Journal Pre-proof

Autosomal Dominantly Inherited Alzheimer Disease: Analysis of genetic subgroups by Machine Learning

Diego Castillo-Barnes, Li Su, Javier Ramírez, Diego Salas-Gonzalez, Francisco J. Martinez-Murcia, Ignacio A. Illan, Fermin Segovia, Andres Ortiz, Carlos Cruchaga, Martin R. Farlow, Chengjie Xiong, Neil R. Graff-Radford, Peter R. Schofield, Colin L. Masters, Stephen Salloway, Mathias Jucker, Hiroshi Mori, Johannes Levin, Juan M. Gorriz, Dominantly Inherited Alzheimer Network (DIAN)

PII: S1566-2535(19)30995-9
DOI: https://doi.org/10.1016/j.inffus.2020.01.001
Reference: INFFUS 1190

To appear in: *Information Fusion*

Received date: 12 December 2019
Revised date: 31 December 2019
Accepted date: 3 January 2020

Please cite this article as: Diego Castillo-Barnes, Li Su, Javier Ramírez, Diego Salas-Gonzalez, Francisco J. Martinez-Murcia, Ignacio A. Illan, Fermin Segovia, Andres Ortiz, Carlos Cruchaga, Martin R. Farlow, Chengjie Xiong, Neil R. Graff-Radford, Peter R. Schofield, Colin L. Masters, Stephen Salloway, Mathias Jucker, Hiroshi Mori, Johannes Levin, Juan M. Gorriz, Dominantly Inherited Alzheimer Network (DIAN), Autosomal Dominantly Inherited Alzheimer Disease: Analysis of genetic subgroups by Machine Learning, *Information Fusion* (2020), doi: https://doi.org/10.1016/j.inffus.2020.01.001

- A comparison in DIAN by considering the 3 genes and Machine Learning.
- Feature selection based on ANOVA followed by Principal Component Analysis (PCA)
- SVM in a nested k-Fold CV resulted in accuracies of 72-74% using PiB PET features
- PSEN1 subgroups vs. NC provided accuracies of 80%, DIAN as an heterogeneous entity

Autosomal Dominantly Inherited Alzheimer Disease: Analysis of genetic subgroups by Machine Learning

Diego Castillo-Barnes[1], Li Su[2], Javier Ramírez[1], Diego Salas-Gonzalez[1], Francisco J. Martinez-Murcia[3], Ignacio A. Illan[1], Fermin Segovia[1], Andres Ortiz[3], Carlos Cruchaga[4], Martin R. Farlow[5], Chengjie Xiong[6], Neil R. Graff-Radford[7], Peter R. Schofield[8], Colin L. Masters[9], Stephen Salloway[10], Mathias Jucker[11], Hiroshi Mori[12], Johannes Levin[13], Juan M. Gorriz[1,2,*], Dominantly Inherited Alzheimer Network (DIAN)[14]

[1] *Department of Signal Theory, Telematics and Communications, University of Granada, Granada (Spain)*

[2] *Department of Psychiatry, University of Cambridge, Cambridge (UK)*

[3] *Department of Communications Engineering, University of Malaga, Malaga (Spain)*

[4] *Department of Psychiatry and Neurology, Washington University School of Medicine, St. Louis, Missouri (USA)*

[5] *Department of Neurology, Indiana University School of Medicine, Indianapolis, Indiana (USA)*

[6] *Division of Biostatistics, Washington University School of Medicine, St. Louis, Missouri (USA)*

[7] *Department of Neurology, Mayo Clinic, Jacksonville, Florida (USA)*

[8] *Neuroscience Research Australia and School of Medical Sciences, University of New South Wales, Sydney (Australia)*

[9] *Florey Institute and University of Melbourne, Victoria (Australia)*

[10] *Butler Hospital, Rhode Island (USA)*

[11] *Department of Cellular Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen (Germany)*

[12] *Department of Clinical Neuroscience, Osaka City University Medical school, Osaka (Japan)*

[13] *Department of Neurology, Ludwig-Maximilians-University of Munich, Munich (Germany).*

[14] *A complete list of DIAN-affiliated faculty and staff can be found at: www.DIAN-info.org/personnel.htm*

Abstract

Despite subjects with Dominantly-Inherited Alzheimer's Disease (DIAD) rep-

*Corresponding author: Juan M. Gorriz. Email: gorriz@ugr.es

resent less than 1% of all Alzheimer's Disease (AD) cases, the Dominantly Inherited Alzheimer Network (DIAN) initiative constitutes a strong impact in the understanding of AD disease course with special emphasis on the presyptomatic disease phase. Until now, the 3 genes involved in DIAD pathogenesis (PSEN1, PSEN2 and APP) have been commonly merged into one group (Mutation Carriers, MC) and studied using conventional statistical analysis. Comparisons between groups using null-hypothesis testing or longitudinal regression procedures, such as the linear-mixed-effects models, have been assessed in the extant literature.

Within this context, the work presented here performs a comparison between different groups of subjects by considering the 3 genes, either jointly or separately, and using tools based on Machine Learning (ML). This involves a feature selection step which makes use of ANOVA followed by Principal Component Analysis (PCA) to determine which features would be realiable for further comparison purposes. Then, the selected predictors are classified using a Support-Vector-Machine (SVM) in a nested k-Fold cross-validation resulting in maximum classification rates of 72-74% using PiB PET features, specially when comparing asymptomatic Non-Carriers (NC) subjects with asymptomatic PSEN1 Mutation-Carriers (PSEN1-MC). Results obtained from these experiments led to the idea that PSEN1-MC might be considered as a mixture of two different subgroups including: a first group whose patterns were very close to NC subjects, and a second group much more different in terms of imaging patterns. Thus, using a k-Means clustering algorithm it was determined both subgroups and a new classification scenario was conducted to validate this process. The comparison between each subgroup *vs.* NC subjects resulted in classification rates around 80% underscoring the importance of considering DIAN as an heterogeneous entity.

Keywords: Dominantly-Inherited Alzheimer's Disease (DIAD), DIAN, Alzheimer's Disease (AD), Neuroimaging, Machine Learning

1. Introduction

Alzheimer's Disease (AD) is neuropathologically defined by the presence of amyloid-β (Aβ)-plaques and by neurofibrillary tangles associated with a suggestive clinical phenotype [1, 2, 3]. Clinically AD is characterized by a progressive loss of memory and other neuropsychiatric changes such as decline in executive functioning and behavioral changes [4, 5].

Since the development of a theoretical model of biomarker changes for AD [6], multiple longitudinal studies about AD have tried to find the exact triggers that could explain the prognosis and evolution of the disease. Clinicopathologic evidence suggests that pathological changes leading to AD such as deposition of Aβ-plaques begin many years prior to onset of cognitive symptoms [7, 8, 9, 10], but it still awaits for further empirical validation. In addition to this, as some more recent works point out, the nature of AD might be mistakenly described until now as different genetic alterations, which are causing the same disease, are expressing themselves through different triggers [11, 12, 13, 14, 3, 15].

Dominantly Inherited Alzheimers Disease (DIAD) only represent about 1% of all AD cases, but it has a marked importance for AD research [16]. This type of AD is caused by known mutations in the Amyloid Precursor Protein (APP) [17], Presenilin-1 (PSEN1) [18, 3] (most frequently found), or Presenilin-2 (PSEN2) [19] genes. DIAD is quite similar to the more common Late Onset AD (LOAD) in many features including clinical presentation and disease course [20, 21, 22, 23, 3, 24]. In this sense, the main difference between DIAD and LOAD is in the age at onset, family history and co-pathologies [25].

To facilitate the study of DIAD and its comparison with LOAD, the National Institute on Aging (NIA) funded the Dominantly Inherited Alzheimer Network (DIAN) [26]. This study assesses people at risk of inheriting an autosomal dominant AD mutation and monitor their evolution through a standardized procedure which includes clinical, cognitive, neuroimaging, CSF and plasma tests among others.

Currently, the DIAN study presents more than 450 subjects with 90 different

mutations in multiple genes[1].

The reason why DIAN is so important for AD research is related to the understanding of the disease natural history. It is certain that a person with DIAD will develop AD in the future, so a follow-up of different biomarkers from DIAD subjects (specially during their initial stages as asymptomatic carriers) could provide a great deal of information which may later be used to make a model of the disease [27, 28]. Findings obtained from DIAD can be extrapolated to the sporadic LOAD so it is expected that better treatments will be prescribed at pre-clinical stages of the disease when damage is minimal and pathogenesis can be slowed down or even prevented from progressing [7, 29, 30, 31, 24, 5].

DIAD research has played an important role in the fact that it is now well-accepted that AD pathogenesis starts 20-30 years before the onset of clinical symptoms [5]. Previous works dedicated to decipher the time course of DIAD have made use of regression models to explore the evolution of AD markers based on a predicted variable called AAOE (Age At Onset Estimated) and its equivalent EYO (Estimated Years to Onset) which is predicted considering the disease onset age of their first-degree relatives [32, 22] or the average age of onset for each mutation type [25].

Over the last decades, neuroscience has transitioned from qualitative reports of case studies from abnormal conditions to quantitative statistical maps to characterize the global pattern of disease. With the inclusion of classical statistics, pathological models were obtained based on classical statistical assumptions on data, which are not always fulfilled [33], making interpretation of the results arguable. Fortunately, the recent use of Machine Learning (ML) techniques in the analysis of different neurological disorders [34] is having a noticeable impact on diagnosis and prognosis of diseases, such as Parkinson's Disease (PD) or AD, even though their complex pathological models are not yet completely understood in limited sample sizes. This conforms a potentially useful screening tool in the diagnosis of unseen subjects or in the development

[1] Or mutation combinations like Exons duplications for example.

of new treatments. The latter, combined with the emergence of highly detailed databases like DIAN, might result in an accurate model of AD disease courses, despite the definition of AD as a mixture of different AD subtypes [35].

With this aim, the work presented here proposes a DIAD analysis based on ML techniques to discern between Non-Carriers (NC) and Mutation-Carriers (MC) participants grouped by which gene is causing (or will cause) the disease. This give us 3 main hypothesis including 1) to check whether all data modalities are necessary for DIAD diagnosis including: Non-Imaging Biomarkers (**NIB**) such as $A\beta_{4(0,2)}$ amyloid, τ protein or the apolipoprotein APOE among others; and imaging features extracted from PiB PET (**PIB**), FDG PET (**FDG**) and MRI scans (**MRI**); 2) to validate if reducing the MC heterogeneity via separating the mutations by their responsible gene has a significative impact on the DIAD diagnosis accuracy; and 3) to determine if the model generated using ML can be used to determine in which period before the symptoms onset group differences are larger.

Final results presented a high performance in classification rates when using PET features with subgrouping MC gene expression even though there may not yet be cognitive symptoms (EYO greater than 5 and 10 years), as well as when clustering asymptomatic PSEN1 MC into two groups: a first group quite similar to the NC group and a second group totally different. The application of a ML approach to these new subsets has resulted in a novel model of the disease and it is expected to have a great impact on DIAD knowledge.

2. Material & Methods

2.1. Participants

DIAN project was founded by Washington University School of Medicine in 2008. This study has been approved by the local institutional review boards of each participating site from United States, Canada, France, Spain, United Kingdom and Australia among other places. Full details of participating sites, enrollment, assessment protocol and infrastructure of DIAN were published in

[26]. The inclusion criteria list includes statements such as be aged 18 (inclusive) or older and the child of an affected individual (clinically or by testing) in a pedigree with a known mutation for DIAD; to have two persons who are not their full-blooded siblings who can serve as collateral sources for the study; or be fluent in a language approved by the DIAN Coordinating Center.

In this work, we have made use of data results from the DIAN Observational Study Data Freeze 11 but excluding previously those participants with least one of the following diagnosed diseases: cerebral stroke (3 subjects), transient ischemic attack (1 subject), dementia by alcoholism (4 subjects), Parkinson's disease (1 subject), traumatic brain injury with chronic deficit/dysfunction (3 subjects), dementia with Lewy bodies, vascular dementia, dementia by unknown causes (3 subjects), frontotemporal dementia, primary progressive aphasia, progressive nonfluent aphasia, semantic dementia, other types of dementia (e.g. logopenic, anomic or transcortical), progressive supranuclear palsy, corticobasal degeneration, Huntington's disease, prionic dementia, Down syndrome, hydrocephalus and central nervous system neoplasm. Besides, in order not to increase the heterogeneity in symptomatic subjects, LOAD cases in DIAN study have been also discarded (13 subjects). As a result, our average dataset consists of a total of 442 subjects (184 males, 41.63%) with an age at baseline of 38.71 $\pm$ 10.98 years[2]. Other disorders such as B12 vitamin deficit, depresion, alcoholism, abuse of other substances, seizures and traumatic brain injury without chronic deficit/dysfunction or thyroid problems have been included due to the large amount of participants with at least one of those diagnostics.

2.2. Genetic groups

Mutations in the APP, PSEN1 and PSEN2 genes were identified in DIAN from DNA extracted from peripheral blood samples as described in [36]. Among all the included participants, 265 were DIAD Mutation Carriers (MC): 202 PSEN1, 22 PSEN2 and 43 APP; while 173 were Non-Carriers (NC) subjects.

[2] Given in terms of mean and standard deviation.

