UNIVERSIDAD NACIONAL DE COLOMBIA

Mining relationships between multi-modal data to characterize Alzheimer's disease manifestation

German Alejandro Pabón Ochoa

Universidad Nacional de Colombia
Faculty of Medicine
Bogotá, Colombia
2024

Mining relationships between multi-modal data to characterize Alzheimer's disease manifestation

German Alejandro Pabón Ochoa

Thesis or degree work presented as a partial requirement to apply for the title of:
Master in Biomedical Engineering

Director
Eduardo Romero Ph.D.

Co-director
Diana Lorena Giraldo Franco Ph.D.

Research Line
Digital Anatomy by Images

Research Group
CIM@LAB - Computer Imaging and Medical Applications Laboratory

Universidad Nacional de Colombia
Faculty of Medicine
Bogotá, Colombia
2024

Se acabó el encanto

Acknowledgments

I express my gratitude to my parents, Gladys and Reinaldo, who have not only been my role models but also my support and constant source of motivation. Their loving words always inspire me to finish what I start. I also extend my thanks to my family, whose aid was crucial during these years, even amid the pandemic; they significantly contributed to making this achievement possible. To my girlfriend, María Fernanda, thank you for being my partner and for being by my side. My loved ones, who are no longer with us, are my stars in the firmament. I thank them for guiding me through this journey.

To my professor, Eduardo Romero, for opening the doors of his research group to me and providing the opportunity for both academic and personal growth. I admire his impressive career, exemplary research, and dedication to all his students.

I appreciate Diana Giraldo, my tutor in this research project, for her time, valuable contributions, and comments. Her collaboration was crucial in identifying and correcting flaws and errors, and her congratulations provided the necessary boost to keep going. Without her collaboration, this project wouldn't be what it is. Also, to all members of the CIM@LAB Research Group, your interest in my work and your significant contributions to telling the story clearly were invaluable.

Last but not least, I thank my past self, who, despite knowing the challenges ahead, decided to embark on this process without fully knowing what would come. This challenge, which I have now overcome, has imparted valuable lessons. Although there were aspects I still needed to learn and others I managed to assimilate, the experience taught me that with effort and determination, I can achieve my goals.

Abstract

Mining relationships between multi-modal data to characterize Alzheimer's disease manifestation

The heterogeneity in the clinical manifestation of Mild Cognitive Impairment (MCI) poses a significant challenge. A comprehensive characterization of Alzheimer's Disease (AD) in this early stage allows for timely detection, prediction of disease progression, and, consequently, intervention and monitoring before clinical diagnosis of dementia. We present a quantitative strategy to characterize neuropsychological alterations in MCI patients at risk of developing AD dementia using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

A set of items from cognitive, functional, and behavioral tests was selected from a sample of cognitively impaired patients. The analysis of abnormal performance and the use of relational metrics allowed us to identify five clusters of items that could represent possible neuropsychological dimensions and could be used to quantitatively describe a cognitively impaired individual. These characteristics are represented by different cognitive and functional domains: 1) constructional praxis; 2) orientation, memory, and daily living tasks; 3) language; 4) attention; and 5) processing speed tasks and executive functions. The proportion of variables exhibited within each characteristic quantifies the neuropsychological abnormality.

The utility of the proposed characterization was evaluated by two tasks. Firstly, predicting disease progression from MCI to AD dementia. We trained and tested a Support Vector Machine Classifier, achieving an accuracy of 0.76 within 36 months. Secondly, identifying MCI subgroups that exhibit diverse neuropsychological profiles and different patterns of gray matter atrophy. Individuals diagnosed with MCI were partitioned into seven subgroups. Upon comparison with cognitively normal individuals, the analysis using Voxel-Based Morphometry revealed specific brain regions with significant differences. We observed a close co-occurrence between cognitive impairments and structural changes. As cognitive and behavioral abnormalities increased, they were associated with more extensive patterns of gray matter atrophy.

This work offers an alternative approach to quantitatively characterize MCI subtypes and comprehend neurodegenerative patterns, providing valuable insights for enhanced characterization in the prodromal stage of Alzheimer's Disease.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, Neuropsychological Test, Prediction, Disease Progression, Quantitative Characterization, Voxel-Based Morphometry, Computational Neuroscience.

Resumen

Extracción de Relaciones Entre Datos Multimodales para Caracterizar la Manifestación de la Enfermedad de Alzheimer.

La heterogeneidad en la manifestación clínica del deterioro cognitivo leve (MCI, por sus siglas en inglés) plantea un desafío significativo. Una caracterización integral de la enfermedad de Alzheimer (AD, por sus siglas en inglés) en esta etapa temprana permite la detección oportuna, la predicción de la progresión de la enfermedad y, en consecuencia, la intervención y el monitoreo antes del diagnóstico clínico de demencia. Presentamos una estrategia cuantitativa para caracterizar alteraciones neuropsicológicas en pacientes con MCI en riesgo de desarrollar demencia por AD, utilizando datos de Alzheimer's Disease Neuroimaging Initiative (ADNI).

Un conjunto de variables de pruebas cognitivas, funcionales y conductuales fue seleccionado de una muestra de pacientes con deterioro cognitivo. El análisis del rendimiento anormal y el uso de métricas relacionales nos permitieron identificar cinco grupos de elementos que podrían representar posibles dimensiones neuropsicológicas de la enfermedad y que podrían utilizarse para describir cuantitativamente a un individuo con deterioro cognitivo. Estas características están representadas por diferentes dominios cognitivos y funcionales: 1) Praxis constructiva, 2) Orientación, Memoria y tareas de la vida diaria, 3) Lenguaje, 4) Atención, y 5) Tareas de velocidad de procesamiento y funciones ejecutivas. La proporción de variables exhibidas dentro de cada característica cuantifica la anormalidad neuropsicológica.

La utilidad de la caracterización propuesta fue evaluada mediante dos tareas. En primer lugar, prediciendo la progresión de la enfermedad desde la MCI hasta la demencia de Alzheimer. Entrenamos y probamos un Clasificador de Máquina de Vectores de Soporte, logrando una precisión del 0,76 dentro de los 36 meses. En segundo lugar, identificando subgrupos de MCI que exhibieron perfiles neuropsicológicos diversos y diferentes patrones de atrofia de materia gris. Individuos diagnosticados con MCI fueron divididos en siete subgrupos. Al comparar con individuos cognitivamente normales, el análisis utilizando Morfometría Basada en Vóxeles reveló regiones cerebrales específicas con diferencias significativas. Observamos una estrecha co-ocurrencia entre los deterioros cognitivos y los cambios estructurales. A medida que aumentaban las anormalidades cognitivas y conductuales, estas se asociaban con patrones más extensos de atrofia de materia gris.

Este trabajo ofrece un enfoque alternativo para caracterizar cuantitativamente los subtipos de MCI y comprender los patrones neurodegenerativos, proporcionando información valiosa para una mejor caracterización en la etapa prodrómica de la enfermedad de Alzheimer.

Palabras clave: Enfermedad de Alzheimer, Deterioro Cognitivo Leve, Pruebas Neuropsicológicas, Predicción, Progresión de la Enfermedad, Caracterización Cuantitativa, Morfometría Basada en Vóxeles, Neurociencia Computacional.

Contents

1 Introduction

Alzheimer's disease (AD) is a chronic, progressive, unremitting, and irreversible neurodegenerative condition [1]. AD is the leading cause of dementia, accounting for 60 to 80 % of dementia cases in individuals older than 65 years of age [2, 3]. The prevalence of AD increases with age and is more common in women than in men, two out of three patients with this dementia are women [4, 2, 5]. By 2050, it is projected that the number of affected individuals will reach 131.5 million worldwide due to increasing longevity [6].

The diagnosis of AD requires a clinical evaluation that assesses the likelihood of AD being the underlying cause of anatomical and cognitive deterioration. A comprehensive and high-confidence diagnosis should include the analysis of multiple sources of information, such as medical history, neuropsychological examinations, selected laboratory tests, and neuroimaging evaluations. There is no single test or biomarker that can definitively establish the diagnosis.

In current conventional clinical settings, the diagnosis is typically made by the physician, who evaluates and weighs the available information [3]. While cognitive and neuroanatomical variables are employed in quantitative analyses of the disease, they are often treated as separate and independent entities. However, these data, derived from different sources of information, offer complementary yet interdependent information, as the underlying pathology is the same [7, 8]. Therefore, the integration and joint analysis of multi-modal data, exploiting their inherent correlations, could provide a more comprehensive, precise, and detailed characterization of different manifestations of the disease [8].

The risk of AD before the onset of clinical symptoms can be estimated using molecular biomarkers such as cerebrospinal fluid (CSF) measures and known genetic risk factors like the apolipoprotein E (APOE4) $\epsilon 4$ genotype [9, 10]. While these biomarkers are precise, they come with drawbacks such as being expensive, requiring invasive procedures like lumbar puncture, and being impractical for extensive screening of large populations [11, 1]. Furthermore, their clinical use is limited to stratifying patients, predicting disease evolution, and assessing treatment response.

The characteristic pathophysiological processes of AD begin long before the onset of symptoms (with a mean duration of 20 years) [4]. These processes constitute a preclinical stage

followed by a prolonged prodromal stage of the disease [12] and involve the accumulation of beta-amyloid (Aβ) protein plaques in extracellular spaces and the aggregation of tau tangles. These amyloid plaques and neurofibrillary tangles appear in various regions of the cerebral cortex and hippocampus, leading to neuronal damage and synaptic degeneration. The initiation of this process has been localized in the temporal lobes, and as the disease progresses, the pathological involvement spreads to other areas of the neocortex, leading to macro-structural changes in brain anatomy such as cortical atrophy. [4].

