S., Liparas, D., Alzheimer's DNI, 2018. How random is the random forest? Random forest algorithm on the service of structural imaging biomarkers for Alzheimer's disease: from Alzheimer's disease neuroimaging initiative (ADNI) database. Neural Regen Res 13, 962. Doi: 10.4103/1673-5374.233433.

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine Learning for Medical Imaging. *Radiographics, 37*, 505–515. https://doi.org/10.1148/rg.2017160130

for the Alzheimer's Disease Neuroimaging Initiative, Ezzati, A., Zammit, A.R., Harvey, D. J., Habeck, C., Hall, C.B., Lipton, R.B., 2019. Optimizing Machine Learning Methods to Improve Predictive Models of Alzheimer's Disease. JAD 71, 1027–1036. Doi: 10.3233/JAD-190262.

for the Alzheimer's Disease Neuroimaging Initiative, Gill, S., Mouches, P., Hu, S., Rajashekar, D., MacMaster, F.P., Smith, E.E., Forkert, N.D., Ismail, Z., 2020. Using Machine Learning to Predict Dementia from Neuropsychiatric Symptom and Neuroimaging Data. JAD 75, 277–288. Doi: 10.3233/JAD-191169.

Grothe, M., Heinsen, H., & Teipel, S. (2013). Longitudinal measures of cholinergic forebrain atrophy in the transition from healthy aging to Alzheimer's disease. *Neurobiology of Aging, 34*, 1210–1220. https://doi.org/10.1016/j.neurobiolaging.2012.10.018

https://adni.loni.usc.edu/, 2022.

https://surfer.nmr.mgh.harvard.edu/, 2022.

https://www.who.int/news-room/fact-sheets/detail/dementia, 2002.

Jongsiriyanyong, S., & Limpawattana, P. (2018). Mild Cognitive Impairment in Clinical Practice: A Review Article. *American Journal of Alzheimer's Disease and Other Dementias, 33*, 500–507. https://doi.org/10.1177/1533317518791401

Karas, G., Sluimer, J., Goekoop, R., van der Flier, W., Rombouts, S. A. R. B., Vrenken, H., … Barkhof, F. (2008). Amnestic Mild Cognitive Impairment: Structural MR Imaging Findings Predictive of Conversion to Alzheimer Disease. *AJNR. American Journal of Neuroradiology, 29*, 944–949. https://doi.org/10.3174/ajnr.A0949

Ledig, C., Schuh, A., Guerrero, R., Heckemann, R. A., & Rueckert, D. (2018). Structural brain imaging in Alzheimer's disease and mild cognitive impairment: Biomarker analysis and shared morphometry database. *Scientific Reports, 8*, 11258. https://doi.org/10.1038/s41598-018-29295-9

Machulda, M. M., Lundt, E. S., Albertson, S. M., Spychalla, A. J., Schwarz, C. G., Mielke, M. M., … Petersen, R. C. (2020). Cortical atrophy patterns of incident MCI subtypes in the Mayo Clinic Study of Aging. *Alzheimer's & Dementia, 16*, 1013–1022. https://doi.org/10.1002/alz.12108

Mofrad, S. A., Lundervold, A., & Lundervold, A. S. (2021). A predictive framework based on brain volume trajectories enabling early detection of Alzheimer's disease. *Computerized Medical Imaging and Graphics, 90*, Article 101910. https://doi.org/10.1016/j.compmedimag.2021.101910

Molinuevo, J.L., Rabin, L.A., Amariglio, R., Buckley, R., Dubois, B., Ellis, K.A., Ewers, M., Hampel, H., Klöppel, S., Rami, L., Reisberg, B., Saykin, A.J., Sikkes, S., Smart, C.M., Snitz, B.E., Sperling, R., Flier, W.M., Wagner, M., Jessen, F., Subjective Cognitive Decline Initiative (SCD-I) Working Group, 2017. Implementation of subjective cognitive decline criteria in research studies. Alzheimer's & Dementia 13, 296–311. Doi: 10.1016/j.jalz.2016.09.012.

Morozova, A., Zorkina, Y., Abramova, O., Pavlova, O., Pavlov, K., Soloveva, K., … Chekhonin, V. (2022). Neurobiological Highlights of Cognitive Impairment in Psychiatric Disorders. *IJMS, 23*, 1217. https://doi.org/10.3390/ijms23031217

Nemoto, K., Yamashita, F., Ohnishi, T., Matsuda, H., & Asada, T. (2006). P2–197: Hypoperfusion in anterior cingulate cortex in the earliest stage of mild cognitive impairment: A community–based study. *Alzheimer's & Dementia, 2*. https://doi.org/10.1016/j.jalz.2006.05.1035

Nestor, S.M., Rupsingh, R., Borrie, M., Smith, M., Accomazzi, V., Wells, J.L., Fogarty, J., Bartha, R., the Alzheimer's Disease Neuroimaging Initiative, 2008. Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. Brain 131, 2443–2454. Doi: 10.1093/brain/awn146.