Marker	Type of test	Assay protocols [9]
$A\beta_{42}$	CSF	INNO, xMAP
$A\beta_{40}$	CSF	INNO
$A\beta_{42}$: $A\beta_{40}$ ratio	CSF	INNO
τ	CSF	xMAP
p-τ	CSF	xMAP
$A\beta_{42}$	Plasma	PL_xMAP
$A\beta_{40}$	Plasma	PL_xMAP
$A\beta_{42}$: $A\beta_{40}$ ratio	Plasma	PL_xMAP
APOE	Genetic	Alleles $\epsilon_{i,j}$ with i,j=2,3,4

Table 1: Non-imaging markers.

The 2 remaining subjects were labelled as Unknown and they were not included in the analysis.

2.3. Non-imaging markers

In the models proposed in [20, 8], fibrillar amyloid-β (Aβ) depositions play a key role in the development of AD [37]. Once amyloid Aβ depositions begin to accumulate, many other biomarkers such as CSF τ protein also becomes abnormal. This stage is followed by a brain atrophy (measured by volumetric MRI tests) and cognitive symptoms.

As mentioned in section 1, several works suggest that there are differences in pathogenesis between DIAD and LOAD [38, 39]. Following recent works [40], and based on availability, we decided to include in this study markers such as $A\beta_{42}$, $A\beta_{40}$, $A\beta_{42}$:$A\beta_{40}$ ratio, τ and p-τ as non-imaging biomarkers (**NIB**) available ro most of our subjects (table 1).

Detailed information about data preparation is available from: https://dian.wustl.edu/our-research/observational-study/.

2.4. Imaging markers

DIAN dataset includes a wide range of imaging variables from MRI and PET images for each participant. Two classes of PET imaging are available for Data Freeze 11: **FDG** (PET imaging using Fluorodeoxyglucose F18 for imaging of

the brain metabolism) and **PiB** (PET imaging using Pittsburgh Compound B for imaging of amyloid depositions). This section gives some details about data acquisition and complementary tools.

2.4.1. MRI acquisition

MRI acquisition was carried out following the Alzheimer Disease Neuroimaging Initiative (ADNI) protocol [6]. Each participant received an accelerated 3D sagittal T1-weighted MPRAGE on a 3T scanner. Resulting images presented a voxel size of $1.1 \times 1.1 \times 1.2$ mm and an adquisition time of approximately $5 - 6$ minutes. All scans have been acquired with a Siemens BioGraph mMR PET-MR 3T scanner or a Siemens Trio 3T MRI scanner depending on the center where tests have been obtained.

All MRI sessions have been processed using the `FreeSurfer, v5.3` analysis suite including all patches, for cortical reconstruction and volumetric segmentation [41]. The technical details of these procedures are described in prior publications [42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53]. More information is available on the website (http://surfer.nmr.mgh.harvard.edu/).

The MRI preprocessing pipeline includes: motion correction and segmentation of the subcortical white matter and deep gray matter volumetric structures on T1 combined with T2 images [48]; intensity normalization; registration to a spherical atlas which utilizes individual cortical folding patterns to match cortical geometry across subjects [45]; and parcellation of the cerebral cortex into units based on gyral and sulcal structure [54]. For each vertex on the cortical surface, thickness was calculated as the shortest distance from the gray/white boundary to the gray/csf boundary [55].

2.4.2. FDG and PiB acquisition

Amyloid imaging was performed with a bolus injection of approximately 15 mCi of [11C]PiB. Dynamic imaging acquisition started either at injection for 70 minutes or 40 minutes post-injection for 30 minutes. For analysis, the PiB PET data between 40 to 70 minutes were used. Metabolic [18F]FDG-PET imaging

that assess the neurodegenerative processess in AD [56, 57, 58], was performed with a 3D dynamic acquisition. It began 40 minutes after a bolus injection of approximately 5 mCi of FDG and lasted for 20 minutes.

PET images were motion corrected and registered to their MRI using the methods described in [59, 60]. For each region-of-interest[3], standardized uptake value ratio (SUVR) was calculated using a cerebellar reference.

Dynamic PET scans are divided into early, middle, and late frames and registered to correct for head motion. The T1 weighted MRI scans are registered to the Talairach atlas and to the PET images, and transformed into an atlas space. Since PET imaging has a spatial resolution around 6 mm, the regional activity measured directly from PET is a linear combination of activity from different regions. This phenomenon is known as the partial volume effect [63, 64] so, it was compensated through the use of a regional spread function for partial volume correction in the analysis [65].

2.5. Summary of input variables

Considering both imaging and non-imaging variables available in DIAN database, a final set of 1755 features were considered in this work including 10 **NIB** tests and 1745 imaging features (521 **FDG**, 184 **MRI** and 1040 **PiB**).

2.6. Symptomatology

Separability between symptomatic and asymptomatic subjects have been carried out taking into account the Clinical Dementia Rating (CDR) scale [66, 67]. While asymptomatic or preclinical subjects are those whose CDR results are equal to 0, sympomatic subjects can be divided according to CDR scale into early symptomatic (CDR = 0.5), mild dementia (CDR = 1), moderate dementia (CDR = 2) and severe dementia (CDR = 3) [68]. Table 2 gives a description of DIAN participants as a function of their baseline CDR results.

[3]Regions obtained using `FreeSurfer, v5.3` following the same protocols as in ADNI cohort study [61, 62].

	Asymptomatic	Symptomatic			
		Early	Mild	Moderate	Severe
NC	173	0	0	0	0
APP	29	9	4	1	0
PSEN1	122	51	20	5	4
PSEN2	20	2	0	0	0

Table 2: CDR in baseline.

2.7. Experiments Definition

To cover all the questions proposed for this work, three experiments have been described as follows:

2.7.1. Experiment 1 - Between-group discrimination ability by data type

Until now, most of the works proposed for DIAN database analysis were focused on the separability between MC and NC subjects using conventional statistical tests like t-Test. As mentioned in section 1, although some of these works have pointed out the relevance of considering the MC separately according to their mutation, we have decided to include a direct comparison between MC and NC as our reference baseline case to improve. It is expected that resulting conclusions could help with the clinical practice for two main reasons: 1) if focusing only on statistically relevant features, to be able to discard non significant ones, it could be reduced the amount of time and costs associated to these kind of tests; and 2) based on this reduced dataset, it could be established a simplified model for DIAD progression. According to these two ideas, the first experiment proposed in this work makes a between-group discrimination to discern between MC and NC to determine which features give a better classification performance (diagnosis) and which of them might be discarded from the point of view of a Machine Learning analysis.

2.7.2. Experiment 2 - Assessment of group heterogeneity in mutation carriers

Once determined which features get a better performance when comparing MC vs. NC, next step will be to determine if subgrouping the MC set according

to the affected genes (PSEN1, PSEN2, APP) improves classification accuracy. With the exception of defined analysis made in PSEN1 MC subjects subgroups [69, 4, 28, 3], there is no statistical evidence as yet that DIAD should be studied as an homogeneous entity including all subjects with at least one of those genes regardless the kind of mutation they present or treat each mutation separately. With this aim, we have made a binary classification between NC vs. PSEN1 carriers as the most populated MC subgroup and a multiclass classification comparing NC vs. PSEN1 vs. PSEN2 and vs. APP with particular attention to those features which gave better results in experiment 1.

2.7.3. Experiment 3 - Comparisons based on EYO

Having fitted the best model of DIAD according to the affected genes, a temporal comparison of these groups was performed. For that, several classifications of subjects subgrouped according to their EYO in steps of 5 years [EYO $\geq$ 15, EYO $\geq$ 10, EYO $\geq$ 5] are also carried out. The objective here is to determine at which range of EYO differences between MC and NC are maximized even when participants do not present any kind of AD cognitive symptom.

2.7.4. Extension of Experiment 2 - Clustering PSEN1-MC

To go deeper into the subject of data heterogeneity in DIAN database, an extra comparison has been added. As most of the asymptomatic MC participants (with special focus on PSEN1 MC) are quite similar to any NC subject, maybe this is resulting in a mixture of two different groups (clusters) including both: subjects quite similar to NC cases and those whose prognosis is clearly different. If it is proven this theory, it will reinforce the explanation of why classification results are so low in comparison with the proposed experiments even regardless the AAEOE value given for any subject.

2.8. Balance of subjects

For the comparisons proposed for each experiment and considering different cohorts of subjects, the total number of participants (NC and MC) in each experiment has been summarized as follows in Table 3:

Exp	Description	NC	MC
1	NC/MC	164	281
	NC/MC **	164	170
2	NC/PSEN1/PSEN2/APP	164	217/20/44
	NC/PSEN1/PSEN2/APP **	164	122/19/29
	NC/PSEN1	164	217
	NC/PSEN1 **	164	122
3	NC/PSEN1	164	217
	NC/PSEN1 **	164	122

Table 3: Balance of subjects in DIAN database for each experiment. Used ** to represent only asymptomatic cases.

2.9. Feature preparation

A general diagram of the data processing pipeline is shown in Figure 1 and it is organised as follows: first, a cross-validation procedure splits input data into two groups (training and test). Second, these subsets are standardized and used as inputs of a feature selection procedure based on an ANOVA test. And finally, once were selected features that matched with the previous conditions, a Principal Component Analysis (PCA) was used in order to reduce the number of features that will be used for their posterior classification.

2.9.1. Data standardization

In order to avoid having input variables with larger ranges that will affect to a posterior multivariate analysis [70], a feature standardization procedure based on rescaling using the z-score formula (1) was performed.

$$z_i = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \qquad i = 1, 2, \ldots, N \tag{1}$$

when x_i is defined to be the original feature i, μ_{x_i} its mean, and σ_{x_i} its standard deviation.

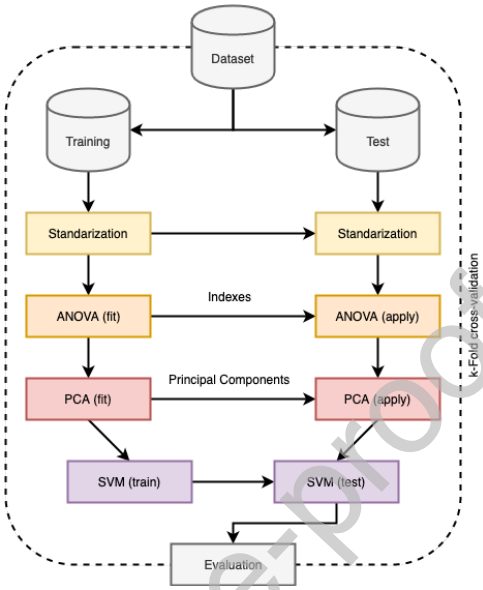

Figure 1: General diagram.

2.9.2. Missing data

Some of the test results in DIAN study are not available for all the subjects. In these cases, it was decided not to include subjects with unknown tests results in our analysis.

2.9.3. Analysis of variance (ANOVA)

The feature selection consists of two phases: the first one is a feature selection algorithm based on an ANOVA test [71, 72] to remove non-informative features for classification. Using this approach, two classes with N observations in each class are compared. For that, suppose $x_{i,j}$ denotes the j-th observation ($j = 1, 2, \ldots, N$) of a feature for the i-th class. As only considering experiments in which two classes are compared, for this work i only could take two possible values: $i = 1, 2$. Using this notation, group means $\bar{x}_i$ and global mean $\bar{x}$ are

defined as follows:

$$\bar{x}_i = \frac{1}{N}\sum_{j=1}^{N} x_{i,j} \qquad \bar{x} = \frac{1}{2N}\sum_{i=1}^{2}\sum_{j=1}^{N} x_{i,j} \tag{2}$$

Then, the estimated total variance of the sample $\hat{\sigma}^2_{\text{Total}}$ is descomposed into within class variance $\hat{\sigma}^2_{\text{within}}$ and the between class variance $\hat{\sigma}^2_{\text{between}}$ as:

$$\begin{aligned} \hat{\sigma}^2_{\text{Total}} &= \textstyle\sum_{i,j}\left(x_{i,j} - \bar{x}\right)^2 \\ \hat{\sigma}^2_{\text{within}} &= \textstyle\sum_{i,j}\left(x_{i,j} - \bar{x}_i\right)^2 \\ \hat{\sigma}^2_{\text{between}} &= N\textstyle\sum_{i}\left(\bar{x}_i - \bar{x}\right)^2 \end{aligned} \tag{3}$$

Using the one-way ANOVA feature selection method which tests the null hypothesis that means obtained from the two classes are equal, and calculating the expected values of equation (3), two new estimators for the variances can be obtained:

$$\begin{aligned} \hat{S}^2_{\text{within}} &= \tfrac{\hat{\sigma}^2_{\text{within}}}{2(N-1)} \\ \hat{S}^2_{\text{between}} &= \tfrac{\hat{\sigma}^2_{\text{between}}}{2(N-1)} \end{aligned} \tag{4}$$

Then, dividing $\hat{S}^2_{\text{between}}$ by $\hat{S}^2_{\text{within}}$, F-statistic is defined. This statistic is therefore used to test the null hypothesis considering an specific significance level α using a one-tailed test of the F-distribution. As highest F_{values} represent the most part of the variance in the target data, values with a high F_{values} are selected. Due to the correspondence between F_{values} and their probability, p-value, to reject the null hypothesis, only those features with p-value below a significance level (α) should be selected.