The prodromal phase, identified as Mild Cognitive Impairment (MCI), is a clinical stage characterized by cognitive decline beyond what is considered normal for age, level of education, and cultural background [12]. Although MCI may exhibit overlapping patterns with aging, it is distinct from the normal healthy aging process [1], and poses a high risk of progression to AD dementia[13, 14]. MCI is characterized by subtle yet measurable cognitive symptoms, often accompanied by subjective complaints [15]. Unlike dementia, MCI does not significantly interfere with daily living [16].

In the symptomatic phase of the disease (with a mean duration of 10 years) [4], individuals exhibit important disturbances across multiple cognitive, functional, and behavioral domains that gradually worsen over time. Progressive memory loss and difficulties in thinking, learning, and language are characteristic. Neuropsychological impairments become so severe over time that they interfere with even the simplest daily tasks and impact basic functional abilities [3].

Timely detection and prediction of AD in its early stages are crucial for providing appropriate and opportune therapeutic interventions [2, 15]. However, this is a challenge because large variability exists in the development and progression of clinical symptoms and pathophysiology among individuals with MCI [17, 18]. The condition involves a wide range of clinical presentations, indicating its heterogeneity in terms of neuropsychological profiles, neurological markers, progression rates, and underlying causes [13, 19, 20]. While a typical presentation is characterized by predominant memory deficits, there are variants with atypical manifestations, including executive function, visuospatial, or language deficits [18, 3, 21]. Cognitive symptoms, as well as functional and behavioral complications, are often undiagnosed, ignored, or misattributed [21].

Conventionally, MCI is classified into two major subtypes: amnestic and non-amnestic, which can further be categorized as single-domain or multi-domain [20, 13, 22]. However, this classification may not fully capture the wide range of variability, as it tends to group individuals with different neuropsychological profiles [23]. Moreover, it relies on diagnostic criteria that may be susceptible to false-positive errors [24].

Recognizing the heterogeneity and identifying subtypes within MCI is crucial for improving patient management [25]. Recent research has identified MCI subtypes using cognitive information and considering patterns of impairment across multiple cognitive domains without solely relying on the amnestic/non-amnestic distinction [26, 27, 24, 17, 28, 29]. These findings highlight the importance of a comprehensive characterization that takes into account various factors to capture the heterogeneity.

The neuropsychological heterogeneity could be reflected in different patterns of change in brain anatomy. Incorporating neuroimaging techniques, such as in vivo magnetic resonance imaging (MRI), provides a noninvasive tool to characterize the MCI. Exploring cerebral atrophy as a complementary approach could help understand the potentially diverse neurodegenerative processes in MCI subtypes.

As far as we know, two previous investigations have examined the association between brain atrophy and cognitive impairment profiles in MCI patients by assessing cortical thickness in pre-defined regions of interest across distinct databases [23, 30]. However, beyond cortical thickness patterns, other analysis techniques without prior assumptions about the location of differences, such as voxel-based morphometry (VBM), enable the identification of brain regions with structural alterations.[31, 32].

1.1. Research Question

How to characterize heterogeneity in early manifestations of Alzheimer's disease?

1.2. Main objective

To build a strategy that exploits relations between multiple sources of information to obtain a quantitative characterization of Alzheimer's Disease manifestation.

1.3. Proposed Approach and Contribution

This work introduces a novel strategy for the quantitative characterization of Alzheimer's disease in its early stages by constructing relational metrics between cognitive, functional, and behavioral tests. A simultaneous analysis of this information captures the heterogeneity of the condition and provides a complementary description of the early neuropsychological symptoms of AD.

We evaluate our strategy by using the proposed neuropsychological characterization in the automated prediction of progression from MCI to AD dementia within the next 36 months,

48 months, and 60 months after baseline neuropsychological assessment.

Furthermore, we utilized the characterization to identify subgroups within MCI with distinct neuropsychological profiles, and we examined their patterns of brain atrophy. These subgroups were identified using a data-driven clustering approach that incorporated follow-up information. To explore gray matter atrophy patterns in MCI subgroups compared to a cognitively normal (CN) group, a comprehensive analysis using Voxel-Based Morphometry (VBM) was conducted on the whole brain. This analysis revealed specific brain regions with a notable percentage of significant differences.

This study contributes by quantitatively characterizing different subtypes of MCI. The findings enhance our understanding of the neurodegenerative patterns linked to the decline in neuropsychological function observed in MCI.

2 Materials

Dataset

This research utilized data from the Alzheimer's Disease Neuroimaging Initiative database (ADNI) (adni.loni.usc.edu). ADNI is a collaborative effort established in 2004 by Principal Investigator Dr. Michael W. Weiner, MD, involving numerous coinvestigators.
ADNI aims to advance our understanding of Alzheimer's disease (AD) and related cognitive disorders through the collection and analysis of longitudinal data. ADNI recruits subjects aged between 55 and 90 from 63 medical institutions, hospitals, or specialized clinics in the United States and Canada that specialize in the study, evaluation, and treatment of neurodegenerative diseases. Prior to data collection, written informed consent is obtained from all participants, emphasizing ethical principles and participant rights.
Subjects undergo a series of tests, including magnetic resonance imaging (MRI), positron emission tomography (PET), cerebrospinal fluid analysis (studying levels of beta-amyloid, total tau, and phosphorylated tau), blood tests (genetic study of the APOE e4 gene), general physical health clinical evaluations, medical history and medication questionnaires, and neuropsychological assessments with cognitive, functional, and behavioral tests. The information is analyzed, and according to specific criteria, subjects may be diagnosed as cognitively normal, having mild cognitive impairment, or having dementia likely due to Alzheimer's disease.
The initiative is characterized by its longitudinal design, which involves comprehensive follow-up assessments at regular intervals. Participants undergo screening, baseline, 3-month, 6-month, and annual visits, allowing researchers to track disease progression.
By leveraging this extensive dataset and longitudinal follow-up, ADNI enables researchers to identify biomarkers, predictive markers, and therapeutic targets associated with the progression of AD and related cognitive disorders. For up-to-date information on ADNI, please visit www.adni-info.org.

2.1. Participants

We took information from 2204 ADNI participants who had data from initial visits. The participants had their respective diagnoses of cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease Dementia (AD) according to the ADNI criteria

[33]. This information was downloaded from the database in August 2019. We analyzed the baseline data to determine if the participants met the inclusion and exclusion criteria for this study.
All participants had to have cognitive, functional, and behavioral test data from their first visit, as well as T1-weighted MRI scans and their diagnosis during follow-up visits. Those who lacked answers to any item of the selected tests were excluded. For the third step of the methodology, those who did not pass the quality control in their analysis of their MRIs were excluded. Additionally, patients diagnosed as Cognitively Normal had to maintain cognitive normality for a minimum of 36 months. Furthermore, they had to maintain their cognition within the parameters of normality throughout the entire follow-up period, and finally, exhibit normal cerebrospinal fluid biomarkers [9].
The subjects finally selected are partitioned into three sets: a normative or control sample, a learning sample, and an evaluation sample. These samples are utilized across various steps of the method, as further detailed in the subsequent section. The normative sample includes subjects who are cognitively normal. The learning sample comprises 30 % of subjects diagnosed with MCI and all subjects with AD dementia. The evaluation sample consists of the remaining 70 % of subjects diagnosed with MCI.

2.2. Cognitive, functional, and behavioral data

ADNI collects a wide range of cognitive, functional, and behavioral data. These tests assess the ability to perform instrumental activities of daily living that require higher cognitive skills and cover performance in various cognitive domains, including memory, orientation, attention, language, praxis, reasoning, and executive functions.
The cognitive tests included in ADNI are: the Alzheimer's Disease Assessment Scale Cognitive (ADAS-Cog13), the Mini Mental State Exam (MMSE), the Logical Memory Test (LMT), the Rey Auditory Verbal Learning Test (AVLT), the Clock Drawing Test (CDT), the Category Fluency Test (CFT), the Trail Making Test (TMT), the American National Adult Reading Test (ANART), the Boston Naming Test (BNT), and the Montreal Cognitive Assessment (MoCA).
The functional and behavioral tests included in ADNI are the Clinical Dementia Rating (CDR), the Functional Assessment Questionnaire (FAQ), and the Everyday Cognition (ECOG).
In this research, we include all cognitive, functional, and behavioral tests, along with their respective test items, available in the database. The ECOG test was excluded because it lacked information for several subjects, which would have significantly reduced the final sample size.
Some test item scores were merged since they belonged to the same task but were recorded individually.
From the MMSE, there were initially 30 scores which were consolidated into 11 items corresponding to the 11 tasks being implemented. Specifically: 5 scores from the time orientation

task were combined to form the item MMORITIME. 5 scores from the spatial orientation task were combined to form the item MMORISPACE. 3 scores from the immediate recall task were combined to form the item MMREGI. 3 scores from the delayed recall task were combined to form the item MMRECALL. 5 scores from the attention task were combined to form the item MMSPELLBKW. 2 scores from the naming task were combined to form the item MMNAM. 3 scores from the command task were combined to form the item MMCOMMAND. The repetition, reading, writing, and drawing tasks each had a single score, so they were taken as individual items, namely MMREPEAT, MMREAD, MMWRITE, and MMDRAW, respectively.
From the MOCA, there were initially 32 scores which were consolidated into 12 items corresponding to the 12 tasks being implemented. Specifically: 3 scores from the visuoconstructional skills (clock) task were combined to form the item MOCACLOCK. 3 scores from the naming task were combined to form the item MOCANAM. 2 scores from the span forward and backward task were combined to form the item MOCADIG. 5 scores from the serial 7s task were combined to form the item MOCASERIAL. 2 scores from the abstraction task were combined to form the item MOCAABS. 5 scores from the delayed recall task were combined to form the item MOCADLREC. 2 scores from the sentence repetition task were combined to form the item MOCAREP. 6 scores from the orientation task were combined to form the item MOCAORI. The alternating trail making, visuoconstructional skills (cube), verbal fluency, and letters and tapping tasks each had a single score, so they were taken as individual items, namely TRAILS, CUBE, MOCAFLUEN, and MOCALET, respectively.
It's important to note that scores from MMSE, LMT, AVLT, CDT, CFT, BNT, and MoCA items were reversed, with higher values indicating greater impairment of cognitive abilities.