Parisot, S., Ktena, S. I., Ferrante, E., Lee, M., Guerrero, R., Glocker, B., & Rueckert, D. (2018). Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease. *Medical Image Analysis, 48*, 117–130. https://doi.org/10.1016/j.media.2018.06.001

Rathore, S., Habes, M., Iftikhar, M. A., Shacklett, A., & Davatzikos, C. (2017). A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer's disease and its prodromal stages. *NeuroImage, 155*, 530–548. https://doi.org/10.1016/j.neuroimage.2017.03.057

Reinvang, I., Grambaite, R., & Espeseth, T. (2012). Executive Dysfunction in MCI: Subtype or Early Symptom. *International Journal of Alzheimer's Disease, 2012*, 1–8. https://doi.org/10.1155/2012/936272

Risacher, S.L., Saykin, A.J., 2019. Neuroimaging in aging and neurologic diseases, in: Handbook of Clinical Neurology. Elsevier, pp. 191–227. Doi: 10.1016/B978-0-12-804766-8.00012-1.

Roberts, R., & Knopman, D. S. (2013). Classification and Epidemiology of MCI. *Clinics in Geriatric Medicine, 29*, 753–772. https://doi.org/10.1016/j.cger.2013.07.003

Yun, R., Lin, C.-C., Shuicai, W.u., Huang, C.-C., Lin, C.-P., & Chao, Y.-P. (2013). A prediction model for cognitive performance in health ageing using diffusion tensor imaging with graph theory. In *In: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 527–530). Osaka: IEEE. https://doi.org/10.1109/EMBC.2013.6609553.

Schmitter, D., Roche, A., Maréchal, B., Ribes, D., Abdulkadir, A., Bach-Cuadra, M., … Krueger, G. (2015). An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *NeuroImage: Clinical, 7*, 7–17. https://doi.org/10.1016/j.nicl.2014.11.001

Sharaev, M., Artemov, A., Kondrateva, E., Ivanov, S., Sushchinskaya, S., Bernstein, A., … Burnaev, E. (2018). Learning Connectivity Patterns via Graph Kernels for fMRI-Based Depression Diagnostics. In *In: 2018 IEEE International Conference on Data Mining Workshops (ICDMW). Presented at the 2018 IEEE International Conference on*

Data Mining Workshops (ICDMW) (pp. 308–314). Singapore, Singapore: IEEE. https://doi.org/10.1109/ICDMW.2018.00051.

Sharaev, M., Artemov, A., Kondrateva, E., Sushchinskaya, S., Burnaev, E., Bernstein, A., … Andreev, A. (2018). MRI-Based Diagnostics of Depression Concomitant with Epilepsy. In *Search of the Potential Biomarkers, in: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). Presented at the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 555–564). Turin, Italy: IEEE. https://doi.org/10.1109/DSAA.2018.00071.

Song, M., Jung, H., Lee, S., Kim, D., & Ahn, M. (2021). Diagnostic Classification and Biomarker Identification of Alzheimer's Disease with Random Forest Algorithm. *Brain Sciences, 11*, 453. https://doi.org/10.3390/brainsci11040453

Suk, H.-I., Shen, D., 2013. Deep Learning-Based Feature Representation for AD/MCI Classification, in: Salinesi, C., Norrie, M.C., Pastor, Ó. (Eds.), Advanced Information Systems Engineering, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 583–590. Doi: 10.1007/978-3-642-40763-5_72.

Suk, H.-I., Wee, C.-Y., Lee, S.-W., & Shen, D. (2016). State-space model with deep learning for functional dynamics estimation in resting-state fMRI. *NeuroImage, 129*, 292–307. https://doi.org/10.1016/j.neuroimage.2016.01.005

Tondelli, M., Wilcock, G. K., Nichelli, P., De Jager, C. A., Jenkinson, M., & Zamboni, G. (2012). Structural MRI changes detectable up to ten years before clinical Alzheimer's disease. *Neurobiology of Aging, 33*, 825.e25–825.e36. https://doi.org/10.1016/j.neurobiolaging.2011.05.018

Venugopalan, J., Tong, L., Hassanzadeh, H. R., & Wang, M. D. (2021). Multimodal deep learning models for early detection of Alzheimer's disease stage. *Scientific Reports, 11*, 3254. https://doi.org/10.1038/s41598-020-74399-w

Yue, L., Hu, D., Zhang, H., Wen, J., Wu, Y., Li, W., … Xiao, S. (2021). Prediction of 7-year's conversion from subjective cognitive decline to mild cognitive impairment. *Human Brain Mapping, 42*, 192–203. https://doi.org/10.1002/hbm.25216