2.9.4. Principal Component Analysis (PCA)

This mathematical procedure [73, 74] provides an approximation of an input dataset $\mathbf{X}$ composed by K participants and M variables in terms of a new set

of N_{Comp} variables with $M \geq N_{\text{Comp}}$. This method uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called Principal Components (or loadings). In order to do that, PCA makes use of the eigenvalue decomposition of its covariance matrix as shown in expression (5):

$$\mathbf{X}^T\mathbf{X} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^T \tag{5}$$

where the columns of $\mathbf{W}$ contain the eigenvectors of $\mathbf{X}^T\mathbf{X}$ and Λ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{X}^T\mathbf{X}$. Therefore, the k^{th} subject of the original data $\mathbf{x}_k$ can be projected to the new space defined by $\mathbf{W}$, obtaining its new set of coordinates $\mathbf{s}_k$:

$$\mathbf{s}_k = \mathbf{x}_k\mathbf{W}^T \tag{6}$$

PCA has been commonly used in applications for AD diagnosis assistance [75, 76, 77]. Most direct consequence to these works is that a reduced number of principal components are necessary and enough to describe the information about AD encoded in the input features.

2.10. Classification

All individual classifications performed in this work make use of Support-Vector-Machine (SVM) classifiers [78, 79]. A SVM classifier is a ML algorithm that splits a given set of binary labeled training data into two subsets. It makes use of an hyperplane that maximizes the distance between the two trained classes. Following this idea, a function $f : \mathbb{R}^p \rightarrow \{1, 0\}$ using p-dimensional patterns $\mathbf{x}_i$ and class labels y_i is built to classify new examples $(\tilde{\mathbf{x}}, y)$:

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_R, y_R) \in \left(\mathbb{R}^p \times \{1, 0\}\right) \tag{7}$$

In order to maximize the distance to the decision boundary that separates the two classes (hyperplane), quadratic programming algorithms are used to minimize the margin cost function J as follows in expression (8) subject to the

inequality constraints in (9). Note that as feature vectors may belong to one of the next three cases: (a) well-classified feature vectors outside the hyperplane margins, (b) well-classified feature vectors inside the hyperplane margins and (c) missclassificated feature vectors. These three cases can be merged into one when introducing a new set of variables, ε_i, also called 'slack variables'. Therefore, the objective will be to maximize the hyperplane margins whereas mantaining $\varepsilon_i \geq 0$ as low as possible including the maximum amount of points well-classified.

$$J(\mathbf{w}, w_0, \varepsilon) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{l}\varepsilon_i \tag{8}$$

$$y_i[\mathbf{w}^T x_i + w_0] \geq 1 - \varepsilon_i, \qquad \varepsilon_i \geq 0, \qquad i = 1, 2, \ldots, l \tag{9}$$

The linear combination of a vectors subset (support vectors) conforms the solution of this problem.

2.11. Cross-validation strategy

To validate the classification results, each dataset has been splited into two groups: a training data group, which was used to train the prediction model, and a test data group, that is then used to measure the classifier's performance. In this work, as the number of participants is large enough, an N-fold cross-validation strategy with $N = 10$ has been selected [80]. In addition to these traditional/empirical methods for the evaluation of the classification performance, it was analyzed the upper bounds of the empirical error based on the theory of uniform convergence of means to their expectations [81], since their competitive performance in heterogeneous data applications, such as the field of neuroscience and neuroimaging. A recent advance in this regard, based on the "in general position" assumption of the data distribution, allows to obtain tighter upper bounds effectively connecting the empirical and actual risks for linear classifiers within a small confidence interval [82].

3. Results

Following the schema proposed in Figure 1 and the comparisons defined in Table 3, several simulations were performed to obtain the number of PCA components (N_{Comp}) which maximize the classification performance considering a previous ANOVA analysis which discards non-significant features for a significance level α.

Depending on the type of input data, four feature selection and classification procedures for each experiment were defined: one for the **NIB** tests and three for imaging features (**FDG**, **MRI** and **PiB**). Once discarded participants without all tests results for each experiment; discarded all tests with a variance equal to 0; having standarizated all remaining columns using the z-score formula as described in (1); and regardless the experiment considered; a total of 10 **NIB**, 520 **FDG**, 184 **MRI** and 1040 **PiB** results were included as input features for each stack.

Moreover, in order to check if these characteristics were suitable for a later ANOVA analysis, Kolmogórov-Smirnov test was applied [83]. The objective of this test was to determine the percentage of features from each group which can be assumed to present a normal distribution. In this work, for all experiments at least 50%-75% of the features matched this condition. Thus, it was admited the use of ANOVA for feature selection.

At this point, for each experiment, the feature selection procedure (ANOVA + PCA) is carried out over training data and evaluated following the 10-fold cross-validation schema. Classification is performed using an SVM classifier with Linear Kernel. All these experiments are calculated considering the 50 most significant features of each input data class using the ANOVA analysis and up to 20 number of components for PCA: $N_{\text{Comp}} = 1, 2, \ldots, 20$.

In the case of experiment 1, it was first decided to make a comparison between NC and MC participants regardless of their CDR. However, and as depicted in Figure 2 (upper), due to the overall small differences obtained following this idea with classification rates not better than 60.39% for **FDG** and less than

60.08% for **PiB**, it was finally decided to compare only asymptomatic subjects for both groups. Moreover, as one of the main objectives for this work was to determinate if subgrouping MC participants in terms of their responsible gene, a second comparison of NC vs asymptomatic PSEN1 vs asymptomatic PSEN2 vs asymptomatic APP was proposed. In that case, the problem was related to the size of each class (only 20 asymptomatic PSEN2 and 29 asymptomatic APP cases) which limits the generalization ability of each classifier[4]. Thus, it was finally decided only to focus on the most populated group and make a direct comparison between NC vs asymptomatic PSEN1 (Experiment 2) which has led to an increase in classification performance above 11.31% when $N_{\text{Comp}} = 15$ as shown in Figure 2 (down). Note that performance represented has been defined in terms of balanced accuracy which is equal to the arithmetic mean of sensitivity (true positive rate) and specificity (true negative rate).

As most of the values obtained showed low classification results, it was analyzed the upper bounds of the empirical error based on the theory of uniform convergence of means to their expectations like explained in section 2.11 for some of the components computed for PCA. In Table 4, it has been included both upper bounds of the empirical error and cross-validation classification performance for a given set of N_{Comp} and each kind of input data from experiment 2. Note that, in order not to overload the table, only some of these combinations have been depicted.

In regard of the results obtained, it was decided to use $N_{\text{Comp}} = 15$ as our reference level for PCA analysis. Using this value, different levels of EYO have been processed from the same point of view (experiment 3) as a way to determine how EYO has an effect on the classification outcome. As depicted in Figure 3, best performance results in terms of balanced accuracy were obtained for: **MRI** at $EYO \geq 15$ years (70.00%); **PiB** at $EYO \geq 10$ years (63.75%); and **PiB** at $EYO \geq 5$ years (60.00%).

[4]This problem appears when the sample size is too small in comparison with the number of input features [84].

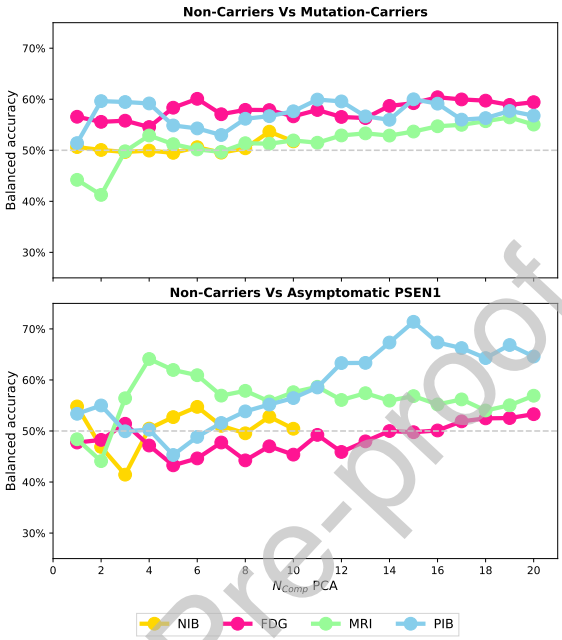

Figure 2: Classification results for experiments 1 and 2 using a SVM classifier with Linear Kernel and considering different types of input data for each N_{Comp} in PCA analysis.

Input data	N_{Comp}	CV accuracies (%)	re-Substitution (%)	Upper bounds (%)$^{(*)}$	h
PIB	**3**	50.00	62.70	16.34	Reject
	10	57.46	85.38	28.21	Accept
	15	71.40	87.50	29.31	Accept
	20	64.60	89.48	30.35	Accept
MRI	**3**	56.41	61.21	16.34	Reject
	10	57.60	65.82	28.21	Reject
	15	56.83	68.01	29.31	Reject
	20	56.90	69.83	30.35	Reject

Table 4: Classification results for NC vs asymptomatic PSEN1 comparison. Neither **FDG** nor **NIB** showed significant results. $^{(*)}$ Reference levels obtained from [82]. CV accuracies column is expressed in terms of the balanced accuracy.

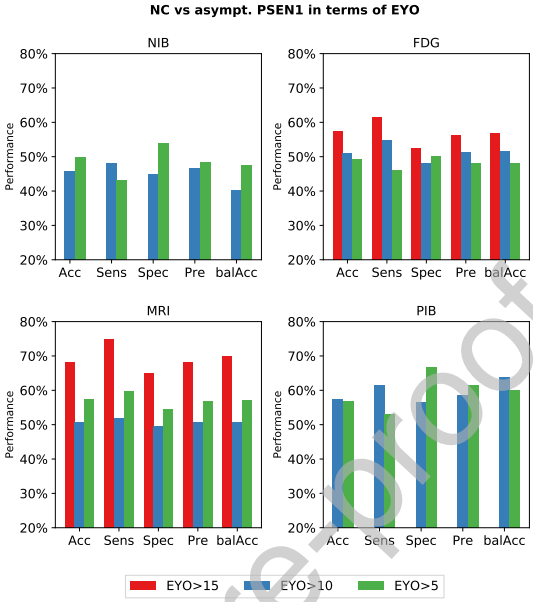

Figure 3: Classification results for experiment 3 depending on the Estimated Years to Onset (EYO) when comparing NC vs. PSEN1 (only asymptomatic cases).

For each experiment and considering all the imaging data types, ROC curves have been calculated as shown in Figure 4 with the exception of EYO analysis which has been included as Supplementary Material. These curves can be used to check whether a fixed number of features selected by ANOVA analysis and a particular value for PCA N_{Comp} might involve a robust solution for the computer aided-diagnosis system. Moreover, to reinforce this idea, Area Under the Curve (AUC) has also been computed as represented in Figure 5 for **PiB**[5] with maximum levels AUC = 0.679 at $N_{\text{Comp}} = 15$ when comparing directly **PiB** features from NC subjects vs asymptomatic PSEN1 participants.

In order to show the PCA results at a more detailed level, analysis of its variance has also been carried out. As depicted in Figure 6 for imaging markers,

[5]AUC for both **FDG** and **MRI** tests have been included as Supplementary Material.

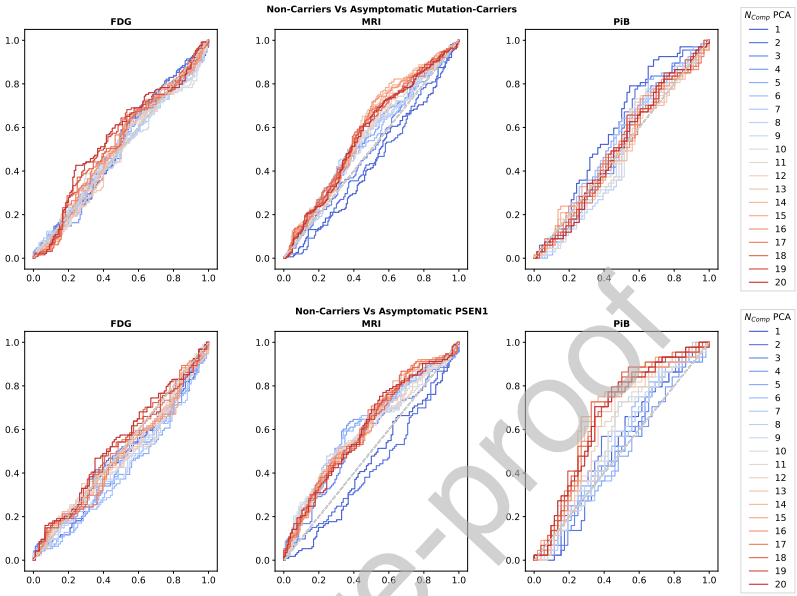

Figure 4: ROC curves obtained considering different types of input features for experiments 1 and 2.

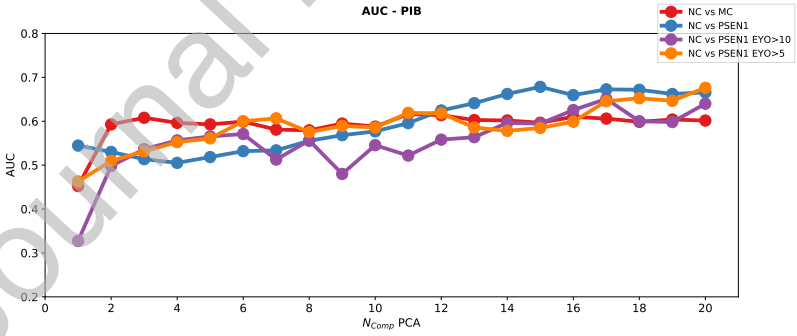

Figure 5: AUC obtained for each comparison performed for experiments 1, 2 and 3 when only using **PiB** imaging features.

both **FDG**, **MRI** and **PiB** reach the 90% of their total variance within the interval $N_{\text{Comp}} \in [0, 10]$. Thus, it can be assumed that increasing N_{Comp} will

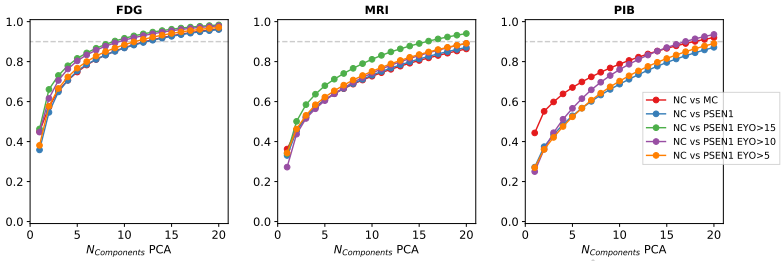

Figure 6: Cumulative variance of PCA relative to N_{Comp}.

not provide significant differences in classification. This is the reason why N_{Comp} has been choosen as a value in the $[0, 20]$ interval.