2.3. Structural MRI

All newly enrolled subjects in ADNI-1 underwent scanning using the ADNI-1.5T scanning protocol, while those in ADNI-GO/ADNI2/ADNI3 underwent scanning using the ADNI-3T scanning protocol. These scans occurred during the screening, 3 months after the screening MRI, and then within 2 weeks before or after the month 6 and subsequent annual visits.
The ADNI MRI Core oversees all MRI image-related tasks, including selecting specific MRI pulse sequences, site qualification, and conducting quality control (QC) checks on all MRI data. Each site had to have at least one MRI scanner qualified for the ADNI MRI Study. ADNI supported MRI scanner protocols from General Electric (GE) Healthcare, Philips Medical Systems, and Siemens Medical Solutions.
ADNI performed quality control on the MRIs at the Mayo Clinic, where two levels of quality control were implemented: adherence to protocol parameters and series-specific quality. Trained analysts subjectively graded the scan quality. All T1 images underwent preprocessing steps by ADNI to ensure comparability. The preprocessed images Include appropriate gradient warping and robust image inhomogeneity corrections.

In the present study, anatomical 3D T1-weighted images labeled with "N3m"were downloaded. This label represents images with all preprocessing steps completed.

3 Methods

The presented methodology is divided into three steps (Fig.**3-1**). Firstly, a quantitative strategy is developed to characterize cognitive and behavioral alterations in subjects. Secondly, this characterization is evaluated by predicting future progression to dementia for MCI subjects. Lastly, as a continuation of the first step, another validation task is performed where a clustering method is applied to group MCI subjects based on their quantitative characterization. We aim to explore the relationship between the neuropsychological profiles and the potential patterns of brain atrophy among the subgroups of MCI.

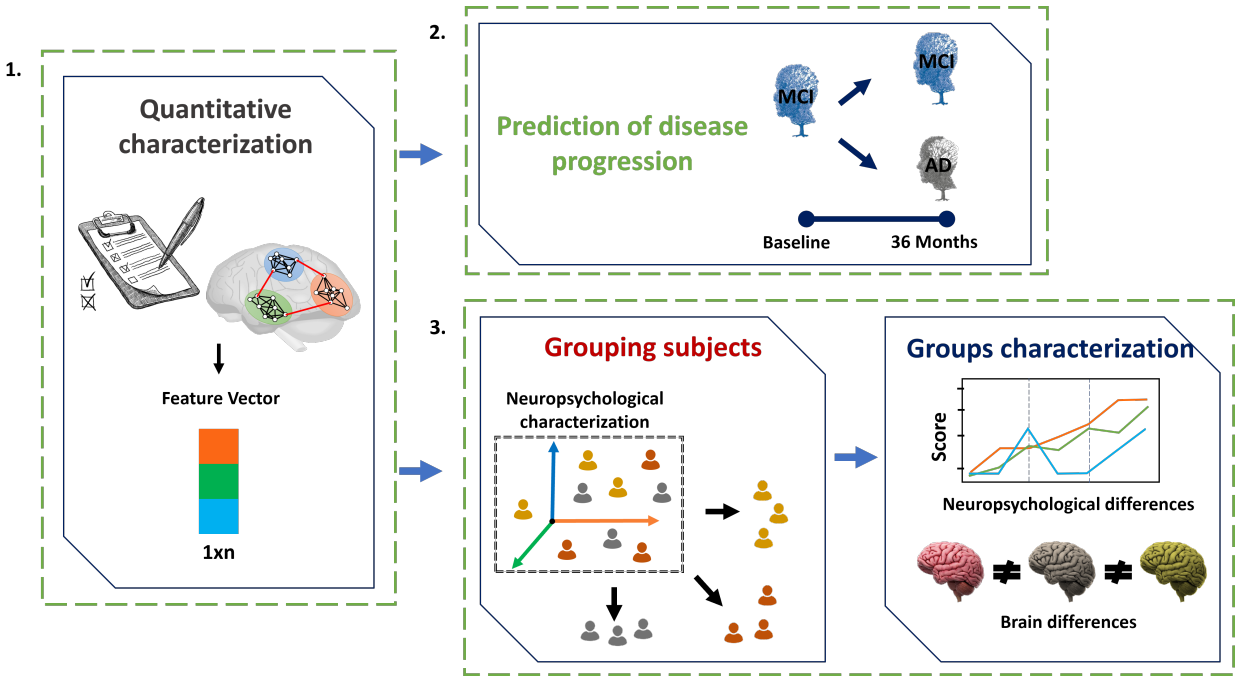

Figure 3-1: Pipeline

3.1. Quantitative characterization

Defining abnormal items

Given that the cognitive, functional, and behavioral items from the tests are measured in different units or scales, they were transformed into standardized z-scores using normative

data.
The process of regression-based standardization involved fitting a linear model for each item, with age and education as regressors, based on observations from the normative sample of control subjects. To calculate the z-scores, the difference between the observed score and the score predicted with the linear model was divided by the standard deviation of the residual, estimated during model fitting. The Standardized Regression-Based (SRB) z-scores [17], expressed in terms of standard deviation units from 0, are given by:

$$\text{z\_score} = \frac{\text{Observed Score} - \text{Predicted Score}}{\text{Standard Error of the Estimate}} \tag{3-1}$$

Then, the corresponding z-scores for neuropsychological items were dichotomized into normal or abnormal performance based on a cutoff. Standardized scores above 1,5 are typically considered abnormal, taking into account the relationship between percentile ranks and standard deviation units for the standard normal distribution [24, 34].

Learning the neuropsychological dimensions

For our proposed characterization, we aim to identify various neuropsychological dimensions based on neuropsychological performance within the disease and quantify the extent of abnormality for each cluster of items in a subject. The underlying assumption is that these dimensions can be learned by clustering items that are simultaneously abnormal in a set of cognitively impaired patients. A notion of distance between the considered items was established by examining the frequency with which abnormal measures co-occur in the learning sample.
Specifically, the distance between a pair of items, denoted as X and Y, is calculated using the conditional probabilities $P(X|Y)$ and $P(Y|X)$ as:

$$d(X,Y) = 1 - \frac{P(X|Y) + P(Y|X)}{2} \tag{3-2}$$

Conditional probabilities, where X=abnormal and Y=abnormal, capture the strength of the relation between items, even when they rarely occur in the population.
Having established this distance for pairs of items, a dendrogram was constructed using hierarchical agglomerative clustering with Ward's method [35]. The optimal number of clusters, delineating the partitioning of variables, was determined using the average silhouette method [36]. The outcome of this step is a partition of items into distinct clusters, each representing different possible neuropsychological dimensions.

Quantifying the manifestation of the disease

After capturing the possible neuropsychological dimensions using the clusters of test items, the next task is to quantify how much each subject manifests each one. An index is calculated

for each dimensions as the proportion of abnormal items exhibited by the subject within the cluster. Then, each subject can be quantitatively described by a vector, the length of which corresponds to the number of the possible neuropsychological dimensions.

3.2. Prediction of disease progression

The usefulness of the proposed characterization was evaluated based on the prediction of disease progression from MCI to AD dementia within the next 36 months, 48 months, and 60 months after baseline neuropsychological assessment. This involved training and testing automated binary classifiers on the evaluation sample of individuals with available diagnostic follow-up data within each specific time frame.

Specifically, Support Vector Machines (SVM) with linear kernels were employed to classify MCI subjects who remained stable and those who progressed to dementia. A random subsampling cross-validation scheme was implemented over 100 iterations. At each iteration, the SVM classifier was trained with a balanced subsample of its classes and tested with the remaining participants from the evaluation sample. For class balancing, 70 % of the subjects from the positive class (those who progressed to dementia) were selected, and an equal number of subjects were taken from the negative class (those who remained stable).

3.3. Grouping subjects

The quantitative characterization was used to identify subgroups of individuals with MCI within the evaluation sample. Multiple grouping strategies were tested, including two different clustering methods (k-means and hierarchical), three distances (Euclidean, Manhattan, and Dot product), and varying the number of groups (ranging from 2 to 8).

By employing purity (3-3) as a metric, we selected the grouping scheme to achieve optimal differentiation between stable MCI subjects and progressive MCI subjects within a diagnostic follow-up period of 36 months, 48 months, and 60 months. This involved ensuring that each subgroup predominantly consisted of subjects sharing the same label. Neuropsychological and anatomical differences are investigated by comparing each MCI subgroup against the control group.

$$Purity = \frac{1}{N}\sum_{i=1}^{K} \max_{j} |C_i \cap L_j| \tag{3-3}$$

Where N is the number of subjects, K is the number of groups, Ci is the set of subjects in group i, and Lj is the set of subjects in class j.