Apart from the proposed schema using ANOVA + PCA + SVM methods, a direct characterization of biomedical tests has also been presented here considering only those features which present a significance level for ANOVA below a certain threshold. Figure 7 shows the most significant imaging features for experiment 2 in terms of ANOVA analysis as example[6]. Moreover, in addition to this plot, a graphical representation of p-value for the first two experiments and considering a significance level for ANOVA $\alpha = 0.05$ has also been included as shown in Figure 8. Note that as lower p-values represent more highlighted differences between groups comparisons, we have made use of the $\log(p\text{-value})$ instead of p-value.

To complete this section, a comment about the study case presented in section 2.7.4 is also provided. When it was considered the scenario where NC and asymptomatic PSEN1 MC were compared, it was analyzed the distribution of the most significative features from both NC and PSEN1 MC classes. As shown in Figure 9, the histograms obtained from each marker revealed two similar distributions[7] with no many differences between them.

[6] Remaining violinplot figures for experiments 1 and 3 have been included as Supplementary Material.

[7] In order not to overload this work, only **FDG** results were depicted.

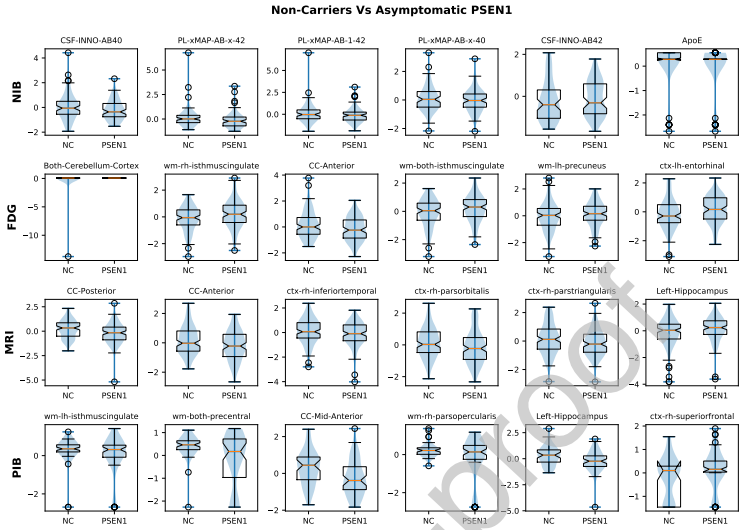

Figure 7: Violinplots and Boxplots for the 6 most significant features of each input group (**NIB**, **FDG**, **MRI** and **PiB**) when comparing NC vs asymptomatic PSEN1 MC.

Although this analysis could be ended here, shapes and sizes of both distributions suggested the existence of at least two subclasses within each group with special emphasys when considering the PSEN1 MC class. Making use of the Akaike Information criterion [85] for k-Means to determine the minimum number of clusters in a class as well as the k-Means algorithm itself, they were computed the two clusters within the PSEN1 subset. As shown in Figure 10, one of the PSEN1 subsets (violet) is usually displaced in relation to the centre of NC class (yellow) unlike the second subset (turquoise). This behaviour remains the same regardless also including symptomatic PSEN1 MC.

In this scenario, the new classification rates obtained when comparing NC vs asymptomatic PSEN1 MC from the first cluster, and the NC vs asymptomatic PSEN1 MC from the second cluster showed significant differences between them. In fact, following the same procedure as perfomed for previous experiments, this resulted in classification rates above 73.56% (first cluster) and 66.08% (sec-

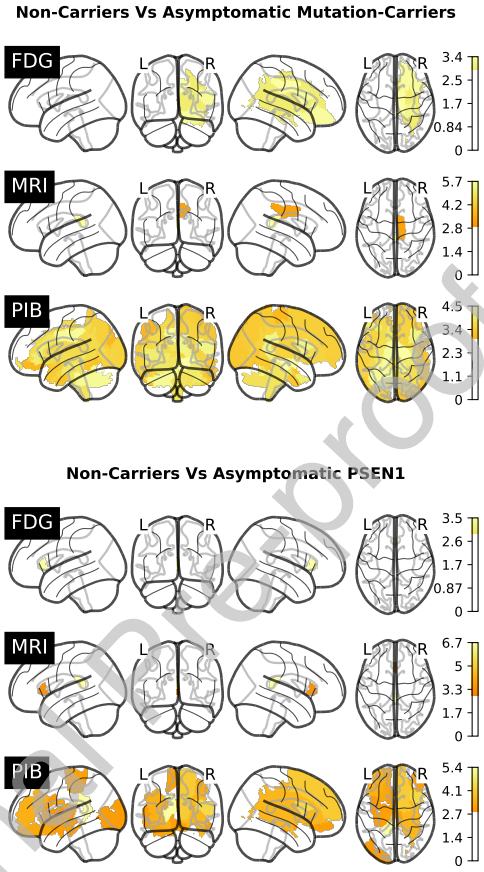

Figure 8: Atlas plot of significant brain regions using the $\log(F_{\text{value}})$ measure. Threshold referred to the significance level 0.05.

ond cluster) when using $N_{\text{Comp}} = 15$ as our reference in previous experiments. This difference is even more highlighted when also including symptomatic cases: 77.69% (first cluster) and 65.44% (second cluster) as depicted in Figure 11.

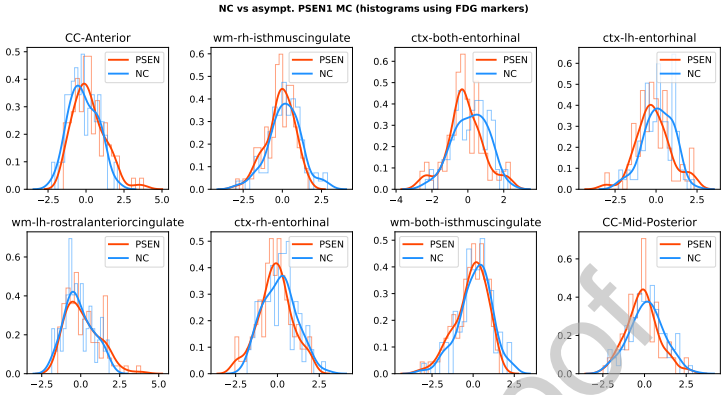

Figure 9: Histograms obtained from the 8 most significative **FDG** imaging markers when comparing NC vs asymptomatic PSEN1 MC.

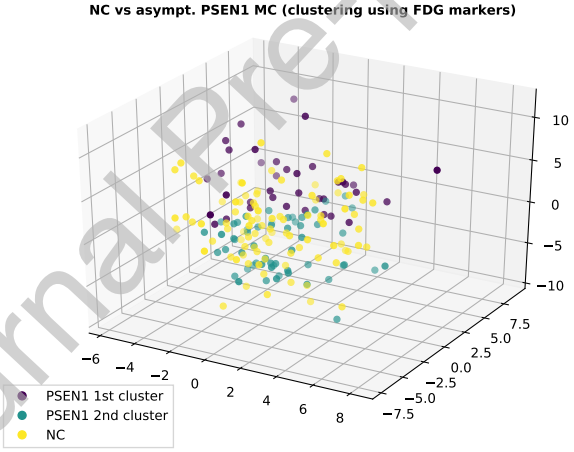

Figure 10: Clustering obtained from **FDG** imaging markers when comparing NC vs asymptomatic PSEN1 MC. This 3D figure has been generated by representing the $N_{\text{Comp}} = 3$ components obtained from PCA analysis step.

4. Discussion & Conclusions

A deep characterization of the clinical and preclinical stages of AD is critical to develop new lines of treatment for the disease [86, 24, 5]. As many works

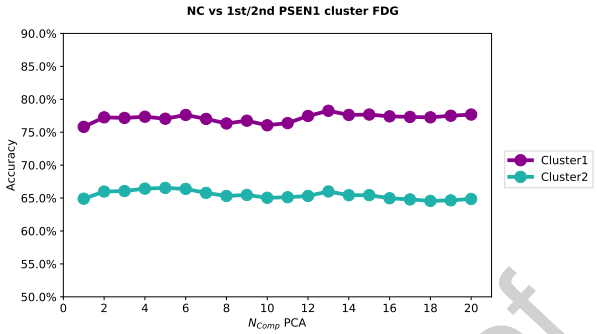

Figure 11: Classification rates obtained from **FDG** imaging markers when comparing NC vs PSEN1 MC. These results underscore the similarity between second PSEN1 cluster with respect NC whereas the first cluster is clearly more differentiated.

point out, AD should not be considered as a single entity [87]. Some of the AD subtypes, including the rare autosomal dominant form might evolve differently and require different ways to face up with the disease. However, most of the current clinical trials carried out to test interventions for AD did not discriminate between disease subtypes (mainly defined genetically in the case of DIAD).

For this work, the 3 experiments as described in section 2.7 section have been carried out with the aim to explore the possibilities offered by ML in analyzing DIAN dataset. In all the experiments presented here, a set of healthy control and non-carriers subjects have been compared with a group of mutation-carriers considering as inputs biomedical markers (**NIB**) and imaging markers (**FDG**, **MRI**, **PiB**). Due to the particular cases considered, experiments are discussed individually as follows in the next subsections. Then, a final section summarizes all the key points in common between them.

4.1. ***Experiment 1*** *- NC vs MC*

Experiment 1 makes a comparison between NC and MC first considering all posible mutation types at every stage. Discussion about this experiment can be started from the analysis of classification rates obtained with a maximum balanced accuracy of 60.39% when using **FDG** imaging markers and 60.08% for

PiB. If focusing on the other kind of tests, this rate decreases to even lower rates regardless of the number of input features selected from the ANOVA analysis or the combination ANOVA+PCA. To this aim, in a second step, it was decided to use only asymptomatic participants as a way to reduce the heterogeneity between subjects. This resulted in classification rates slightly higher than 60% when using **PiB** but still to far away from our objective. In this scenario, although **PiB** presented a better behaviour, results obtained have to be treated with caution since the low rates obtained and the lack of robustness shown in Figure 4: **PiB** curves present a considerable difference with respect the diagonal line but with a low AUC (maximum AUC equal to 0.6138 when using a high N_{Comp} for PCA). In the light of these findings, it can be assumed that **PiB** features ensure an appropriate separation of MC and NC but within a noisy environment in which missclasification errors happen very often. If unmasking data heterogeneity, better classification rates are expected to be obtained.

Despite of having some highly significant brain regions obtained from the ANOVA analysis selection such as those highlighted in Figure 8; superposition and data heterogeneity is limiting effectiveness of ML analysis for this experiment. Nevertheless, these results are quite similar to obtained in other works with special attention to central brain and right hemisphere. For example, [86] proved significant differences for carriers at *temporal lobe* (both medially in the region of the *hippocampus* and laterally in the *temporal neocortex*), *cuneus*, *precuneus*, *cingulate*, *putamen* and *thalamus* also referred by [88, 89, 16].

4.2. ***Experiment 2*** *- NC vs subgrouped MC*

As a way to highlight the relevance of better comparing NC vs subgrouped MC in future DIAN studies, a second comparison between NC vs PSEN1 vs PSEN2 vs APP MC was performed. However, due to the low number of subjects in PSEN2 and APP classes as depicted in 2, it was finally decided only to compare NC vs PSEN1 MC participants among them.

In regard of Figure 2 results for NC vs asymptomatic PSEN1 MC, when selecting a high N_{Comp} for PCA, classification results increase up to 71.4%

(**PiB** features). This supposes an increase above 13.16% with respect the NC vs asymptomatic MC comparison. In fact, regardless of the Estimated Years to Onset (EYO) from any patient, this difference in **PiB** and **MRI** scans is susceptible of being used to determine whether a new subject is more likely to be developing DIAD or if this assumption still remains far in time.

Finally, owing to AUC results in Figure 5, the solution proposed shows a robust behaviour. A clear illustration of this can be seen in the region between $N_{\text{Comp}} = 4$ and $N_{\text{Comp}} = 15$ where the increasing in AUC values take place in a linear way. Indeed, even considering a N_{Comp} out of these margins, none of the AUC values raise (or lower) dramatically.

As depicted in Figure 8, now the most highlighted regions are located at: *cerebellum*, *entorhinal*, *cingulate* (both caudal-ant., isthmus, post. and rostral-ant.), *medial orbitofrontal*, *frontal pole*, *middletemporal*, *parahippocampal*, *corpus callosum* (both ant., mid-ant. and post.), *pars triangularis*, *precentral*, *ventral DC*, *brain-stem* and *insula*. In comparison with the current literature related to this experiment, [4] described a significant deposition of $A\beta$ amiloid in WM but it did not specified which were the particular regions where these results were obtained. Only [69] referred changes at *cerebellum* corresponding with one of our most significant regions for **PiB** markers. Note that as those works only were referred to PSEN1 MC but without specifying ther CDR grade, these results are not totally comparable to ours. The only work that presented a similar comparison was [28] where making use of brain imaging and fluid biomarkers, they characterized 18 asymptomatic children with E280A PSEN1 mutation. In that case, it was proven that MC children were distinguished from control individuals by higher levels in $A\beta_{42}$ and $A\beta_{42} : A\beta_{40}$ ratio plus differences measured in *precuneus* area, *post. cingulate*, *medial temporal lobe* and *hippocampus*. These results also correspond with the presented here although **NIB** result have not been significant for our ML model.

4.3. ***Experiment 3*** - *NC vs asymptomatic PSEN1 MC (EYO consideration)*

Once it has been proven that comparing NC vs MC subgrouped by their responsible gene is more effective from the point of view of ML analysis specially for **PiB** features, next step has been to determine if this significance is variable in time. For that, considering different levels of EYO applied to asymptomatic PSEN1 MC, input features have been classified and compared among them as shown in Figure 3.

In view of the results obtained, what first calls the attention is the lack of classification results for **PiB** and **NIB** at $EYO \geq 15$ years. In [20], authors referred changes for signaling findings at *neostriatum* area. However, using our balanced muestral size of $[35, 35]$ participants, not enough differences between NC and asymptomatic PSEN1 MC with $EYO \geq 15$ years were found. The same reasoning applies to **NIB** features despite findings given by other works such as [20, 89, 9, 14, 90] where it was stated that $A\beta$ depositions (and increasing of CSF τ levels and brain atrophy) occurs at least 15 years before symptoms onset when comparing NC vs MC. On the contrary, both **MRI** and **FDG** classifications have returned classification rates above 70.50% specially when considering **FDG** scan results. Indeed, ANOVA analysis carried out has underscored the following brain regions: *paracentral*, *postcentral*, *insula*, *putamen*, *cingulate* (both. caudal-ant. and rostral-ant.), *fusiform gyrus*, *medial orbitofrontal*, *parahippocampal*, *corpus callosum* (both central, mid-ant. and post.), *choroid-plexus*, *accumbens*, *inf. lat. vent.*, *inf. temporal*, *frontal pole*, *cuneus*, *pars orbitalis*, *lingual* and *supramarginal*. Some of these areas corresponds with the findings discovered by other works like [20] using **MRI** scans at this stage.

At a posterior stage of DIAD pathogenesis, when introducing $EYO \geq 10$ cases ($[64, 64]$ participants), the first data patterns using **PiB** features and/or **NIB** begin to be significant and SVM classifiers start to discern between NC and MC cases[8].

It is known that *hippocampal* volume reductions are observed as early as

[8] At $N_{\text{Comp}} = 15$, AUC $= 0.595$ for **PiB** features

10 years before expected onset [86, 20]. Moreover, there are also references to changes in **FDG** and **PiB** located at: *precuneus*, *entorhinal*, *precentral*, *parietal lobule* (both lateral and post-central) and *post. cingulate* regions; as well as significant changes for **MRI** features related to atrophy of *accumbens* and *amygdala* [32, 89]. Most of these patterns fit with the majority of the highlighted regions by our ANOVA analysis but also including: *poscentral*, *frontal pole*, *paracentral*, *corpus callosum*, *pericalcarine*, *cerebellum*, *inf. lat. vent.*, *choroid-plexus*, *pars opecularis*, *lateral orbitofrontal*, *insula*, *supramarginal* and *middletemporal*. With reference to **NIB** features, although [9] pointed out significant concentrations of $A\beta_{42}$, τ and p-τ markers in MC group at very early EYO points, these results have not been relevant in this case.

Finally, with respect $EYO \geq 5$ results, it is well known that decreased volumes of the *thalamus*, *caudate*, *temporal lobe*, *parietal lobe* and *occipital lobe* are reported in presymptomatic PSEN1 mutation carriers about 5 years before symptom onset [86, 13, 90]. Other works such as presented in [20, 89, 14] confirm these results, add new regions like *precuneus* and *entorhinal* areas to the list, and even suggest significant changes in cerebral hypometabolism and hippocampal atrophy during this stage of the disease. All of these results are consistent with ours but extending our findings to: *cerebellum*, *ventral DC*, *corpus callosum*, *accumbens*, *choroid-plexus*, *pars opecularis*, *insula*, *amygdala*, *caudal middlefrontal*, *precentral*, *paracentral*, *poscentral* and *post. cingulate*.

4.4. Extension of Experiment 2

One last question that remains to be answered is if classification results for experiments 1, 2 and 3 could be improved. Although it has been proved 1) that PET scans are the most suitable kind of test for DIAD, and 2) that reducing the heterogeneity of the disease (for example by differentiating participants according to their responsible gene) we can get a more precise model of the pathology; similarity between NC and asymptomatic MC at early stages of the disease should be further explored. This is the main reason why the extension of experiment 2 proposed in section 2.7.4 has been suggested as a way to measure

the relationship between NC subsets and those without any kind of mutation. In this scenario, distribution analysis of the most significant features for DIAD could be a good starting point. As shown in Figure 9 for **FDG**, a preliminary assessment of NC and PSEN1 distributions suggests that both classes are very similar. Although this behaviour was not as good as expected, in view of the distribution tails from PSEN1 distributions it was stated that maybe this shape might be related to a mixture of two/more distributions. To confirm this idea, a clustering approach based on the use of k-Means algorithm was applied to the data as a way to determine if one/more of these PSEN1 subclasses could belong to the same distribution of NC participants or if they could be considered as an independent entity. Fortunately, as depicted in Figure 11 when using **FDG** imaging markers[9], the second statement was confirmed. In fact, whereas one of the PSEN1 clusters is hard to distinguish from NC with classification rates barely above 64%, the other one improve these rates with balanced accuracy rates close to 80%. This procedure has been also repeated for **NIB**, **MRI** and **PiB** confirming this trend though not as much margin as using **FDG** including both symptomatic and asymptomatic cases. For example, when only using asymptomatic PSEN1 MC cases, PSEN1 subclasses got a difference margin up to 6.1% whereas **FDG** got 7.5%.

Relative to demographics extracted from each cluster, both populations are closely similar between them in terms of sex, age or MMSE as summarized in Table 5 (both symptomatic and asymptomatic cases) and Table 6 (only asymptomatic). This fact reflects that clustering is not based on the separation of participants according to their sex, age, or even their degree of disease; but only about activation measures. Although no significant differences were found with respect their first-degree relatives age at onset or the average age at onset obtained for the PSEN1 group, a further analysis of these subgroups is proposed for future work.

[9] As the input markers type which gave the best results.

Cluster	Group	N	Age	MMSE
First	Male	18	40.17 ± 11.38	26.67 ± 5.04
	Female	28	38.61 ± 10.88	26.68 ± 6.03
	Both	46	39.21 ± 11.10	26.67 ± 5.67
Second	Male	23	42.65 ± 10.57	26.65 ± 4.98
	Female	51	39.37 ± 9.76	26.81 ± 4.74
	Both	74	40.39 ± 10.14	26.76 ± 4.82

Table 5: Demographics obtained from the two clusters computed with k-Means for NC vs PSEN1 MC.

Cluster	Group	N	Age	MMSE
First	Male	16	41.69 ± 10.70	26.94 ± 5.32
	Female	21	37.38 ± 9.72	26.71 ± 5.00
	Both	37	39.24 ± 10.38	26.81 ± 5.48
Second	Male	20	44.20 ± 10.43	26.65 ± 4.98
	Female	37	39.27 ± 9.16	26.81 ± 4.74
	Both	57	41.00 ± 9.91	27.79 ± 4.68

Table 6: Demographics obtained from the two clusters computed with k-Means for NC vs asymptomatic PSEN1 MC.

4.5. ***General conclusions***

When this work was proposed, its main objective was to answer three questions: first, which kind of clinical test was the most relevant for DIAN diagnosis; second, considering MC subjects all together or separately (3 genes whose mutation is responsible of DIAD) if this could aid us to improve the model of the disease; and third, to confirm if these ideas could be used for a longitudinal analysis from the point of view of ML algorithms and to test its validity. In this sense, there is no other work trying to perform a ML analysis using this database so any conclusion stated here can be used for future works in the study of DIAN.

Owing to classification results obtained from each experiment performed, the first conclusion that we can assume is that all the imaging tests can provide relevant information concerning the DIAD prognosis and/or its pathogenesis with special emphasis on **PiB** features. This is indicating that even subtle dif-

ferences in Aβ plaques deposited in neuronal tissue are more related to a higher separation between MC and non-MC than any other kind of test considered for this work. In fact, regardless the experiment under consideration, differences between NC and MC are larger using **PiB** than any other kind of input marker with an exception: results obtained whe comparing NC vs asymptomatic PSEN1 MC with $EYO \geq 15$ years in experiment 3. Nevertheless, these results should be taken with care due to the kind of input markers used: SUVR values obtained using the analysis suite `FreeSurfer`. It is no exaggeration to state that a proper image preprocessing particularly designed for AD study might bring new evidences about pathogenesis and prognosis of the disease better than SUVR values do. Indeed, an specific analysis of those regions might be also important for future works even when a no direct evidence of AD has appeared yet (asymptomatic subjects) [14].

With respect to the improvement achieved when subgrouping DIAD participants in DIAN study, it is obvious that comparisons using subgrouped MC outperforms the results obtained for NC vs MC from experiment 1. Although it would have been desirable to compare NC vs PSEN1 vs PSEN2 and APP MC in a multiclass classification schema, the improvement reached when comparing NC vs asymptomatic PSEN1 MC instead of just using NC vs asymptomatic MC has been noteworthy (11.31% of classification increase for experiment 2). Further, if taking into account the results obtained for the extension of experiment 2, it has been demonstrated that even comparisons within PSEN1 MC with respect the control group (NC) show a strong differentiation among two PSEN1 subclasses in terms of imaging response. This fact has not been described before beyond the proposal made by [3] who suggested the existence of two or more differenciated groups for PSEN1 MC from the point of view of their symptomatology. Now, as this assumption was mathematically confirmed, this result constitutes an important milestone for the understanding of AD for two simple reasons: 1) it helps to corroborate findings in previous works, and 2) it establishes the basis for future comparisons even despite the low discriminative information extracted from DIAD markers but with special emphasis on PET

imaging markers.

In conclusion, despite subjects with an inherited autosomal dominant AD mutation represent less than 1% of AD persons, the study carried out by DIAN initiative constitutes a strong impact in the understanding of AD with special emphasis on the disease course [26, 30]. As exposed during the discussion of all experiments, merging all MC subtypes as one unique MC group (even regardless the stage of their pathogenesis) might lead to a loss in statistical power. In the current work, it has been shown that our model fits better with DIAD progression even at its earliest stages. In this sense, since the proposal of a theoretical biomarker changes model for AD by [6], though several works have pointed out the relevance of different markers even 20 years before first symptoms onset, the application of deeper mathematical tools like ML models aid to discard some of them and to concentrate only of those clearly relevant. For all these reasons, it is expected that the use of new DIAN database updates joined with a deep image processing of **FDG**, **MRI** but, above all, **PiB** scans; provide insights into DIAD and LOAD and could even potentially be employed as read out in future treatment trials[10].

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author contributions

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

[10]For example, the recent trials involving anti-$A\beta$ antibodies referred by [24].

Acknowledgment

This work was supported by the MINECO/FEDER under the TEC2015-64718-R and RTI2018-098913-B-I00 projects and the Ministry of Economy, Innovation, Science and Employment of the Junta de Andalucía under the Excellence Project P11-TIC-7103.

LS is supported by Alzheimers Research UK Senior Research Fellowship (ARUK-SRF2017B-1).

HM is supported by Japan AMED.

Data collection and sharing for this project was supported by The Dominantly Inherited Alzheimers Network (DIAN, U19AG032438) funded by the National Institute on Aging (NIA), the German Center for Neurodegenerative Diseases (DZNE), Raul Carrea Institute for Neurological Research (FLENI), Partial support by the Research and Development Grants for Dementia from Japan Agency for Medical Research and Development, AMED, and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI). This manuscript has been reviewed by DIAN Study investigators for scientific content and consistency of data interpretation with previous DIAN Study publications. We acknowledge the altruism of the participants and their families and contributions of the DIAN research and support staff at each of the participating sites for their contributions to this study.

References

[1] P. H. St George-Hyslop, J. L. Haines, L. A. Farrer, R. Polinsky, C. V. Broeckhoven, A. Goate, D. R. C. McLachlan, H. Orr, A. C. Bruni, S. Sorbi, I. Rainero, J. F. Foncin, D. Pollen, J.-M. Cantu, R. Tupler, N. Voskresenskaya, R. Mayeux, J. Growdon, V. A. Fried, R. H. Myers, L. Nee, H. Backhovens, J.-J. Martin, M. Rossor, M. J. Owen, M. Mullan, M. E. Percy, H. Karlinsky, S. Rich, L. Heston, M. Montesi, M. Mortilla, N. Nacmias, J. F. Gusella, J. A. Hardy, Genetic linkage studies suggest that alzheimer's disease is not a single homogeneous disorder, Nature 347 (1990) 194.

doi:10.1038/347194a0.
URL http://dx.doi.org/10.1038/347194a0

[2] J. Ringman, S. Monsell, D. Ng, Y. Zhou, A. Nguyen, G. Coppola, V. Van Berlo, M. Mendez, S. Tung, S. Weintraub, M. Mesulam, E. Bigio, D. Gitelman, A. Fisher-Hubbard, R. Albin, H. Vinters, Neuropathology of autosomal dominant alzheimer disease in the national alzheimer coordinating center database, Journal of Neuropathology and Experimental Neurology 75 (3) (2016) 284–290. doi:10.1093/jnen/nlv028.