3.3.1. Neuropsychological differences

The statistical differences in indices for each neuropsychological dimension between the resulting subgroups and the CN group were investigated using Mann-Whitney Wilcoxon tests. A significance level for group differences was set at $p < 0{,}05$ after Bonferroni correction for multiple comparisons.

3.3.2. Brain differences

The analysis of structural MRI data was conducted using Voxel-Based Morphometry (VBM) [37] implemented in Statistical Parametric Mapping 12 (SPM12) software within MATLAB R2017b (MathWorks Inc., Natick, MA, USA).
VBM involves the following steps. First, the T1-weighted images are segmented into gray matter (GM), white matter, and cerebrospinal fluid, and the quality of all segmentations is visually assessed. We utilize the unified segmentation-normalization approach in SPM12, which determines the deformations required to align the images to a common stereotaxic space through affine and non-linear transformations. The images are then resampled to a voxel size of $1{,}5 \times 1{,}5 \times 1{,}5$ mm and normalized to the Montreal Neurological Institute (MNI) T1-weighted template. Afterward, the warped gray matter images are smoothed by convolving them with an isotropic Gaussian kernel with a full width at half maximum (FWHM) of 10 mm.
For the statistical analysis, a linear model was constructed, including age, sex, number of years of education, and total intracranial volume as covariates. This model was employed to investigate GM differences between each MCI subgroup and CN using two-sample t-tests with a significance threshold of $p < 0.05$, corrected for multiple comparisons by the family-wise error (FWE) method. VBM analysis enables the detection of brain differences and highlights areas with significant GM atrophy.
To determine the percentage of significant differences in specific brain regions for each MCI subgroup, the test statistic maps were segmented using the Automated Anatomical Labelling Atlas 3 (AAL3v1) [38]. This atlas is designed to be in the same three-dimensional space as the MNI template, ensuring accurate alignment between labeled brain structures in the atlas and the test statistic maps already in the MNI space. After extracting the regions from the test statistic maps, the proportion of voxels showing statistically significant differences is calculated for each region in the atlas.

4 Results

4.1. Demographic characteristics

Table **4-1** provides an overview of diagnostic groups and demographic features of the participants who met the inclusion and exclusion criteria.

Table 4-1: Demographic characteristics

Base line diagnosis	N	Sex (% Male)	Age (years)	Education (years)	ADNI cohorts	
					ADNI2	ADNIGO
CN	111	45.9	72.0 ±5.4	16.6 ±2.5	110	1
MCI	399	55.1	71.5 ±7.3	16.2 ±2.6	299	100
AD	102	59.8	74.6 ±7.7	15.7 ±2.6	102	0
	612	54.2	72.1 ±7.2	16.2 ±2.6	511	101

Abbreviations: N, number of subjects; CN, Cognitively normal; MCI, Mild cognitive impairment; AD, Alzheimer's disease Dementia.

In relation to subjects diagnosed as cognitively normal, they had to have a minimum diagnostic follow-up of 36 months, which continues over time. The table **4-2** shows the total follow-up in months and the number of CN subjects who comply with each follow-up period. Thus, their average total follow-up is 70 months, with a standard deviation of 16.22 months.

Table 4-2: Months of diagnostic follow-up for the CN subjects

Time frame	36	48	60	72	84	96
Total Subjects	4	19	20	26	32	10

The data partition into 3 sets is shown in Table **4-3**.

Table 4-3: Data partition

	Normative	Learning	Evaluation
CN	111	0	0
MCI	0	123	276
AD	0	102	0
	111	225	276

4.2. Selected cognitive, functional, and behavioral data

Once the linear model was adjusted for each item in the regression-based standardization process, ten items (4 from the MMSE, 4 from the CDR, and 2 from the FAQ) were excluded from the analysis due to their variances and regression coefficients being 0.

In this way, a total of 46 items were selected from the 10 cognitive tests, as follows: 13 from ADAS-Cog13, 7 from the MMSE, 2 from the LMT, 5 from the AVLT, 2 from the CDT, 1 from the CFT, 2 from the TMT, 1 from the ANART, 1 from the BNT, and 12 from the MoCA. On the other hand, 10 items were included from the 2 functional and behavioral tests, as follows: 2 from the CDR and 8 from the FAQ.

4.3. Possible neuropsychological dimensions

Using the established metric between test items and the hierarchical agglomerative clustering method, the items can be partitioned into 5 clusters, an optimal partition according to the average silhouette criterion (Figure **4-1**).

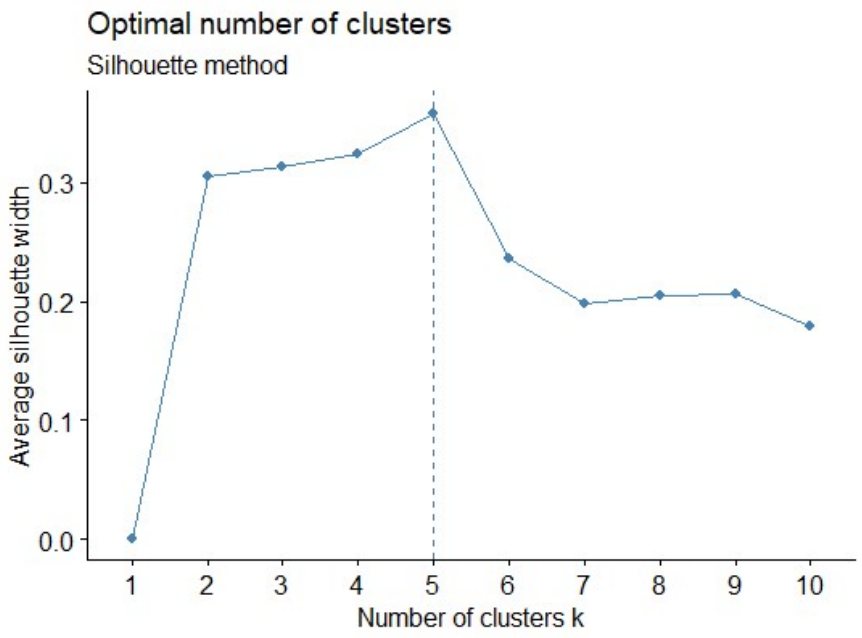

Figure 4-1: Optimal number of clusters

The resulting partition of the dendrogram is illustrated in Figure **4-2**. These clusters indicate possible neuropsychological dimensions can be described based on the cognitive, functional, and behavioral items they contain (Table **4-4**).

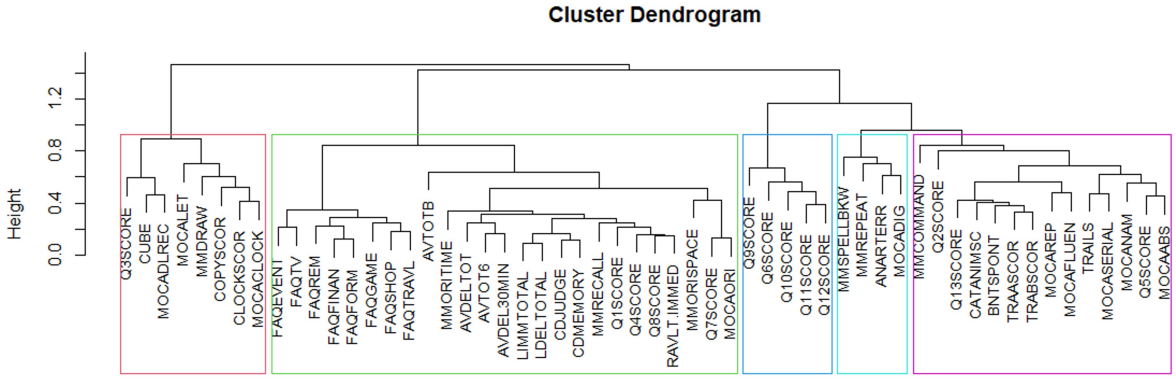

Figure 4-2: Partition of the Hierarchical Clustering

From left to right in Figure **4-2**:

- First dimension: is predominantly represented by items related to constructional praxis, involving drawing tasks of figures or patterns, along with memory and attention items from tasks requiring word recall and vigilance.

- Second dimension: primarily includes memory items from immediate and delayed recall memory tasks, all functional and behavioral items reflecting performance in daily living tasks, and all orientation items from tests assessing orientation in time and space.

- Third dimension: comprises items from ADAS-Cog13, predominantly related to language and linked with executive functions and praxis items. These correspond to tasks involving comprehension and spoken language, remembering instructions, and ideational praxis.

- Fourth dimension: contains two language items and two attention items from tasks involving reading, repeating, and spelling.

- Fifth dimension: encompasses a variety of domains, with a focus on language items coupled with executive functions, reasoning, attention, and memory items from tasks involving naming, repetition, lexical access, following commands, and processing speed.