[3] N. S. Ryan, J. M. Nicholas, P. S. J. Weston, Y. Liang, T. Lashley, R. Guerreiro, G. Adamson, J. Kenny, J. Beck, L. Chavez-Gutierrez, B. de Strooper, T. Revesz, J. Holton, S. Mead, M. N. Rossor, N. C. Fox, Clinical phenotype and genetic associations in autosomal dominant familial alzheimer's disease: a case series, The Lancet Neurology 15 (13) (2016) 1326–1335. doi:https://doi.org/10.1016/S1474-4422(16)30193-4.
URL http://www.sciencedirect.com/science/article/pii/S1474442216301934

[4] J. M. Ringman, A. Goate, C. L. Masters, N. J. Cairns, A. Danek, N. Graff-Radford, B. Ghetti, J. C. Morris, Genetic heterogeneity in alzheimer disease and implications for treatment strategies, Current Neurology and Neuroscience Reports 14 (11) (2014) 499. doi:10.1007/s11910-014-0499-8.
URL https://doi.org/10.1007/s11910-014-0499-8

[5] R. N. Martins, V. Villemagne, H. R. Sohrabi, P. Chatterjee, T. M. Shah, G. Verdile, P. Fraser, K. Taddei, V. B. Gupta, S. R. Rainey-Smith, E. Hone, S. Pedrini, W. L. Lim, I. Martins, S. Frost, S. Gupta, S. O'Bryant, A. Rembach, D. Ames, K. Ellis, S. J. Fuller, B. Brown, S. L. Gardener, B. Fernando, P. Bharadwaj, S. Burnham, S. M. Laws, A. M. Barron, K. Goozee, E. J. Wahjoepramono, P. R. Asih, J. D. Doecke, O. Salvado, A. I. Bush, C. C. Rowe, S. E. Gandy, C. L. Masters, Alzheimer's disease: a journey from amyloid peptides and oxidative stress, to biomarker

technologies and disease prevention strategies-gains from aibl and dian cohort studies, Journal of Alzheimer's Disease 62 (3) (2018) 965–992. doi:10.3233/JAD-171145.

[6] C. R. Jack, D. S. Knopman, W. J. Jagust, L. M. Shaw, P. S. Aisen, M. W. Weiner, R. C. Petersen, J. Q. Trojanowski, Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade, The Lancet Neurology 9 (1) (2010) 119–128. doi:https://doi.org/10.1016/S1474-4422(09)70299-6.
URL http://www.sciencedirect.com/science/article/pii/S1474442209702996

[7] R. A. Sperling, P. S. Aisen, L. A. Beckett, D. A. Bennett, S. Craft, A. M. Fagan, T. Iwatsubo, C. R. Jack, J. Kaye, T. J. Montine, D. C. Park, E. M. Reiman, C. C. Rowe, E. Siemers, Y. Stern, K. Yaffe, M. C. Carrillo, B. Thies, M. Morrison-Bogorad, M. V. Wagster, C. H. Phelps, Toward defining the preclinical stages of alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease, Alzheimer's & Dementia 7 (3) (2011) 280–292. doi:https://doi.org/10.1016/j.jalz.2011.03.003.
URL http://www.sciencedirect.com/science/article/pii/S1552526011000999

[8] C. R. Jack, D. S. Knopman, W. J. Jagust, R. C. Petersen, M. W. Weiner, P. S. Aisen, L. M. Shaw, P. Vemuri, H. J. Wiste, S. D. Weigand, T. G. Lesnick, V. S. Pankratz, M. C. Donohue, J. Q. Trojanowski, Tracking pathophysiological processes in alzheimer's disease: an updated hypothetical model of dynamic biomarkers, The Lancet Neurology 12 (2) (2013) 207–216. doi:https://doi.org/10.1016/S1474-4422(12)70291-0.
URL http://www.sciencedirect.com/science/article/pii/S1474442212702910

[9] A. M. Fagan, C. Xiong, M. S. Jasielec, R. J. Bateman, A. M. Goate,

T. L. S. Benzinger, B. Ghetti, R. N. Martins, C. L. Masters, R. Mayeux, J. M. Ringman, M. N. Rossor, S. Salloway, P. R. Schofield, R. A. Sperling, D. Marcus, N. J. Cairns, V. D. Buckles, J. H. Ladenson, J. C. Morris, D. M. Holtzman, , Longitudinal change in csf biomarkers in autosomal-dominant alzheimer's disease, Science Translational Medicine 6 (226). doi:10.1126/scitranslmed.3007901.
URL http://stm.sciencemag.org/content/6/226/226ra30

[10] S. J. B. Vos, F. Verhey, L. Frlich, J. Kornhuber, J. Wiltfang, W. Maier, O. Peters, E. Rther, F. Nobili, S. Morbelli, G. B. Frisoni, A. Drzezga, M. Didic, B. N. M. van Berckel, A. Simmons, H. Soininen, I. Koszewska, P. Mecocci, M. Tsolaki, B. Vellas, S. Lovestone, C. Muscio, S.-K. Herukka, E. Salmon, C. Bastin, A. Wallin, A. Nordlund, A. de Mendona, D. Silva, I. Santana, R. Lemos, S. Engelborghs, S. Van der Mussele, , Y. Freund-Levi, . K. Wallin, H. Hampel, W. van der Flier, P. Scheltens, P. J. Visser, Prevalence and prognosis of alzheimers disease at the mild cognitive impairment stage, Brain 138 (5) (2015) 1327–1338. doi:10.1093/brain/awv029.
URL http://dx.doi.org/10.1093/brain/awv029

[11] M. Hutton, J. Pérez-Tur, J. Hardy, Genetics of alzheimer's disease, Essays In Biochemistry 33 (1998) 117–131. doi:10.1042/bse0330117.
URL http://essays.biochemistry.org/content/33/117

[12] R. I. Scahill, G. R. Ridgway, J. W. Bartlett, J. Barnes, N. S. Ryan, S. Mead, J. Beck, M. J. Clarkson, S. J. Crutch, J. M. Schott, et al., Genetic influences on atrophy patterns in familial alzheimer's disease: a comparison of app and psen1 mutations, Journal of Alzheimer's disease 35 (1) (2013) 199–212.

[13] N. S. Ryan, G.-J. Biessels, L. Kim, J. M. Nicholas, P. A. Barber, P. Walsh, P. Gami, H. R. Morris, A. J. Bastos-Leite, J. M. Schott, J. Beck, S. Mead, L. Chavez-Gutierrez, B. de Strooper, M. N. Rossor, T. Revesz, T. Lashley, N. C. Fox, Genetic determinants of white matter hyperintensities and amyloid angiopathy in familial alzheimer's disease, Neurobiology of Ag-

ing 36 (12) (2015) 3140–3151. doi:10.1016/j.neurobiolaging.2015.08.026.

[14] M. Suárez-Calvet, M. Á. Araque Caballero, G. Kleinberger, R. J. Bateman, A. M. Fagan, J. C. Morris, J. Levin, A. Danek, M. Ewers, C. Haass, , Early changes in csf strem2 in dominantly inherited alzheimer's disease occur after amyloid deposition and neuronal injury, Science Translational Medicine 8 (369). doi:10.1126/scitranslmed.aag1767.
URL http://stm.sciencemag.org/content/8/369/369ra178

[15] G. Di Fede, M. Catania, E. Maderna, R. Ghidoni, L. Benussi, E. Tonoli, G. Giaccone, F. Moda, A. Paterlini, I. Campagnani, S. Sorrentino, L. Colombo, A. Kubis, E. Bistaffa, B. Ghetti, F. Tagliavini, Molecular subtypes of alzheimers disease, Scientific reports. 8 (1). doi:10.1038/s41598-018-21641-1.

[16] N. P. Oxtoby, A. L. Young, D. M. Cash, T. L. S. Benzinger, A. M. Fagan, J. C. Morris, R. J. Bateman, N. C. Fox, J. M. Schott, D. C. Alexander, Data-driven models of dominantly-inherited alzheimer's disease progression, Brain 141 (5) (2018) 1529–1544. doi:10.1093/brain/awy050.
URL http://dx.doi.org/10.1093/brain/awy050

[17] E. Levy, M. Carman, I. Fernandez-Madrid, M. Power, I. Lieberburg, S. van Duinen, G. Bots, W. Luyendijk, B. Frangione, Mutation of the alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, dutch type, Science 248 (4959) (1990) 1124–1126. doi:10.1126/science.2111584.
URL http://science.sciencemag.org/content/248/4959/1124

[18] R. Sherrington, E. I. Rogaev, Y. Liang, E. A. Rogaeva, G. Levesque, M. Ikeda, H. Chi, C. Lin, G. Li, K. Holman, T. Tsuda, L. Mar, J. F. Foncin, A. C. Bruni, M. P. Montesi, S. Sorbi, I. Rainero, L. Pinessi, L. Nee, I. Chumakov, D. Pollen, A. Brookes, P. Sanseau, R. J. Polinsky, W. Wasco, H. A. R. Da Silva, J. L. Haines, M. A. Pericak-Vance, R. E. Tanzi, A. D. Roses, P. E. Fraser, J. M. Rommens, P. H. St George-Hyslop, Cloning of a

gene bearing missense mutations in early-onset familial alzheimer's disease, Nature 375 (1995) 754. doi:10.1038/375754a0.
URL http://dx.doi.org/10.1038/375754a0

[19] E. Levy-Lahad, W. Wasco, P. Poorkaj, D. Romano, J. Oshima, W. Pettingell, C. Yu, P. Jondro, S. Schmidt, K. Wang, e. al., Candidate gene for the chromosome 1 familial alzheimer's disease locus, Science 269 (5226) (1995) 973–977. doi:10.1126/science.7638622.
URL http://science.sciencemag.org/content/269/5226/973

[20] R. J. Bateman, C. Xiong, T. L. Benzinger, A. M. Fagan, A. Goate, N. C. Fox, D. S. Marcus, N. J. Cairns, X. Xie, T. M. Blazey, D. M. Holtzman, A. Santacruz, V. Buckles, A. Oliver, K. Moulder, P. S. Aisen, B. Ghetti, W. E. Klunk, E. McDade, R. N. Martins, C. L. Masters, R. Mayeux, J. M. Ringman, M. N. Rossor, P. R. Schofield, R. A. Sperling, S. Salloway, J. C. Morris, Clinical and biomarker changes in dominantly inherited alzheimer's disease, New England Journal of Medicine 367 (9) (2012) 795–804. doi:10.1056/NEJMoa1202753.
URL https://doi.org/10.1056/NEJMoa1202753

[21] E. M. Reiman, Y. T. Quiroz, A. S. Fleisher, K. Chen, C. Velez-Pardo, M. Jimenez-Del-Rio, A. M. Fagan, A. R. Shah, S. Alvarez, A. Arbelaez, M. Giraldo, N. Acosta-Baena, R. A. Sperling, B. Dickerson, C. E. Stern, V. Tirado, C. Munoz, R. A. Reiman, M. J. Huentelman, G. E. Alexander, J. B. Langbaum, K. S. Kosik, P. N. Tariot, F. Lopera, Brain imaging and fluid biomarker analysis in young adults at genetic risk for autosomal dominant alzheimer's disease in the presenilin 1 e280a kindred: a case-control study, The Lancet Neurology 11 (12) (2012) 1048–1056. doi:https://doi.org/10.1016/S1474-4422(12)70228-4.
URL http://www.sciencedirect.com/science/article/pii/S1474442212702284

[22] T. JB, B. MR, B. RJ, et al, Functional connectivity in autosomal dominant

and late-onset alzheimer's disease, JAMA Neurology 71 (9) (2014) 1111–1122. doi:10.1001/jamaneurol.2014.1654.
URL http://dx.doi.org/10.1001/jamaneurol.2014.1654

[23] M. Tang, D. C. Ryman, E. McDade, M. S. Jasielec, V. D. Buckles, N. J. Cairns, A. M. Fagan, A. Goate, D. S. Marcus, C. Xiong, R. F. Allegri, J. P. Chhatwal, A. Danek, M. R. Farlow, N. C. Fox, B. Ghetti, N. R. Graff-Radford, C. Laske, R. N. Martins, C. L. Masters, R. P. Mayeux, J. M. Ringman, M. N. Rossor, S. P. Salloway, P. R. Schofield, J. C. Morris, R. J. Bateman, Neurological manifestations of autosomal dominant familial alzheimer's disease: a comparison of the published literature with the dominantly inherited alzheimer network observational study (dian-obs), The Lancet Neurology 15 (13) (2016) 1317–1325. doi:https://doi.org/10.1016/S1474-4422(16)30229-0.
URL http://www.sciencedirect.com/science/article/pii/S1474442216302290

[24] R. J. Bateman, T. L. Benzinger, S. Berry, D. B. Clifford, C. Duggan, A. M. Fagan, K. Fanning, M. R. Farlow, J. Hassenstab, E. M. McDade, S. Mills, K. Paumier, M. Quintana, S. P. Salloway, A. Santacruz, L. S. Schneider, G. Wang, C. Xiong, The dian-tu next generation alzheimer's prevention trial: Adaptive design and disease progression model, Alzheimer's & Dementia 13 (1) (2017) 8–19. doi:https://doi.org/10.1016/j.jalz.2016.07.005.
URL http://www.sciencedirect.com/science/article/pii/S1552526016300486

[25] D. C. Ryman, N. Acosta-Baena, P. S. Aisen, T. Bird, A. Danek, N. C. Fox, A. Goate, P. Frommelt, B. Ghetti, J. B. Langbaum, F. Lopera, R. Martins, C. L. Masters, R. P. Mayeux, E. McDade, S. Moreno, E. M. Reiman, J. M. Ringman, S. Salloway, P. R. Schofield, R. Sperling, P. N. Tariot, C. Xiong, J. C. Morris, R. J. Bateman, , Symptom onset in autosomal dominant alzheimer disease, Neurology 83 (3) (2014) 253–260. doi:10.1212/WNL.

0000000000000596.
URL http://n.neurology.org/content/83/3/253

[26] J. C. Morris, P. S. Aisen, R. J. Bateman, T. L. S. Benzinger, N. J. Cairns, A. M. Fagan, B. Ghetti, A. M. Goate, D. M. Holtzman, W. E. Klunk, E. McDade, D. S. Marcus, R. N. Martins, C. L. Masters, R. Mayeux, A. Oliver, K. Quaid, J. M. Ringman, M. N. Rossor, S. Salloway, P. R. Schofield, N. J. Selsor, R. A. Sperling, M. W. Weiner, C. Xiong, K. L. Moulder, V. D. Buckles, Developing an international network for alzheimer research: The dominantly inherited alzheimer network, Clinical investigation 2 (10) (2012) 975–984. doi:10.4155/cli.12.93.

[27] J. M. Ringman, L.-J. Liang, Y. Zhou, S. Vangala, E. Teng, S. Kremen, D. Wharton, A. Goate, D. S. Marcus, M. Farlow, B. Ghetti, E. McDade, C. L. Masters, R. P. Mayeux, M. Rossor, S. Salloway, P. R. Schofield, J. L. Cummings, V. Buckles, R. Bateman, J. C. Morris, the Dominantly Inherited Alzheimer Network, Early behavioural changes in familial alzheimer's disease in the dominantly inherited alzheimer network, Brain 138 (4) (2015) 1036–1045. doi:10.1093/brain/awv004.
URL http://dx.doi.org/10.1093/brain/awv004

[28] Y. Quiroz, A. Schultz, K. Chen, et al, Brain imaging and blood biomarker abnormalities in children with autosomal dominant alzheimer disease: A cross-sectional study, JAMA Neurology 72 (8) (2015) 912–919. doi:10.1001/jamaneurol.2015.1099.
URL http://dx.doi.org/10.1001/jamaneurol.2015.1099

[29] S. Mills, J. Mallmann, A. Santacruz, A. Fuqua, M. Carril, P. Aisen, M. Althage, S. Belyew, T. Benzinger, W. Brooks, V. Buckles, N. Cairns, D. Clifford, A. Danek, A. Fagan, M. Farlow, N. Fox, B. Ghetti, A. Goate, D. Heinrichs, R. Hornbeck, C. Jack, M. Jucker, W. Klunk, D. Marcus, R. Martins, C. Masters, R. Mayeux, E. McDade, J. Morris, A. Oliver, J. Ringman, M. Rossor, S. Salloway, P. Schofield, J. Snider,

P. Snyder, R. Sperling, C. Stewart, R. Thomas, C. Xiong, R. Bateman, Preclinical trials in autosomal dominant ad: Implementation of the dian-tu trial, Revue Neurologique 169 (10) (2013) 737–743. doi:https://doi.org/10.1016/j.neurol.2013.07.017.
URL http://www.sciencedirect.com/science/article/pii/S003537871300876X

[30] K. L. Moulder, B. J. Snider, S. L. Mills, V. D. Buckles, A. M. Santacruz, R. J. Bateman, J. C. Morris, Dominantly inherited alzheimer network: facilitating research and clinical trials, Alzheimer's Research & Therapy 5 (5) (2013) 48. doi:10.1186/alzrt213.
URL https://doi.org/10.1186/alzrt213

[31] Y. Su, T. M. Blazey, C. J. Owen, J. J. Christensen, K. Friedrichsen, N. Joseph-Mathurin, Q. Wang, R. C. Hornbeck, B. M. Ances, A. Z. Snyder, L. A. Cash, R. A. Koeppe, W. E. Klunk, D. Galasko, A. M. Brickman, E. McDade, J. M. Ringman, P. M. Thompson, A. J. Saykin, B. Ghetti, R. A. Sperling, K. A. Johnson, S. P. Salloway, P. R. Schofield, C. L. Masters, V. L. Villemagne, N. C. Fox, S. Frster, K. Chen, E. M. Reiman, C. Xiong, D. S. Marcus, M. W. Weiner, J. C. Morris, R. J. Bateman, T. L. S. Benzinger, D. I. A. Network, Quantitative amyloid imaging in autosomal dominant alzheimer's disease: Results from the dian study group, PLOS ONE 11 (3) (2016) 1–14. doi:10.1371/journal.pone.0152082.
URL https://doi.org/10.1371/journal.pone.0152082

[32] R. J. Bateman, P. S. Aisen, B. D. Strooper, N. C. Fox, C. A. Lemere, J. M. Ringman, S. Salloway, R. A. Sperling, M. Windisch, C. Xiong, Autosomal-dominant alzheimer's disease: a review and proposal for the prevention of alzheimer's disease, Alzheimer's Research & Therapy 3 (1) (2010) 1. doi:10.1186/alzrt59.

[33] M. A. Lindquist, B. Caffo, C. Crainiceanu, Ironing out the statistical wrin-

kles in ten ironic rules, NeuroImage 81 (2013) 499–502. doi:10.1016/j.neuroimage.2013.02.056.

[34] K. Sakai, K. Yamada, Machine learning studies on major brain diseases: 5-year trends of 2014-2018, Japanese Journal of Radiology 37 (1) (2018) 34–72. doi:10.1007/s11604-018-0794-4.

[35] C. R. Jack, D. A. Bennett, K. Blennow, M. C. Carrillo, B. Dunn, S. B. Haeberlein, D. M. Holtzman, W. Jagust, F. Jessen, J. Karlawish, E. Liu, J. L. Molinuevo, T. Montine, C. Phelps, K. P. Rankin, C. C. Rowe, P. Scheltens, E. Siemers, H. M. Snyder, R. Sperling, C. Elliott, E. Masliah, L. Ryan, N. Silverberg, NIA-AA research framework: Toward a biological definition of alzheimer's disease, Alzheimer's & Dementia 14 (4) (2018) 535–562. doi:10.1016/j.jalz.2018.02.018.

[36] C. Cruchaga, S. Chakraverty, K. Mayo, F. L. M. Vallania, R. D. Mitra, K. Faber, J. Williamson, T. Bird, R. Diaz-Arrastia, T. M. Foroud, B. F. Boeve, N. R. Graff-Radford, P. St. Jean, M. Lawson, M. G. Ehm, R. Mayeux, A. M. Goate, for the NIA-LOAD/NCRAD Family Study Consortium, Rare variants in app, psen1 and psen2 increase risk for ad in late-onset alzheimer's disease families, PLoS ONE 7 (2). doi:10.1371/journal.pone.0031039.
URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3270040/

[37] A. S. Fleisher, K. Chen, Y. T. Quiroz, L. J. Jakimovich, M. G. Gomez, C. M. Langois, J. B. Langbaum, N. Ayutyanont, A. Roontiva, P. Thiyyagura, W. Lee, H. Mo, L. Lopez, S. Moreno, N. Acosta-Baena, M. Giraldo, G. Garcia, R. A. Reiman, M. J. Huentelman, K. S. Kosik, P. N. Tariot, F. Lopera, E. M. Reiman, Florbetapir pet analysis of amyloid-beta deposition in the presenilin 1 e280a autosomal dominant alzheimer's disease kindred: a cross-sectional study, The Lancet Neurology 11 (12) (2012) 1057–1065. doi:10.1016/S1474-4422(12)70227-2.
URL http://dx.doi.org/10.1016/S1474-4422(12)70227-2

[38] K. G. Mawuenyega, W. Sigurdson, V. Ovod, L. Munsell, T. Kasten, J. C. Morris, K. E. Yarasheski, R. J. Bateman, Decreased clearance of cns - amyloid in alzheimer's disease, Science 330 (6012) (2010) 1774–1774. doi:10.1126/science.1197623.
URL http://science.sciencemag.org/content/330/6012/1774

[39] M. Pera, D. Alcolea, R. Sánchez-Valle, C. Guardia-Laguarta, M. Colom-Cadena, N. Badiola, M. Suárez-Calvet, A. Lladó, A. A. Barrera-Ocampo, D. Sepulveda-Falla, R. Blesa, J. L. Molinuevo, J. Clarimón, I. Ferrer, E. Gelpi, A. Lleó, Distinct patterns of app processing in the cns in autosomal-dominant and sporadic alzheimer disease, Acta Neuropathologica 125 (2) (2013) 201–213. doi:10.1007/s00401-012-1062-9.
URL https://doi.org/10.1007/s00401-012-1062-9

[40] L. Yan, C. Y. Liu, K.-P. Wong, S.-C. Huang, W. J. Mack, K. Jann, G. Coppola, J. M. Ringman, D. J. Wang, Regional association of pcasl-mri with fdg-pet and pib-pet in people at risk for autosomal dominant alzheimer's disease, NeuroImage: Clinical 17 (2018) 751–760. doi:https://doi.org/10.1016/j.nicl.2017.12.003.
URL http://www.sciencedirect.com/science/article/pii/S221315821730308X

[41] M. Reuter, N. J. Schmansky, H. D. Rosas, B. Fischl, Within-subject template estimation for unbiased longitudinal image analysis, NeuroImage 61 (4) (2012) 1402–1418. doi:https://doi.org/10.1016/j.neuroimage.2012.02.084.
URL http://www.sciencedirect.com/science/article/pii/S1053811912002765

[42] A. M. Dale, M. I. Sereno, Improved localizadon of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach, Journal of Cognitive Neuroscience 5 (2) (1993) 162–176. doi:10.1162/jocn.1993.5.2.162.

[43] A. M. Dale, B. Fischl, M. I. Sereno, Cortical Surface-Based Analysis, NeuroImage 9 (2) (1999) 179–194. doi:10.1006/nimg.1998.0395.

[44] B. Fischl, M. I. Sereno, A. M. Dale, Cortical surface-based analysis, NeuroImage 9 (2) (1999) 195–207. doi:10.1006/nimg.1998.0396.

[45] B. Fischl, M. I. Sereno, R. B. Tootell, A. M. Dale, High-resolution intersubject averaging and a coordinate system for the cortical surface, Human Brain Mapping 8 (4) (1999) 272–284. doi:10.1002/(sici)1097-0193(1999)8:4<272::aid-hbm10>3.0.co;2-4.

[46] B. Fischl, A. M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images, Proceedings of the National Academy of Sciences 97 (20) (2000) 11050–11055. doi:10.1073/pnas.200033797.

[47] B. Fischl, A. Liu, A. Dale, Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex, IEEE Transactions on Medical Imaging 20 (1) (2001) 70–80. doi:10.1109/42.906426.

[48] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A. M. Dale, Whole brain segmentation, Neuron 33 (3) (2002) 341–355. doi:10.1016/s0896-6273(02)00569-x.

[49] B. Fischl, Automatically parcellating the human cerebral cortex, Cerebral Cortex 14 (1) (2004) 11–22. doi:10.1093/cercor/bhg087.

[50] B. Fischl, D. H. Salat, A. J. van der Kouwe, N. Makris, F. Ségonne, B. T. Quinn, A. M. Dale, Sequence-independent segmentation of magnetic resonance images, NeuroImage 23 (2004) S69–S84. doi:10.1016/j.neuroimage.2004.07.016.

[51] F. Ségonne, A. Dale, E. Busa, M. Glessner, D. Salat, H. Hahn, B. Fischl, A hybrid approach to the skull stripping problem in MRI, NeuroImage 22 (3) (2004) 1060–1075. doi:10.1016/j.neuroimage.2004.03.032.

[52] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany, P. Maguire, D. Rosas, N. Makris, A. Dale, B. Dickerson, B. Fischl, Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer, NeuroImage 32 (1) (2006) 180–194. doi:10.1016/j.neuroimage.2006.02.051.

[53] J. Jovicich, S. Czanner, D. Greve, E. Haley, A. van der Kouwe, R. Gollub, D. Kennedy, F. Schmitt, G. Brown, J. MacFall, B. Fischl, A. Dale, Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data, NeuroImage 30 (2) (2006) 436–443. doi:10.1016/j.neuroimage.2005.09.046.

[54] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, R. J. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, NeuroImage 31 (3) (2006) 968–980. doi:10.1016/j.neuroimage.2006.01.021.

[55] B. Fischl, FreeSurfer, NeuroImage 62 (2) (2012) 774–781. doi:10.1016/j.neuroimage.2012.01.021.