Table 4-4: Items per dimensions

Dimension 1	Q3SCORE	ADAS-Cog13 Constructional Praxis
	CUBE	MoCA Visouconstructional Skills
	MOCADLREC	MoCA Delayed Recall
	MOCALET	MoCA Letters and Tapping
	MMDRAW	MMSE Construction
	COPYSCOR	CDT Clock Copying, Copy Condition
	CLOCKSCOR	CDT Clock Drawing, Command Condition
	MOCACLOCK	MoCA Visouconstructional Skills
Dimension 2	FAQEVENT	FAQ Keeping track of current events
	FAQTV	FAQ Paying attention to and understanding a TV program, book, or magazine
	FAQREM	FAQ Remembering appointments, family occasions, holidays, medications
	FAQFINAN	FAQ Writing checks, paying bills, or balancing a checkbook
	FAQFORM	FAQ Assembling tax records, business affairs, or other papers
	FAQGAME	FAQ Playing a game of skill such as bridge or chess, working on a hobby
	FAQSHOP	FAQ Shopping alone for clothes, household necessities, or groceries
	FAQTRAVL	FAQ Traveling out of the neighborhood, driving, or arranging to take public transportation
	AVTOTB	AVLT Immediate Recall List B
	MMORITIME	MMSE Orientation
	AVDELTOT	AVLT Recognition List A
	AVTOT6	AVLT Delayed Recall List A
	AVDEL30MIN	AVLT 30 Minute Delayed Recall List A
	LIMMTOTAL	LMT Immediate Recall
	LDELTOTAL	LMT Delayed Recall
	CDJUDGE	CDR Judgment and Problem Solving
	CDMEMORY	CDR Memory
	MMRECALL	MMSE Immediate Recall
	Q1SCORE	ADAS-Cog13 Word Recall
	Q4SCORE	ADAS-Cog13 Delayed Word Recall
	Q8SCORE	ADAS-Cog13 Word Recognition
	RAVLTIMMED	AVLT Immediate Recall List A
	MMORISPACE	MMSE Orientation
	Q7SCORE	ADAS-Cog13 Orientation
	MOCAORI	MoCA Orientation
Dimension 3	Q9SCORE	ADAS-Cog13 Remembering Test Instructions
	Q6SCORE	ADAS-Cog13 Ideational Praxis
	Q10SCORE	ADAS-Cog13 Comprehension of Spoken Language
	Q11SCORE	ADAS-Cog13 Word-Finding Difficulty
	Q12SCORE	ADAS-Cog13 Spoken Language Ability
Dimension 4	MMSPELLBKW	MMSE Attention
	MMREPEAT	MMSE Repetition
	ANARTERR	ANART Reading a List Of Words
	MOCADIG	MoCA Digit Span Forward and Backward
Dimension 5	MMCOMMAND	MMSE Command
	Q2SCORE	ADAS-Cog13 Commands/Instructions
	Q13SCORE	ADAS-Cog13 Number Cancellation
	CATANIMSC	CFT Verbal Fluency Given a Semantic Category
	BNTSPONT	BNT Naming
	TRAASCOR	TMT Part A Processing Speed and Executive Function
	TRABSCOR	TMT Part B Processing Speed and Executive Function
	MOCAREP	MoCA Sentence Repetition
	MOCAFLUEN	MoCA Verbal Fluency Given a Letter
	TRAILS	MoCA Alternating Trail
	MOCASERIAL	MoCA Serial 7s
	MOCANAM	MoCA Naming
	Q5SCORE	ADAS-Cog13 Naming
	MOCAABS	MoCA Abstraction

The indices for the five possible neuropsychological dimensions were computed across all participants in the evaluation sample. The mean and standard deviation of these indices are outlined in Table **4-5**. It should be noted that for individuals with MCI, potential AD symptoms are subtle; hence, lower values for the vectors of indices are expected.

Table 4-5: Manifestation indices

Dimensions	Dimension 1	Dimension 2	Dimension 3	Dimension 4	Dimension 5
Indices	0.19 ±0.17	0.35 ±0.19	0.08 ±0.14	0.15 ±0.21	0.20 ±0.15

4.4. Prediction of disease progression

As the months passed, the evaluation sample decreased due to patients either passing away or choosing to no longer participate in the ADNI study. Consequently, the sample of subjects differs across various follow-up periods. Table **4-6** shows the distribution of the evaluation sample for training and testing for each diagnostic follow-up time frame.

Table 4-6: Distribution of evaluation sample for Training and Testing for each time frame

Diagnostic follow-up in month	Stable	Converter	Train		Test	
			Stable	Converter	Stable	Converter
36	148	72	52	52	96	20
48	117	83	60	59	57	24
60	82	86	62	62	20	24

The performance of the quantitative characterization in the task of predicting progression from MCI to dementia for each follow-up time frame was evaluated with classification performance metrics: sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV). The mean and standard deviation across iterations of the cross-validation scheme are presented in Table **4-7**.

Table 4-7: Performance of disease progression prediction.

Time	Sensivity	Specifity	Accuracy	PPV	NPV
36	**0.67 ±0.09**	**0.78 ±0.05**	**0.76 ±0.04**	0.40 ±0.06	**0.92 ±0.02**
48	0.65 ±0.10	0.78 ±0.05	0.74±0.03	0.56 ±0.05	0.84 ±0.03
60	0.65 ±0.09	0.76 ±0.07	0.70±0.06	**0.76 ±0.06**	0.65 ±0.06

According these results, the prediction of progression from MCI to dementia seems to be more accurate for the 36-month follow-up period. It is important to note that the positive

and negative predictive values are influenced by the class imbalance in the test subjects, as illustrated in Table **4-6** and reflected in the confusion matrices (Table **4-8**).

Table 4-8: Confusion matrix for each time frame

Diagnostic follow-up in month	Prediction	Ground truth	
		Stable	Converter
36	Stable	75.4 ± 4.8	6.5 ± 1.8
	Converter	20.6 ± 4.8	13.4 ± 1.8
48	Stable	44.5 ± 3.0	8.2 ± 2.5
	Converter	12.5 ± 3.0	15.7 ± 2.5
60	Stable	15.2 ± 1.5	8.2 ± 2.3
	Converter	4.8 ± 1.55	15.7 ± 2.3

4.5. Grouping subjects

Three subjects from the normative sample and six subjects from the evaluation sample were excluded because their magnetic resonance images did not pass the quality control during the segmentation step. Table **4-9** shows the distribution of the evaluation sample for each diagnostic follow-up period, subtracting the patients excluded for the present results.

Table 4-9: Distribution of evaluation sample for Training and Testing for each time frame

Diagnostic follow-up in month	Stable	Converter
36	146	68
48	116	78
60	81	81

Different grouping strategies were applied to separate the MCI subjects based on their diagnostic follow-up. The purity was calculated for each strategy in the time frames of 36 months, 48 months, and 60 months (Table **4-10**, Table **4-11**, and Table **4-12**, respectively).
When analyzing the final groups of the groupings with greater purity. We realized that this metric does not accurately reflect how the classes are distributed among the groups. High overall purity does not necessarily imply high intragroup purity. Intragroup purity is defined as the dominance of a class within a group. Therefore, we conducted additional verification on the groupings that had overall purities greater than or equal to 0,74, which is the lowest purity among the highest purities of the different follow-up times. The highest purity for the 36-month follow-up was 0.78, for 48 months, it was 0.75, and for 60 months, it was 0.74.

Table 4-10: Purity calculated for each grouping strategy for the 36-month follow-up.

Number of	Euclidean		Manhattan		Dot product	
Groups	K-mean	Hierarchical	K-means	Hierarchical	K-means	Hierarchical
2	0,74	0,68	0,73	0,68	0,76	0,68
3	0,71	0,68	0,72	0,71	0,78	0,68
4	0,74	0,68	0,76	0,71	0,75	0,69
5	0,73	0,68	0,72	0,71	0,77	0,69
6	0,75	0,73	0,74	0,71	0,78	0,70
7	0,76	0,73	0,73	0,71	0,78	0,70
8	0,75	0,74	0,77	0,71	0,78	0,71

Table 4-11: Purity calculated for each grouping strategy for the 48-month follow-up.

Number of	Euclidean		Manhattan		**Dot product**	
Groups	K-mean	Hierarchical	K-means	Hierarchical	**K-means**	Hierarchical
2	0,67	0,68	0,65	0,59	0,72	0,59
3	0,67	0,68	0,64	0,69	0,73	0,60
4	0,71	0,68	0,69	0,69	0,73	0,60
5	0,67	0,68	0,69	0,69	0,72	0,61
6	0,69	0,68	0,70	0,69	0,74	0,61
7	0,71	0,68	0,72	0,69	**0,75**	0,62
8	0,72	0,70	0,74	0,70	0,73	0,62

Table 4-12: Purity calculated for each grouping strategy for the 60-month follow-up.

Number of	Euclidean		Manhattan		Dot product	
Groups	K-mean	Hierarchical	K-means	Hierarchical	K-means	Hierarchical
2	0,58	0,64	0,59	0,51	0,69	0,50
3	0,66	0,64	0,66	0,56	0,66	0,50
4	0,64	0,64	0,66	0,61	0,69	0,51
5	0,65	0,64	0,66	0,61	0,68	0,51
6	0,66	0,64	0,68	0,61	0,68	0,52
7	0,69	0,68	0,67	0,64	0,70	0,53
8	0,69	0,69	0,69	0,64	0,74	0,53

Of the 20 grouping strategies with purities greater than or equal to 0.74, we selected the strategy for the 48-month follow-up that uses k-means as the clustering method, dot product as the distance measure, and a total of 7 groups. The overall purity for this strategy is 0.75.

When analyzing the distribution of patients among the groups and the predominance of classes, we verified that this strategy had the highest minimum intragroup purity value compared to the other strategies. The resulting MCI subgroups are described in Table **4-13**.

Table 4-13: Subgroups of MCI subjects.