[56] L. Mosconi, Brain glucose metabolism in the early and specific diagnosis of alzheimer's disease, European Journal of Nuclear Medicine and Molecular Imaging 32 (4) (2005) 486–510. doi:10.1007/s00259-005-1762-7.
URL https://doi.org/10.1007/s00259-005-1762-7

[57] J. J. Gomar, C. Conejero-Goldberg, P. Davies, T. E. Goldberg, Extension and refinement of the predictive value of different classes of markers in adni: Four-year follow-up data, Alzheimer's & Dementia 10 (6) (2014) 704–712. doi:https://doi.org/10.1016/j.jalz.2013.11.009.
URL http://www.sciencedirect.com/science/article/pii/S1552526014000090

[58] N. Smailagic, M. Vacante, C. Hyde, S. Martin, O. Ukoumunne, C. Sachpekidis, 18f-fdg pet for the early diagnosis of alzheimer's disease dementia and other dementias in people with mild cognitive impairment (mci), The Cochrane Library.

[59] D. J. Rowland, J. R. Garbow, R. Laforest, A. Z. Snyder, Registration of [18f]FDG microPET and small-animal MRI, Nuclear Medicine and Biology 32 (6) (2005) 567–572. doi:10.1016/j.nucmedbio.2005.05.002.

[60] S. A. Eisenstein, J. M. Koller, M. Piccirillo, A. Kim, J. A. V. Antenor-Dorsey, T. O. Videen, A. Z. Snyder, M. Karimi, S. M. Moerlein, K. J. Black, J. S. Perlmutter, T. Hershey, Characterization of extrastriatal d2 in vivo specific binding of [18f](n-methyl)benperidol using PET, Synapse 66 (9) (2012) 770–780. doi:10.1002/syn.21566.

[61] T. Benzinger, A. Fagan, J. Chhatwal, J. C. Morris, M. Rossor, R. Bateman, Regional variability of imaging biomarkers in dian, Alzheimer's & Dementia 10 (4) (2014) 127. doi:10.1016/j.jalz.2014.04.054.

[62] N. Franzmeier, J. Ren, R. J. Bateman, J. C. Morris, J. Levin, M. Jucker, T. L. Benzinger, I. Yakushev, N. Koutsouleris, M. Ewers, CROSS-VALIDATED BIOMARKER-BASED PREDICTION OF 4-YEAR RATE OF COGNITIVE DECLINE IN NON-DEMENTED SUBJECTS AT RISK OF AD, Alzheimer's & Dementia 14 (7) (2018) P75. doi:10.1016/j.jalz.2018.06.2154.

[63] V. Frouin, C. Comtat, A. Reilhac, M.-C. Grgoire, Correction of partial-volume effect for pet striatal imaging: Fast implementation and study of robustness, Journal of Nuclear Medicine 43 (2002) 1715–1726.

[64] O. Rousset, A. Rahmim, A. Alavi, H. Zaidi, Partial volume correction strategies in PET, PET Clinics 2 (2) (2007) 235–249. doi:10.1016/j.cpet.2007.10.005.

[65] O. G. Rousset, Y. Ma, A. C. Evans, Correction for partial volume effects in pet: principle and validation., Journal of nuclear medicine : official publication, Society of Nuclear Medicine 39 (1998) 904–911.

[66] L. Berg, Clinical dementia rating (cdr), Psychopharmacology Bulletin 24 (1988) 637–639.

[67] J. C. Morris, Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the alzheimer type, International Psychogeriatrics 9 (S1) (1997) 173–176. doi:10.1017/S1041610297004870.

[68] C. Laske, H. R. Sohrabi, M. S. Jasielec, S. Mller, N. K. Koehler, S. Grber, S. Frster, A. Drzezga, F. Mueller-Sarnowski, A. Danek, M. Jucker, R. J. Bateman, V. Buckles, A. J. Saykin, R. N. Martins, J. C. Morris, D. I. A. N. (DIAN), Diagnostic value of subjective memory complaints assessed with a single item in dominantly inherited alzheimer's disease: Results of the dian study, BioMed Research International (2015) 7doi:10.1155/2015/828120. URL http://dx.doi.org/10.1155/2015/828120

[69] C. A. Lemere, F. Lopera, K. S. Kosik, C. L. Lendon, J. Ossa, T. C. Saido, H. Yamaguchi, A. Ruiz, A. Martinez, L. Madrigal, L. Hincapie, J. C. A. L., D. C. Anthony, E. H. Koo, A. M. Goate, D. J. Selkoe, J. C. A. V., The e280a presenilin 1 alzheimer mutation produces increased aβ42 deposition and severe cerebellar pathology, Nature Medicine 2 (10) (1996) 1146–1150. doi:10.1038/nm1096-1146.

[70] A. B. Graf, S. Borer, Normalization in support vector machines (2001) 277–282doi:10.1007/3-540-45404-7_37.

[71] G. Heiman, Research Methods in Statistics, 2002.

[72] J. Ramírez, J. Górriz, A. Ortiz, F. Martinez-Murcia, F. Segovia, D. Salas-Gonzalez, D. Castillo-Barnes, I. Illan, C. Puntonet, Ensemble of random forests one vs. rest classifiers for mci and ad prediction using anova cortical and subcortical feature selection and partial

least squares, Journal of Neuroscience Methods 302 (2018) 47–57. doi:https://doi.org/10.1016/j.jneumeth.2017.12.005.
URL http://www.sciencedirect.com/science/article/pii/S0165027017304223

[73] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometrics and Intelligent Laboratory Systems 2 (1) (1987) 37–52. doi:https://doi.org/10.1016/0169-7439(87)80084-9.
URL http://www.sciencedirect.com/science/article/pii/0169743987800849

[74] P. Markiewicz, J. Matthews, J. Declerck, K. Herholz, Robustness of multivariate image analysis assessed by resampling techniques and applied to FDG-PET scans of patients with alzheimer's disease, NeuroImage 46 (2) (2009) 472–485. doi:10.1016/j.neuroimage.2009.01.020.

[75] M. López, J. Ramírez, J. Górriz, I. Illán, D. Salas-Gonzalez, F. Segovia, R. Chaves, Svm-based cad system for early detection of the alzheimer's disease using kernel pca and lda, Neuroscience Letters 464 (3) (2009) 233–238. doi:https://doi.org/10.1016/j.neulet.2009.08.061.
URL http://www.sciencedirect.com/science/article/pii/S0304394009011677

[76] I. Illán, J. Górriz, J. Ramírez, D. Salas-Gonzalez, M. López, F. Segovia, R. Chaves, M. Gómez-Rio, C. Puntonet, 18f-fdg pet imaging analysis for computer aided alzheimer's diagnosis, Information Sciences 181 (4) (2011) 903–916. doi:https://doi.org/10.1016/j.ins.2010.10.027.
URL http://www.sciencedirect.com/science/article/pii/S0020025510005323

[77] F. Segovia, C. Bastin, E. Salmon, J. M. Górriz, J. Ramírez, C. Phillips, Combining pet images and neuropsychological test data for automatic diagnosis of alzheimer's disease, PLOS ONE 9 (2) (2014) 1–8. doi:https:

//doi.org/10.1371/journal.pone.0088687.
URL https://doi.org/10.1371/journal.pone.0088687

[78] V. N. Vapnik, Statistical Learning Theory, Ed. 1, John Wiley and Sons, New York, USA, 1998.

[79] T. Joachims, Text categorization with support vector machines: Learning with many relevant features, in: Machine Learning: ECML-98, Springer Berlin Heidelberg, 1998, pp. 137–142. doi:10.1007/bfb0026683.

[80] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection., in: IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence, Vol. 2 of IJCAI'95, 1995, pp. 1137–1145.
URL http://dl.acm.org/citation.cfm?id=1643031.1643047

[81] V. Vapnik, Estimation of Dependences Based on Empirical Data, no. 1, Springer New York, 2006. doi:10.1007/0-387-34239-7.

[82] J. M. Górriz, J. Ramirez, J. Suckling, On the computation of distribution-free performance bounds: Application to small sample sizes in neuroimaging, Pattern Recognition 93 (2019) 1 – 13. doi:https://doi.org/10.1016/j.patcog.2019.03.032.
URL http://www.sciencedirect.com/science/article/pii/S0031320319301402

[83] Kolmogorov–smirnov test, in: The Concise Encyclopedia of Statistics, Springer New York, pp. 283–287. doi:10.1007/978-0-387-32833-1_214.

[84] J. Ramírez, J. Górriz, D. Salas-Gonzalez, A. Romero, M. López, I. Álvarez, M. Gómez-Río, Computer-aided diagnosis of alzheimer's type dementia combining support vector machines and discriminant set of features, Information Sciences 237 (2013) 59–72. doi:10.1016/j.ins.2009.05.012.

[85] H. Akaike, A new look at the statistical model identification, IEEE Transactions on Automatic Control 19 (6) (1974) 716–723. doi:10.1109/TAC.1974.1100705.

[86] D. M. Cash, G. R. Ridgway, Y. Liang, N. S. Ryan, K. M. Kinnunen, T. Yeatman, I. B. Malone, T. L. Benzinger, C. R. Jack, P. M. Thompson, B. F. Ghetti, A. J. Saykin, C. L. Masters, J. M. Ringman, S. P. Salloway, P. R. Schofield, R. A. Sperling, N. J. Cairns, D. S. Marcus, C. Xiong, R. J. Bateman, J. C. Morris, M. N. Rossor, S. Ourselin, N. C. Fox, The pattern of atrophy in familial alzheimer's disease, Neurology 81 (16) (2013) 1425–1433. doi:10.1212/WNL.0b013e3182a841c6.
URL http://n.neurology.org/content/81/16/1425

[87] B. Dubois, H. Hampel, H. H. Feldman, P. Scheltens, P. Aisen, S. Andrieu, H. Bakardjian, H. Benali, L. Bertram, K. Blennow, K. Broich, E. Cavedo, S. Crutch, J.-F. Dartigues, C. Duyckaerts, S. Epelbaum, G. B. Frisoni, S. Gauthier, R. Genthon, A. A. Gouw, M.-O. Habert, D. M. Holtzman, M. Kivipelto, S. Lista, J.-L. Molinuevo, S. E. O'Bryant, G. D. Rabinovici, C. Rowe, S. Salloway, L. S. Schneider, R. Sperling, M. Teichmann, M. C. Carrillo, J. Cummings, C. R. Jack, Preclinical alzheimer's disease: Definition, natural history, and diagnostic criteria, Alzheimer's & Dementia 12 (3) (2016) 292–323. doi:https://doi.org/10.1016/j.jalz.2016.02.002.
URL http://www.sciencedirect.com/science/article/pii/S1552526016000509

[88] J. M. Ringman, J. O'Neill, D. Geschwind, L. Medina, L. G. Apostolova, Y. Rodriguez, B. Schaffer, A. Varpetian, B. Tseng, F. Ortiz, J. Fitten, J. L. Cummings, G. Bartzokis, Diffusion tensor imaging in preclinical and presymptomatic carriers of familial alzheimer's disease mutations, Brain 130 (7) (2007) 1767–1776. doi:10.1093/brain/awm102.

[89] T. L. S. Benzinger, T. Blazey, C. R. Jack, R. A. Koeppe, Y. Su, C. Xiong, M. E. Raichle, A. Z. Snyder, B. M. Ances, R. J. Bateman, N. J. Cairns,

A. M. Fagan, A. Goate, D. S. Marcus, P. S. Aisen, J. J. Christensen, L. Ercole, R. C. Hornbeck, A. M. Farrar, P. Aldea, M. S. Jasielec, C. J. Owen, X. Xie, R. Mayeux, A. Brickman, E. McDade, W. Klunk, C. A. Mathis, J. Ringman, P. M. Thompson, B. Ghetti, A. J. Saykin, R. A. Sperling, K. A. Johnson, S. Salloway, S. Correia, P. R. Schofield, C. L. Masters, C. Rowe, V. L. Villemagne, R. Martins, S. Ourselin, M. N. Rossor, N. C. Fox, D. M. Cash, M. W. Weiner, D. M. Holtzman, V. D. Buckles, K. Moulder, J. C. Morris, Regional variability of imaging biomarkers in autosomal dominant alzheimer's disease, Proceedings of the National Academy of Sciences 110 (47) (2013) 4502–4509. doi:10.1073/pnas.1317918110.
URL http://www.pnas.org/content/110/47/E4502

[90] S. Lee, F. Viqar, M. E. Zimmerman, A. Narkhede, G. Tosto, T. L. Benzinger, D. S. Marcus, A. M. Fagan, A. Goate, N. C. Fox, N. J. Cairns, D. M. Holtzman, V. Buckles, B. Ghetti, E. McDade, R. N. Martins, A. J. Saykin, C. L. Masters, J. M. Ringman, N. S. Ryan, S. Frster, C. Laske, P. R. Schofield, R. A. Sperling, S. Salloway, S. Correia, C. Jack, M. Weiner, R. J. Bateman, J. C. Morris, R. Mayeux, A. M. B. and, White matter hyperintensities are a core feature of alzheimer's disease: Evidence from the dominantly inherited alzheimer network, Annals of Neurology 79 (6) (2016) 929–939. doi:10.1002/ana.24647.

Declaration of interest:

All authors disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work or state if there are no interests to declare.

Prof. JM Gorriz

Author statement:

Diego Castillo-Barnes, Li Su, Javier Ramırez[1], Juan M. G´orriz : all terms

Diego Salas-Gonzalez[1], Francisco J. Martinez-Murcia[3],Ignacio A. Ill´an[1], Fermin Segovia[1], Andres Ortiz[3]: Data Curation, Conceptualization

DIAN Sterring Committee*: Resources, Project Administration; Funding Acquisition;

*it will be specifically defined if the paper is finally accepted (a list of more than 20 authors). The complete list of authors should be completed in collaboration with the Pastora consortium according to their publication policy.