Subgroup	N	Stable (%)	Converters (%)	Sex (% Male)	Age (Years)
A	19	14 (74 %)	5 (26 %)	63.1	71.1±6.2
B	50	36 (72 %)	14 (28 %)	50.0	72.2±7.6
C	19	18 (95 %)	1 (5 %)	36.8	70.7±5.5
D	46	33 (72 %)	13 (28 %)	60.8	71.6±7.7
E	29	8 (28 %)	21 (72 %)	51.7	70.2±7.8
F	23	6 (26 %)	17 (74 %)	65.2	73.9±6.2
G	8	1 (12 %)	7 (87 %)	50.0	75.9±4.1
Total	194	116 (60 %)	78 (40 %)	54.6	71.9±7.1

Abbreviations: N, number of subjects.

Consequently, data from 302 individuals was included in this step (Table **4-14**).

Table 4-14: Evaluation subsample

Group	N	Sex (% M)	Age (years)	Education (years)	MMSE
CN	108	44.4	72.0 ±5.5	16.6 ±2.6	29.13±1.07
MCI	194	54.6	71.9 ±7.1	16.3 ±2.5	27.91±1.71
	302	50.9	71.9 ±6.6	16.4 ±2.5	28.35±1.62

Abbreviations: N, number of subjects.
MMSE, Total Score Mini-Mental State Examination.

4.5.1. Neuropsychological differences

The distribution of each neuropsychological feature value for each MCI subgroup is shown in Fig. **4-3**. A score of 1 indicates the highest abnormality for that feature. We can observe that all subgroups exhibit significantly abnormal performance compared to cognitively normal in the feature related to orientation, memory, and activities of daily living, as in the case

of subgroup A, which only shows impairment in that particular feature. Subgroup B also exhibits impairments in processing speed and executive function. Subgroup C demonstrates deficits in the aforementioned feature as well as in attention, processing speed, and executive functions. Subgroups D and E display abnormal performance in all dimensions except attention. Lastly, subgroups F and G exhibit abnormalities in all dimensions.

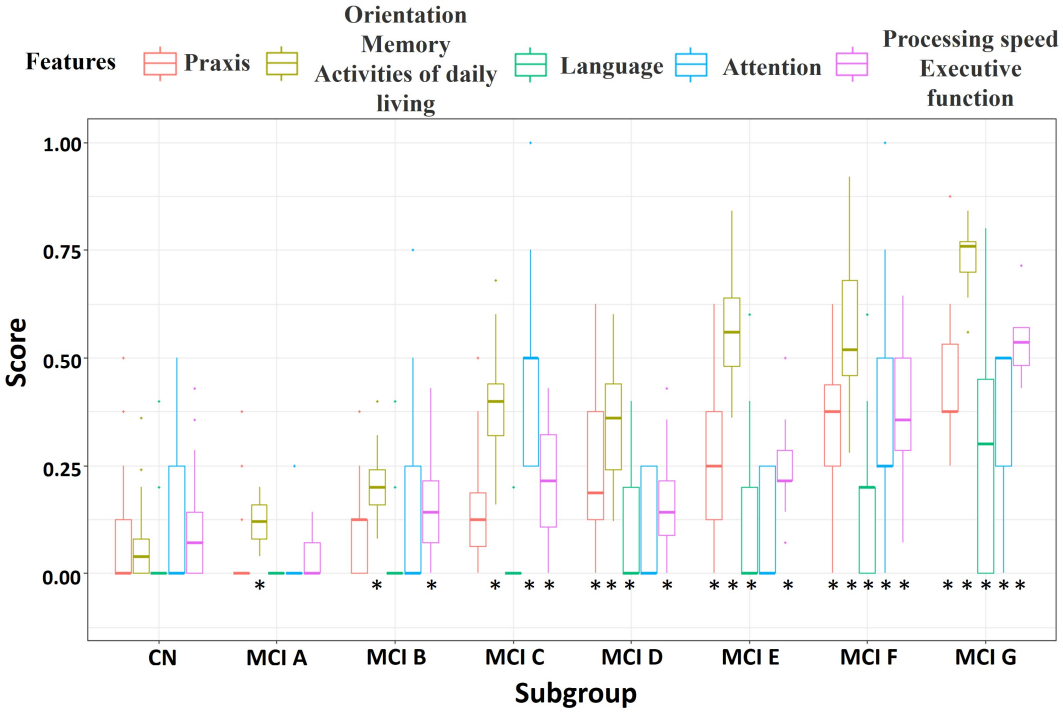

Figure 4-3: Neuropsychological performance for the MCI subgroups.
*Dimension significantly different relative to CN ($p < 0{,}05$) after Bonferroni correction.

4.5.2. Brain differences

Fig. **4-4** illustrates the test statistic maps, highlighting the areas of significantly reduced GM density for the MCI subgroups relative to CN group. No significant differences were found between subgroups A and C when compared to CN. For subgroups B, D, and E, differences in subcortical and limbic lobe regions stand out, with the amygdala and hippocampus being particularly representative. Subgroups F and G exhibit a more widespread pattern of differences in brain regions. Subgroup G is characterized by larger areas of brain regions that differentiate it from the CN group.

Table **4-15** shows the regions where the percentage of voxels with significant differences is greater than 1 % of the total regional area when comparing the resulting MCI subgroups

against CN. We observe a gradual increase in the percentage of voxels with significant differences in the amygdala, hippocampus, parahippocampal gyrus, and the superior segment of the temporal pole when comparing these regions across the different subgroups. This trend suggests a possible progressive involvement of these brain structures as the condition of the subjects deteriorates. The increasing percentage of affected voxels highlights the escalating impact of atrophy on these critical areas associated with memory and emotional regulation. There does not appear to be a predominance of significant differences in either hemisphere. The distribution of significant differences is relatively balanced between the left and right hemispheres, indicating that the observed changes are not confined to one side of the brain but are instead distributed across both hemispheres. This suggests that the neurodegenerative processes affecting the regions studied impact both hemispheres similarly.

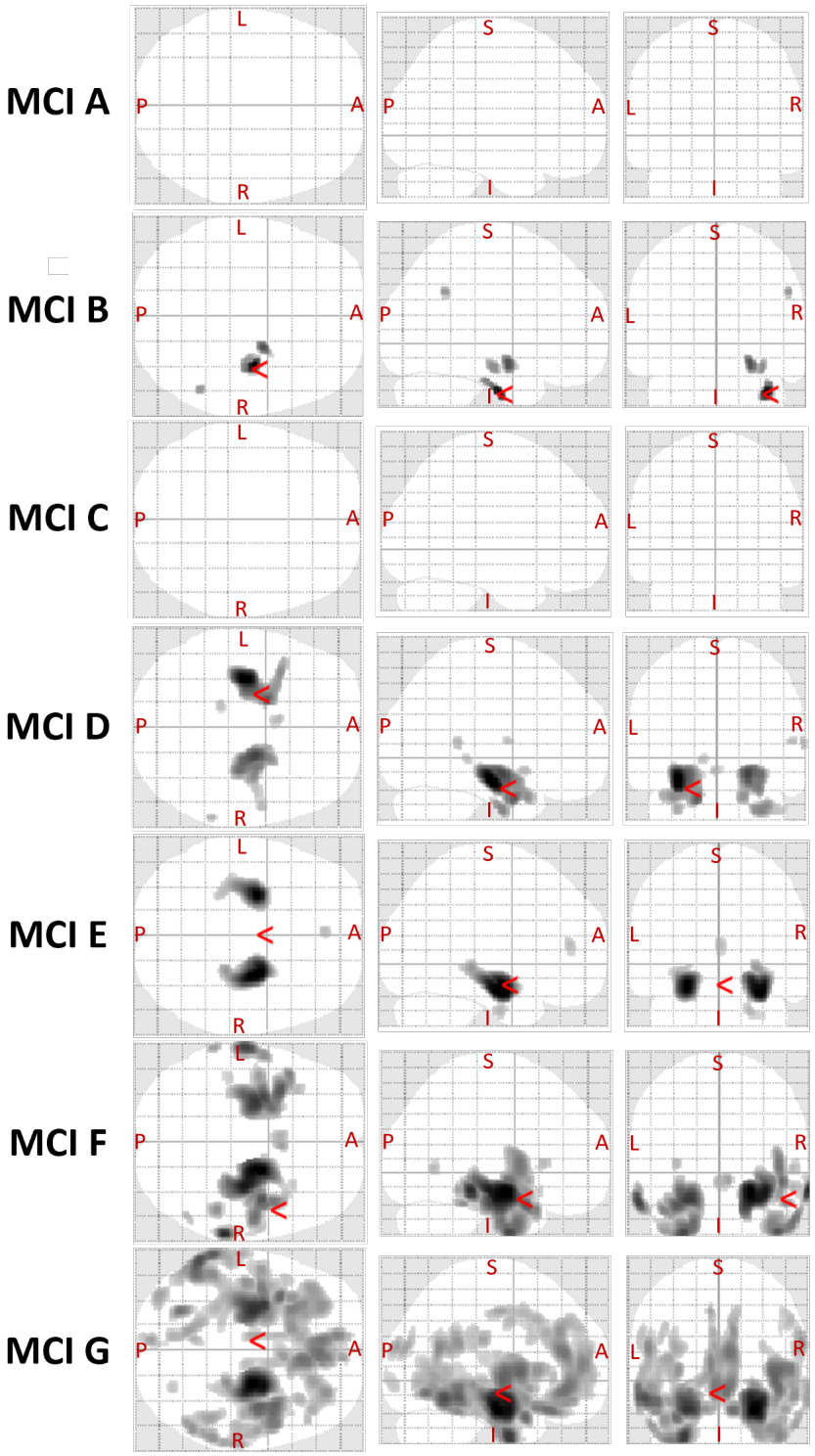

Figure 4-4: Test statistic maps for the MCI subgroups relative to CN group.
Note: L, Left; R, Right; A, Anterior; P, Posterior; S, Superior, and I, Inferior.

Table 4-15: Percent per region by hemisphere (Right/left) with significant differences when comparing each MCI subgroup against CN.

Regions	B	D	E	F	G
Subcortical gray nuclei					
Nucleus accumbens	-	-	-	4.1/2.8	9.8/-
Limbic lobe					
Amygdala	3.7/-	9.5/16.0	11.5/14.4	12.6/18.5	19.7/21.5
Anterior cingulate & paracingulate gyri	-	-	-	-	12.7/13.2
Hippocampus	1.8/-	11.9/14.6	16.9/14.2	18.5/15.3	17.8/19.5
Middle cingulate & paracingulate gyri	-	-	-	-	2.5/2.0
Parahippocampal gyrus	-	3.1/4.8	2.5/-	8.2/8.4	12.9/10.9
Temporal pole: middle	-	-	-	-/3.5	-
Temporal pole: superior	-	-/1.9	-	1.5/5.7	2.3/2.7
Insula					
Insula	-	-	-	5.6/-	12.3/6.0
Central region					
Rolandic operculum	-	-	-	-	5.6/4.2
Frontal Lobe					
Anterior orbital gyrus	-	-	-	-	4.8/7.3
Gyrus rectus	-	-	-	-	9.5/5.3
Lateral orbital gyrus	-	-	-	-	-/3.7
Medial orbital gyrus	-	-	-	-	5.4/4.8
Olfactory cortex	-	-	-	1.7/1.7	4.1/4.2
Posterior orbital gyrus	-	-	-	-	3.6/14.7
Superior frontal gyrus, medial	-	-	-	-	2.4/2.3
Superior frontal gyrus, medial orbital	-	-	-	-	10.2/11.9
Temporal Lobe					
Heschl's gyrus	-	-	-	-	4.6/5.0
Inferior temporal gyrus	-	-	-	2.7/2.3	4.1/7.8
Middle temporal gyrus	-	-	-	1.4/-	2.4/4.0
Superior temporal gyrus	-	-	-	1.0/-	3.3/-
Parietal Lobe					
SupraMarginal gyrus	-	-	-	-	2.2/2.1
Occipital Lobe					
Calcarine fissure	-	-	-	-	-/1.3
Cuneus	-	-	-	-	-/1.7
Fusiform gyrus	-	-	-	1.5/2.8	3.3/6.2
Inferior occipital gyrus	-	-	-	-	-/2.1
Lingual gyrus	-	-	-	-	-/4.5
Middle occipital gyrus	-	-	-	-	-/1.6

5 Discussion

We present a quantitative characterization of AD in its early stages using cross-sectional information from cognitive, functional, and behavioral tests. Our approach identifies abnormal items and employs a relational metric to group these items, establishing five possible neuropsychological dimensions. For each neuropsychological dimension, we calculated a quantitative index. To evaluate their usefulness, we used a vector of five indices to predict the progression from MCI to AD dementia within the next 36 months, 48 months, and 60 months.

The five neuropsychological dimensions group together abnormal neuropsychological items from different domains. There was one predominant domain, accompanied by other variables corresponding to test tasks designed to evaluate other domains. This finding could be explained by the possible inherent dependencies between cognitive, functional, and behavioral domains. Neuropsychological tests usually contain tasks to evaluate one specific domain; however, performance on a task does not strictly depend on a single domain.

The identified neuropsychological dimensions may be capturing impairments in more complex abilities that require a combination of cognitive, functional, and behavioral skills. The first dimension indicates that low performance in constructional praxis is associated with memory and attention failures. The second dimension reflects the association between performance in activities of daily life, memory, and orientation. The last three dimensions indicate language abilities; however, one of them evaluates more global functions of comprehension and expressive language, another demonstrates specific performance in certain tasks, and the last one describes a more complex combination of language in association with a greater number of domains.

It should be noted that in the early stages of cognitive impairment, in patients with MCI, it is possible to predict progression to AD dementia. The performance of the automated prediction of disease progression demonstrates the usefulness of the proposed characterization strategy for disease prognosis. Furthermore, it relies only on neuropsychological data that can be easily and inexpensively acquired, making it a practical tool for patient monitoring.

In addition, this study identified seven MCI subgroups, each displaying heterogeneous neuropsychological profiles and distinct patterns of GM atrophy. We opted for the term 'subgroup' instead of 'subtype' to describe these groups of MCI subjects. This choice was made

due to the uncertainty about whether the observed heterogeneity arises from different stages of the condition, represents genuinely distinct disease phenotypes, or is a combination of both.

The dimension of orientation, memory, and activities of daily living appears impaired in all MCI subgroups, indicating its characteristic nature within this condition. This finding aligns with the fact that these items are associated with the diagnostic criteria for MCI [33], as it encompasses items from functional performance assessments, as well as from tests such as the Mini-Mental State Examination, Clinical Dementia Rating, and the Logical Memory II subscale of the Wechsler Memory Scale–Revised.

Subgroups A and C did not show significant anatomical differences, however, they did present cognitive variations when compared to the CN group, with subgroup C showing a greater divergence than subgroup A. This finding suggests the presence of neuropsychological impairments not necessarily linked to detectable cerebral abnormalities. In contrast, subgroups F and G demonstrated a widespread pattern of cortical atrophy, aligning with their neuropsychological profile, which is characterized by impaired performance across all dimensions.

The observed relation between the extent of GM atrophy and the neuropsychological deficits in MCI provides evidence for the close co-occurrence of structural changes and cognitive impairments. Our findings suggest a progressive pattern of increasing severity across the MCI subgroups from A to G, with subgroup A showing the least impairment and subgroup G displaying the greatest level of compromise. The extent of GM atrophy begins in limbic system structures and subsequently extends to other regions of the brain.

An important finding in the results is the similarity in cognitive profiles and brain atrophy regions between subgroups D and E. However, these subgroups consist of individuals with different disease progressions. Subgroup D consists of stable patients, while Subgroup E comprises those progressing to dementia within 48 months. The differentiating factor seems to be their performance in orientation, memory, and daily living tasks, with Subgroup E exhibiting poorer results.

A possible limitation of our methodology is that the cut-offs for abnormality rely on the assumption that standardized z-scores have a normal distribution. A more detailed analysis would be required to evaluate how slight variations in the cut-offs could impact the presented results. The strengths of the study include the consideration of a wide range of cognitive and functional tasks, enabling the comprehensive characterization of patterns of neuropsychological deficits. To avoid assessing multiple domains with total scores, an approach with subscores was adopted. The implementation of meticulous quality control measures in MRI image processing ensures reliable results. Additionally, the analysis involved a comparison

of whole-brain gray matter tissue, allowing for the capture of brain differences without restrictions on specific regions of interest.

6 Conclusions

In summary, our approach successfully learns relationships between test items based on neuropsychological performance within the disease, identifying neuropsychological dimensions with which it is possible to quantify neuropsychologically MCI subjects and predict their progression to AD dementia. The quantitative characterization demonstrates a predictive accuracy of 0.76 for progression from MCI to dementia within the next 36 months. Additionally, we identified seven distinct subgroups of MCI, each exhibiting unique neuropsychological profiles and patterns of brain atrophy. Notably, the impairment of orientation, memory, and activities of daily living was consistently observed across all subgroups. Our findings reveal a significant relationship between cognitive impairment and brain atrophy, with an increasing extent of gray matter atrophy corresponding to the severity of cognitive and behavioral abnormalities. This gradient of severity underscores the heterogeneity within the MCI population and emphasizes the importance of considering different clinical presentations when evaluating and treating patients with MCI.

6.1. Academic products

The contributions and results of this work were presented at two international conferences:

- *"Mining relations between neuropsychological data to characterize Alzheimer's disease"* in Proceedings of the 17th International Seminar on Medical Information Processing and Analysis (SIPAIM), Campinas (Brazil), November 2021. Accepted and published [39].
- *"Exploring anatomical brain heterogeneity in subjects with mild cognitive impairment"* in Proceedings of the 19th International Seminar on Medical Information Processing and Analysis (SIPAIM), Mexico City (Mexico), November 2023. Accepted and published [40].

6.2. Future Work

Future work could explore the relationship between the quantitative neuropsychological characterization of the disease and known biomarkers of AD pathology. Additionally, the present study did not delve into statistical differences between the resulting subgroups, which could

be a focus of future research. Similarly, longitudinal patterns of atrophy and neuropsychological profiles within these MCI subgroups could be analyzed. Understanding the trajectory of cognitive decline and its association with patterns of atrophy can provide valuable insights into the progression of MCI and its underlying mechanisms.

References

[1] Martina Zvěřová. Alzheimer's disease and blood-based biomarkers–potential contexts of use. *Neuropsychiatric disease and treatment*, 14:1877–1882, 2018.

[2] Alzheimer's Association et al. 2020 alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 16(3):391+, 2020.

[3] Alireza Atri. The alzheimers disease clinical spectrum: diagnosis and management. *Medical Clinics*, 103(2):263–293, 2019.

[4] Colin L. Masters, Randall Bateman, Kaj Blennow, Christopher C. Rowe, Reisa A. Sperling, and Jeffrey L. Cummings. Alzheimer's disease. *Nature Reviews Disease Primers*, 1:15056, 2015.

[5] Maria Ferretti, M Florencia Iulita, Enrica Cavedo, Patrizia Chiesa, Annemarie Schumacher Dimech, Antonella Chadha, Francesca Baracchi, Hélène Girouard, Sabina Misoch, Ezio Giacobini, Herman Depypere, and Harald Hampel. Sex differences in alzheimer disease — the gateway to precision medicine. *Nature Reviews Neurology*, 14, 07 2018.

[6] Martin James Prince, Anders Wimo, Maelenn Mari Guerchet, Gemma Claire Ali, Yu-Tzu Wu, and Matthew Prina. *World Alzheimer Report 2015-The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends.* Alzheimer's Disease International, 2015.

[7] Zhang Daoqiang, Shen Dinggang, and The Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in alzheimer's disease. *NeuroImage*, 59(2):895–907, 2012.

[8] Serra Angela, Galdi Paola, and Tagliaferri Roberto. Multi-view learning in biomedical applications. In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pages 265–280. Elsevier, 2019.

[9] Oskar Hansson, John Seibyl, Erik Stomrud, Henrik Zetterberg, John Q. Trojanowski, Tobias Bittner, Valeria Lifke, Veronika Corradini, Udo Eichenlaub, Richard Batrla, Katharina Buck, Katharina Zink, Christina Rabe, Kaj Blennow, and Leslie M. Shaw. Csf biomarkers of alzheimer's disease concord with amyloid-β pet and predict clinical progression: A study of fully automated immunoassays in biofinder and adni cohorts. *Alzheimer's & Dementia*, 14(11):1470–1481, 2018.

[10] Joey Annette Contreras, Vahan Aslanyan, Melanie D. Sweeney, Lianne M.J. Sanders, Abhay P. Sagare, Berislav V. Zlokovic, Arthur W. Toga, S. Duke Han, John C. Morris, Anne Fagan, Parinaz Massoumzadeh, Tammie L. Benzinger, and Judy Pa. Functional connectivity among brain regions affected in alzheimer's disease is associated with csf tnf- α in apoe4 carriers. *Neurobiology of Aging*, 86:112–122, 2020.

[11] Christian Humpel. Identifying and validating biomarkers for alzheimer's disease. *Trends in biotechnology*, 29(1):26–32, 2011.

[12] Douglas W Scharre. Preclinical, prodromal, and dementia stages of alzheimer's disease. *Pract Neurol*, pages 36–47, 2019.

[13] Berndt Winblad, Katie Palmer, Miia Kivipelto, Vesna Jelic, Laura Fratiglioni, L-O Wahlund, Agneta Nordberg, Lars Bäckman, Michael Albert, Ove Almkvist, et al. Mild cognitive impairment–beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. *Journal of internal medicine*, 256(3):240–246, 2004.

[14] Reisa A Sperling, Jason Karlawish, and Keith A Johnson. Preclinical alzheimer disease—the challenges ahead. *Nature Reviews Neurology*, 9(1):54–58, 2013.

[15] Emily C. Edmonds, Carrie R. McDonald, Anisa Marshall, Kelsey R. Thomas, Joel Eppig, Alexandra J. Weigand, Lisa Delano-Wood, Douglas R. Galasko, David P. Salmon, and Mark W. Bondi. Early versus late mci: Improved mci staging using a neuropsychological approach. *Alzheimer's & Dementia*, 15(5):699–708, 2019.

[16] Ronald C Petersen, Glenn E Smith, Stephen C Waring, Robert J Ivnik, Eric G Tangalos, and Emre Kokmen. Mild cognitive impairment: clinical characterization and outcome. *Archives of neurology*, 56(3):303–308, 1999.

[17] Joel S. Eppig, Emily C. Edmonds, Laura Campbell, Mark Sanderson-Cimino, Lisa Delano-Wood, and Mark W. Bondi. Statistically derived subtypes and associations with cerebrospinal fluid and genetic biomarkers in mild cognitive impairment: A latent profile analysis. *Journal of the International Neuropsychological Society*, 23(7):564–576, 2017.

[18] Nanbo Sun, Elizabeth C. Mormino, Jianzhong Chen, Mert R. Sabuncu, and B.T. Thomas Yeo. Multi-modal latent factor exploration of atrophy, cognitive and tau heterogeneity in alzheimer's disease. *NeuroImage*, 201:116043, 2019.

[19] Charles DeCarli. Mild cognitive impairment: prevalence, prognosis, aetiology, and treatment. *The Lancet Neurology*, 2(1):15–21, 2003.

[20] Ronald C Petersen. Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, 256(3):183–194, 2004.

[21] Devendra Goyal, Donna Tjandra, Raymond Q Migrino, Bruno Giordani, Zeeshan Syed, Jenna Wiens, Alzheimer's Disease Neuroimaging Initiative, et al. Characterizing heterogeneity in the progression of alzheimer's disease using longitudinal clinical and neuroimaging biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10:629–637, 2018.

[22] Ronald C Petersen, Rosebud O Roberts, David S Knopman, Bradley F Boeve, Yonas E Geda, Robert J Ivnik, Glenn E Smith, and Clifford R Jack. Mild cognitive impairment: ten years later. *Archives of neurology*, 66(12):1447–1455, 2009.

[23] Emily C Edmonds, Joel Eppig, Mark W Bondi, Kelly M Leyden, Bailey Goodwin, Lisa Delano-Wood, Carrie R McDonald, Alzheimer's Disease Neuroimaging Initiative, et al. Heterogeneous cortical atrophy patterns in mci not captured by conventional diagnostic criteria. *Neurology*, 87(20):2108–2116, 2016.

[24] Emily C Edmonds, Lisa Delano-Wood, Lindsay R Clark, Amy J Jak, Daniel A Nation, Carrie R McDonald, David J Libon, Rhoda Au, Douglas Galasko, David P Salmon, et al. Susceptibility of the conventional criteria for mild cognitive impairment to false-positive diagnostic errors. *Alzheimer's & Dementia*, 11(4):415–424, 2015.

[25] Tiffany F Hughes, Beth E Snitz, and Mary Ganguli. Should mild cognitive impairment be subtyped? *Current opinion in psychiatry*, 24(3):237–242, 2011.

[26] Lisa Delano-Wood, Mark W Bondi, Joshua Sacco, Norm Abeles, Amy J Jak, David J Libon, and Andrea Bozoki. Heterogeneity in mild cognitive impairment: Differences in neuropsychological profile and associated white matter lesion pathology. *Journal of the International Neuropsychological Society*, 15(6):906–914, 2009.

[27] David J Libon, Sharon X Xie, Joel Eppig, Graham Wicas, Melissa Lamar, Carol Lippa, Brianne M Bettcher, Catherine C Price, Tania Giovannetti, Rod Swenson, et al. The heterogeneity of mild cognitive impairment: A neuropsychological analysis. *Journal of the International Neuropsychological Society*, 16(1):84–93, 2010.

[28] Mary M Machulda, Emily S Lundt, Sabrina M Albertson, Walter K Kremers, Michelle M Mielke, David S Knopman, Mark W Bondi, and Ronald C Petersen. Neuropsychological subtypes of incident mild cognitive impairment in the mayo clinic study of aging. *Alzheimer's & Dementia*, 15(7):878–887, 2019.

[29] Diana L Giraldo, Jan Sijbers, Eduardo Romero, and Alzheimer's Disease Neuroimaging Initiative. Quantification of cognitive impairment to characterize heterogeneity of

patients at risk of developing alzheimer's disease dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 13(1):e12237, 2021.

[30] Mary M Machulda, Emily S Lundt, Sabrina M Albertson, Anthony J Spychalla, Christopher G Schwarz, Michelle M Mielke, Clifford R Jack Jr, Walter K Kremers, Prashanthi Vemuri, David S Knopman, et al. Cortical atrophy patterns of incident mci subtypes in the mayo clinic study of aging. *Alzheimer's & Dementia*, 16(7):1013–1022, 2020.

[31] John Ashburner and Karl J Friston. Voxel-based morphometry—the methods. *Neuroimage*, 11(6):805–821, 2000.

[32] John Ashburner and Karl J Friston. Why voxel-based morphometry should be used. *Neuroimage*, 14(6):1238–1243, 2001.

[33] R. C. Petersen, P. S. Aisen, L. A. Beckett, M. C. Donohue, A. C. Gamst, D. J. Harvey, C. R. Jack, W. J. Jagust, L. M. Shaw, A. W. Toga, J. Q. Trojanowski, and M. W. Weiner. Alzheimer's disease neuroimaging initiative (adni) clinical characterization. *Neurology*, 74(3):201–209, 2010.

[34] Maura Mitrushina, Kyle B Boone, Jill Razani, and Louis F D'Elia. *Handbook of normative data for neuropsychological assessment*. Oxford University Press, 2005.

[35] Joe Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963.

[36] Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[37] J. Ashburner and K.J. Friston. Voxel based morphometry. In Larry R. Squire, editor, *Encyclopedia of Neuroscience*, pages 471–477. Academic Press, Oxford, 2009.

[38] Edmund T Rolls, Chu-Chung Huang, Ching-Po Lin, Jianfeng Feng, and Marc Joliot. Automated anatomical labelling atlas 3. *Neuroimage*, 206:116189, 2020.

[39] German A Pabón, Diana L Giraldo, and Eduardo Romero. Mining relations between neuropsychological data to characterize alzheimer's disease. In *17th International Symposium on Medical Information Processing and Analysis*, volume 12088, pages 190–197. SPIE, 2021.

[40] German A. Pabón Ochoa, Diana L. Giraldo, and Eduardo Romero. Exploring anatomical brain heterogeneity in subjects with mild cognitive impairment. In *2023 19th International Symposium on Medical Information Processing and Analysis (SIPAIM)*, pages 1–5, 2023.